



UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE CIÊNCIAS MÉDICAS

THAÍS FENZ ARAUJO

**AS BASES MOLECULARES DA PICNODISOSTOSE E ANÁLISE DA
DISTRIBUIÇÃO DAS MUTAÇÕES NO ESTADO DO CEARÁ E EM OUTRAS
REGIÕES DO BRASIL**

CAMPINAS

2016

THAÍS FENZ ARAUJO

**AS BASES MOLECULARES DA PICNODISOSTOSE E A ANÁLISE DA
DISTRIBUIÇÃO DAS MUTAÇÕES NO ESTADO DO CEARÁ E EM OUTRAS
REGIÕES DO BRASIL**

Dissertação de mestrado apresentada à Faculdade de Ciências Médicas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de mestra em Ciências Médicas, área de concentração de Ciências Biomédicas.

ORIENTADORA: PROFA. DRA. DENISE PONTES CAVALCANTI

ESTE EXEMPLAR CORRESPONDE À VERSÃO
FINAL DA DISSERTAÇÃO DEFENDIDA PELA
ALUNA THAÍS FENZ ARAUJO E ORIENTADA PELA
PROFA. DRA. DENISE PONTES CAVALCANTI.

CAMPINAS

2016

Agência(s) de fomento e nº(s) de processo(s): CAPES, 3300017023p6

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Ciências Médicas
Maristella Soares dos Santos - CRB 8/8402

Ar15b Araujo, Thaís Fenz, 1988-
As bases moleculares da Pcnodisostose e análise da distribuição das mutações no estado do Ceará e em outras regiões do Brasil / Thaís Fenz Araujo. – Campinas, SP : [s.n.], 2016.

Orientador: Denise Pontes Cavalcanti.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Ciências Médicas.

1. Pcnodisostose. 2. Cathepsina K. 3. Consanguinidade. I. Cavalcanti, Denise Pontes, 1957-. II. Universidade Estadual de Campinas. Faculdade de Ciências Médicas. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: The molecular basis of Pcnodysostosis and the distribution of mutations in state of Ceará and in other regions of Brazil

Palavras-chave em inglês:

Pcnodysostosis

Cathepsin K

Consanguinity

Área de concentração: Ciências Biomédicas

Titulação: Mestra em Ciências Médicas

Banca examinadora:

Denise Pontes Cavalcanti [Orientador]

Louis Bernard Klaczko

Carmen Silvia Bertuzzo

Cláudia Vianna Mauerer Morelli

Debora Romeo Bertola

Data de defesa: 25-02-2016

Programa de Pós-Graduação: Ciências Médicas

BANCA EXAMINADORA DA DEFESA DE MESTRADO

THAÍS FENZ ARAUJO

ORIENTADORA: DENISE PONTES CAVALCANTI

MEMBROS:

1. PROFA. DRA. DENISE PONTES CAVALCANTI

2. PROF. DR. LOUIS BERNARD KLACZKO

3. PROFA. DRA. CARMEN SILVIA BERTUZZO

Programa de Pós-Graduação em Ciências Médicas da Faculdade de Ciências Médicas da Universidade Estadual de Campinas.

A ata de defesa com as respectivas assinaturas dos membros da banca examinadora encontra-se no processo de vida acadêmica do aluno.

DATA DA DEFESA: 25/02/2016

AGRADECIMENTOS

A essência de alguns sentimentos não pode ser traduzida, mesmo que sejam utilizadas as mais belas palavras da forma mais graciosa que nossa língua permitir. A gratidão é um desses sentimentos, que apenas quem vivencia consegue compreender. Espero que todos aqueles citados aqui saibam de sua importância nessa etapa da minha jornada.

Agradeço a meus pais e irmãos por sempre estarem presentes em minha vida, por me apoiarem e me incentivarem a conquistar meus objetivos. Aos familiares e amigos agradeço pelos momentos de alegria e também pelos momentos de dificuldade compartilhados, pois são exatamente nessas horas que temos a oportunidade de superar obstáculos e adquirir experiência.

Em especial agradeço a Profa. Dra. Denise Pontes Cavalcanti, por ter me recebido como aluna e confiado a mim a responsabilidade de estudar e desenvolver um projeto tão interessante, com o qual aprendi muito. Também agradeço pela imensa paciência e pelos inúmeros ensinamentos.

Aos amigos e companheiros de trabalho, agradeço por toda ajuda dentro e fora do laboratório, bem como pelos momentos de descontração compartilhados, em especial agradeço à Luciana Bonadia, Cynthia, Karina, Thatiane, Dora e Carolina pela imensa ajuda e contribuição intelectual e ao Vagner por todo apoio emocional durante esta etapa.

Agradeço aos colaboradores do projeto, pelo esforço em recrutar pacientes, documentos e informações, sem os quais não seria possível a realização desse estudo e aos membros da banca, por todas as sugestões, que auxiliaram na produção final desse manuscrito. Agradeço também aos próprios pacientes, por aceitarem participar do estudo e às agências de fomento, CAPES e CNPq, pelo auxílio financeiro concedido.

RESUMO

A Picnodisostose é uma doença autossômica recessiva rara classificada entre as displasias esqueléticas com aumento da densidade óssea e tendência a fraturas. Além dessas características, fazem parte do quadro: face típica, suturas cranianas e fontanelas abertas, braquidactilia e osteólise das falanges distais. Apesar de ser considerada rara, recentemente mais de 20 famílias, a maioria consanguínea, com pelo menos um afetado por Picnodisostose foram observadas no Estado do Ceará (CE). Essa alta frequência justificou o presente trabalho, cuja investigação molecular teve como principal objetivo testar a hipótese de possível efeito fundador. A investigação se estendeu para mais 15 famílias de outras regiões. Foram investigados 39 indivíduos de 33 famílias, sendo 18 do Ceará. Utilizando o sequenciamento Sanger do gene *CTSK* foram identificadas seis mutações diferentes, cinco previamente descritas -p.G146R, p.G194S, p.R241X, p.A277V, p.C318Y- e uma inédita, p.W29Mfs*10, cuja concentração na região noroeste do CE chamou atenção. A análise das mutações juntamente com dados familiares, origem das famílias e reavaliação do fenótipo permitiram as seguintes conclusões: 1 – A alta taxa de endocruzamento nas famílias explica a alta frequência de Picnodisostose no CE; 2 – A origem das famílias analisadas mostra que a maior parte das mutações observadas nas outras regiões do Brasil derivam de um processo de migração interna da população, sobretudo do fluxo nordeste – sudeste; 3 – O estudo de haplótipos reforçou a hipótese levantada de efeito fundador para a mutação inédita; 4 – A frequência de mutações em homozigose nas famílias do Ceará, maior que a frequência de consanguinidade referida, sugere que a real taxa de consanguinidade dessas famílias é da ordem de 80%; 5 – O fenótipo da Picnodisostose é bastante homogêneo, independentemente do tipo de mutação no gene *CTSK*. **PALAVRAS- CHAVE:** Picnodisostose, Catepsina K, consanguinidade.

ABSTRACT

Pycnodysostosis is a rare autosomal recessive disease classified among the skeletal dysplasias with increase of bone density and tendency to fractures. The following clinical and radiological features are also present in Pycnodysostosis: typical facial appearance, open sutures and fontanelles, brachydactyly, and osteolysis of the distal phalanges. In spite of the rarity of the Pycnodysostosis, recently more than 20 families, most consanguineous, with at least one affected by Pycnodysostosis, were observed in the State of Ceará (CE), in the Brazilian northeast. This cluster prompted us to perform a molecular investigation of the *CTSK* gene in these families, under the work hypothesis of a possible founder effect. This investigation also included 15 families from other Brazilian regions. We have studied 39 individuals from 33 families, of which 18 were from Ceará. The sequencing of the *CTSK* gene by the Sanger method identified six different mutations, being five previously described - G146R, G194S, R241X, A277V, C318Y, and a novel variant - W29Mfs*10. This latter appears to be concentrated in the northwest of the Ceará. The molecular analysis associated with the investigation of the origin of the maternal and paternal families, as well as the phenotype review, allowed us to the following conclusions. 1 – The high frequency of Pycnodysostosis in the Ceará is a consequence of the high inbreeding of the families in that region. 2 – The analysis of the origin of the families shows that the most of the found mutations in the other Brazilian regions should come from Northeast by a process of intern migration of the population. 3 – The haplotype study reinforces the hypothesis of founder effect related to the novel mutation. 4 – The high frequency of homozygous mutations among the patients from Ceará (close to 80%) suggests that the real inbreeding should be much higher than the referred consanguinity (33,3%). 5 – The Pycnodysostosis phenotype is homogeneous, despite of the type of mutation in *CTSK* gene. **KEYWORDS:** Pycnodysostosis, *CTSK*, inbreeding.

LISTA DE ILUSTRAÇÕES

Figura 1. Estrutura do gene <i>CTSK</i>	19
Figura 2. Distribuição das mutações encontradas nos pacientes provenientes do Ceará (1 a 18)	31
Figura 3. Distribuição das mutações encontradas nos provenientes de outras localidades (19 a 33)	33
Figura 4. Eletroferograma mostrando a duplicação em homozigose da base timina no éxon 2 (c.83dupT)	36
Figura 5. Distribuição de alguns marcadores no cromossomo 1 de acordo com as distâncias gênica (<i>Genethon</i>) e física (<i>Ensembl</i>).....	38
Figura 6. Haplótipo dos nove pacientes portadores da mutação inédita (p.W29Mfs*10). ..	39

LISTA DE TABELAS

Tabela 1. Resumo das mutações encontradas no gene <i>CTSK</i> em pacientes com Picnodisostose, descritas em artigos publicados até Agosto de 2015.....	20
Tabela 2. Dados clínicos e moleculares das famílias estudadas.....	30
Tabela 3. Procedência e origem dos pacientes estudados.....	34
Tabela 4. Combinação das mutações encontradas nos probandos das 33 famílias estudadas.....	35
Tabela 5. Comprimento dos marcadores nos alelos com a mutação inédita e nos demais alelos.	40
Tabela 6. Revisão fenotípica dos pacientes estudados	41
Apêndice 1:Tabela 7. Primers construídos para estudo molecular do gene <i>CTSK</i>.....	61
Apêndice 2. Tabela 8.Primers referentes aos marcadores utilizados no estudo de haplótipos.....	62

LISTA DE ABREVIATURAS E SIGLAS

A: Aminoácido Alanina

AL: Alagoas

BA: Bahia

c.: Localização no cDNA (número da base nitrogenada)

C: Aminoácido Cisteína

cDNA: DNA complementar, sintetizado a partir de uma RNAm, portanto, sem regiões intrônicas

CTSK: Catepsina K

D: Aminoácido Ácido aspártico

dbSNP (*Single Nucleotide Polymorphism Database*): banco de dados

Del: deleção

DNA: Ácido desoxirribonucleico

dNTP: Desoxirribonucleotídeos fosfatados

DP: Desvio Padrão

Dup: duplicação

E: Aminoácido Glutamato

EDTA (*Ethylenediamine tetraacetic acid*): agente quelante

F: Aminoácido Fenilalanina

FCM: Faculdade de Ciências Médicas

fs* (*frameshift mutation*): Mutação de erro de matriz de leitura

G: Aminoácido Glicina

GAGs: Glisossaminoglicanos

H: Aminoácido Histidina

HGMD (*Human Gene Mutation Database*): banco de dados

HGVS (*Human Genome Variation Society*)

HIAS: Hospital Infantil Albert Sabin

I: Aminoácido Isoleucina

Ins: Inserção

K: Aminoácido Lisina

Kb: Kilobases (unidade de comprimento do fragmento de DNA, equivalente a mil pares de bases)

L: Aminoácido Leucina

M: Aminoácido Metionina

MA: Maranhão

MG: Minas Gerais

MgCl₂: Cloreto de Magnésio

Mitf (*Microphthalmia associated transcription factor*): Fator de transcrição associado à microfalmia

N: Aminoácido Asparagina

NM (*RefSeq Human_Messenger*): Número referência do transcrito

NP (*RefSeq Human_Protein*): Número referência da proteína humana sequenciada

OCD: Osteocondrodisplasia

OMIM (*Online Mendelian Inheritance in Man*): banco de dados online de doenças mendelianas

p.: Localização na proteína (número do aminoácido)

P: Aminoácido Prolina

pb: pares de base (unidade de comprimento do fragmento de DNA)

PB: Paraíba

PCR (*Polymerase Chain Reaction*): Reação em cadeia da polimerase

pH: potencial hidrogeniônico, indica acidez

PR: Paraná

Q: Aminoácido Glutamina

R: Aminoácido Arginina

RN: Rio Grande do Norte

RNA_m: Ácido ribonucleico mensageiro

RS: Rio Grande do Sul

S: Aminoácido Serina

SP: São Paulo

SPARC (*secreted protein acidic and rich in cysteine*): Glicoproteína de ligação

T: Aminoácido Treonina

Taq (*Thermus aquaticus*): espécie de bactéria capaz de tolerar altas temperaturas

TCLE: Termo de Consentimento Livre e Esclarecido

UNICAMP: Universidade Estadual de Campinas

V: Aminoácido Valina

W: Aminoácido Triptofano

X: códon de terminação da tradução

Y: Aminoácido Tirosina

SUMÁRIO

1.	INTRODUÇÃO	14
1.1	BREVE REVISÃO DA LITERATURA	15
1.1.1	A Picnodisostose	15
1.1.2	Do gene à proteína.....	16
1.1.3	Mutações descritas na literatura	18
1.1.4	A prevalência da Picnodisostose	22
1.2	DADOS LOCAIS.....	22
2.	OBJETIVOS	23
2.1	GERAL	23
2.2	ESPECÍFICOS	23
3.	CASUÍSTICA, MATERIAIS E MÉTODOS	24
3.1	CASUÍSTICA.....	24
3.2	MÉTODOS	24
3.2.1	Desenho e Padronização dos <i>Primers</i>	24
3.2.2	Obtenção do material biológico.....	25
3.2.3	Amplificação do DNA por reação em cadeia da Polimerase (PCR).....	25
3.2.4	Sequenciamento direto do gene <i>CTSK</i> pelo método de Sanger	26
3.2.5	Análise do Sequenciamento	26
3.2.7	Revisão de Fenótipos	27
4.	RESULTADOS	28
4.1	DESCRIÇÃO DA CASUÍSTICA	28
4.2	MUTAÇÕES ENCONTRADAS.....	28
4.3	Combinação das mutações encontradas em todos os probandos da casuística	35
4.4	Breve caracterização da mutação inédita p.W29Mfs*10 (c.83dupT)	36
4.5	Breve caracterização das mutações já descritas na literatura	36

4.6	Análise de Haplótipos	37
4.7	Revisão de Fenótipos	40
4.8	Produção Científica	43
5.	DISCUSSÃO	44
5.1	A Prevalência da Picnodisostose é subestimada?	44
5.2	A Picnodisostose no Ceará.....	44
5.3	Picnodisostose em outras regiões do Brasil	46
5.4	Mutações encontradas e sua origem geográfica.....	47
5.5	Análise de haplótipos ligados à mutação inédita.....	49
5.6	Análise fenotípica	51
7.	CONCLUSÕES.....	53
8.	REFERÊNCIAS BIBLIOGRÁFICAS	54
9.	APÊNDICES	61
10.	ANEXOS	85
	Anexo 1: Pareceres dos Comitês de Ética.....	85
	Anexo 2: Modelo do Termo de Consentimento Livre e Esclarecido utilizado.....	89

INTRODUÇÃO

As displasias esqueléticas ou OsteoCondroDisplasias (OCD) são entendidas como um grupo heterogêneo de doenças genéticas que afetam a formação e desenvolvimento de ossos e cartilagens, cuja prevalência tem sido estimada em torno de 3 a cada 10.000 nascimentos [1, 2].

Entre as mais de 350 OCD conhecidas, encontra-se a Picnodisostose (OMIM 265800), uma doença autossômica recessiva rara pertencente ao grupo das displasias com aumento da densidade óssea e, portanto, com tendência a fraturas [3].

No escopo de um projeto de pesquisa sobre diagnóstico das OCD, em outubro de 2013, foi organizada uma visita ao Hospital Infantil Albert Sabin (Fortaleza, CE) para discussão dos casos locais de displasias esqueléticas e nesta ocasião foi observada uma alta frequência de pacientes com Picnodisostose, chamando atenção para a necessidade de uma investigação mais acurada dessa observação.

O objetivo geral desse trabalho foi, portanto, estudar as bases moleculares das famílias com pelo menos um afetado por Picnodisostose no Estado do Ceará, de modo a testar a hipótese de efeito fundador nessa região. Também foi incluída nesse trabalho a investigação de famílias com afetados por Picnodisostose oriundas de outras regiões do território nacional.

1.1 BREVE REVISÃO DA LITERATURA

1.1.1 A Picnodisostose

As displasias esqueléticas ou OsteoCondroDisplasias (OCD) formam um grupo heterogêneo de doenças raras que afetam o crescimento e desenvolvimento de ossos e cartilagens [1]. Geralmente as OCD são decorrentes de mutações monogênicas que produzem anomalias na estrutura de proteínas, outros produtos gênicos e/ou processos metabólicos, que em geral interferem na regulação da placa de crescimento [4]. Cerca de 40% das OCD são diagnosticadas ao nascimento, cuja prevalência estimada é de 3 a cada 10.000 [1]. O restante normalmente é diagnosticado na infância [1].

A última revisão da nosologia das doenças genéticas esqueléticas mantém a classificação dessas condições por meio de critérios clínico-radiológicos e moleculares em 42 grupos ou famílias, dos quais 32 representam as OCDs [3]. Entre um dos grupos de OCD estão relacionadas 20 diferentes condições que apresentam em comum o aumento da densidade dos ossos [3]. Dentre essas, encontra-se a Picnodisostose (OMIM 265800) [3], uma condição autossômica recessiva rara e com distribuição igual entre os sexos [3, 5].

Acredita-se que o primeiro caso de Picnodisostose tenha sido relatado por Montanari em 1923, sob o nome de Acondroplasia Cleido-Crânio-Digital, visto que o autor supunha se tratar de uma acondroplasia atípica com anormalidades na clavícula, crânio e falanges distais (Montanari *apud* 6). Apenas em 1962 Maroteaux e Lamy definiram essa condição como uma nova doença óssea, nomeando-a de *Pycnodysostosis* (do grego: *pycnos* = denso; *dys* = anormal; *osteon* = osso). Com esta definição alguns autores no passado também atribuíram o nome de síndrome de Maroteaux-Lamy para essa condição (Maroteaux *apud* 6).

No ano de 1965 Maroteaux e Lamy sugeriram que o famoso pintor francês Henri de Toulouse-Lautrec (1864-1901) fora afetado por essa doença, iniciando um amplo debate acerca do correto diagnóstico do pintor (Maroteaux *apud* 7). Segundo Maroteaux e Lamy *apud* Hodder [7], a baixa estatura, a fragilidade óssea, a consanguinidade parental e o provável defeito no fechamento das fontanelas, que talvez seria um dos motivos do pintor sempre estar usando um chapéu, foram

consideradas evidências suficientes para sustentar o diagnóstico de Picnodisostose [7–9].

As principais características clínico-radiológicas dessa doença, presentes em mais de 50% dos pacientes, são: baixa estatura ($-1,77DP < -5,2DP$); aumento da densidade óssea, ocasionando tendência a fraturas; atraso no fechamento das fontanelas e das suturas cranianas; proeminência dos ossos frontal e parietal; ângulo mandibular obtuso; hipoplasia maxilar; palato em ogiva e mãos e pés pequenos e alargados nas extremidades, com osteólise das falanges distais [5, 9, 10]. Aproximadamente um terço dos pacientes possuem olhos proeminentes com esclera azulada e podem apresentar defeitos odontológicos, como mau alinhamento dos dentes, atraso na erupção dos dentes permanentes, com persistência dos dentes decíduos, periodontite e cáries frequentes [10]. Unhas displásicas, hipoplasia distal da clavícula e complicações respiratórias, como a apnéia obstrutiva, também ocorrem com frequência [5, 10, 11]. Essa doença não afeta a cognição e a maioria dos pacientes tem uma expectativa de vida normal [5, 10].

Apesar da rica caracterização dos achados fenotípicos dos pacientes com Picnodisostose, esse diagnóstico pode ser confundido com outros, sendo a Osteopetrose e a Disostose Cleidocraniana os principais diagnósticos diferenciais [10, 12]. O atraso no fechamento das fontanelas; a proeminência dos ossos frontal, parietal e occipital; a hipopneumatização dos seios paranasais; palato em ogiva e alterações na clavícula podem ser encontrados em ambas - displasia Cleidocraniana e na Picnodisostose [10]. Os achados de densidade óssea aumentada com consequente tendência a fraturas observados na Picnodisostose, também são comuns na Osteopetrose [10].

O diagnóstico correto e precoce da Picnodisostose é importante, seja para a orientação e prevenção de fraturas, como para conduzir as complicações odontológicas e respiratórias, propiciando uma melhor qualidade de vida a esses pacientes [5, 10, 11, 13].

1.1.2 Do gene à proteína

As bases moleculares da Picnodisostose começaram a ser desvendadas com o estudo genotípico de dezesseis indivíduos afetados, pertencentes a uma família árabe Israelita, por Gelb e colaboradores em 1995. [14, 15]. Esses autores encontraram uma região de homozigose por descendência no braço longo do

cromossomo 1 (1q21) em todos os indivíduos afetados dessa família [14]. Dentre os genes candidatos situados nesta região encontrava-se o gene *CTSK*, responsável por codificar uma potente enzima lisossômica, a protease Catepsina K, que participa nos processos de remodelação do osso e é altamente expressa nos osteoclastos [15]. Os primeiros sequenciamentos do gene *CTSK* em indivíduos com fenótipo sugestivo de Picnodisostose foram realizados no ano de 1996 e, de fato, evidenciaram mutações nesse gene [8, 15].

O gene *CTSK* possui aproximadamente 12 Kilobases (Kb) e contém oito éxons (GenBank NM_000396.3), sendo o primeiro éxon não codificante. O códon iniciador da tradução (ATG) se localiza no éxon 2 e o códon de terminação da tradução (TGA) se localiza na porção média do éxon 8 (Fig. 1) [5, 9]. Esse gene codifica a Catepsina K, uma enzima lisossômica membro da família das proteases cisteína - tipo papaína [16].

Essa protease é sintetizada como uma pré-pró proteína inativa, de 329 aminoácidos, composta por três regiões: a pré-região, codificada pela parte inicial do éxon 2, constituída por 15 aminoácidos e que funciona como uma sequência sinal responsável por direcionar a proteína até os lisossomos; a pró-região, codificada pela parte distal do éxon 2, éxon 3 e parte inicial do 4, formada por 99 aminoácidos, e a região madura ativa, codificada pelos éxons 4, 5, 6, 7 e parte do 8, a qual é constituída por 215 aminoácidos (Fig. 1) [5, 9].

A clivagem da pró-região, que ocorre em pH ácido, é responsável pela ativação da proteína e sua expressão é regulada através de múltiplos mecanismos, que incluem macromoléculas inibidoras, glicosaminoglicanos (GAGs), fatores de transcrição da família *Mitf* (*Microphthalmia transcription factor*) e parâmetros ambientais como pH e potencial redox [9, 16].

Essa endoprotease é altamente expressa nos osteoclastos, onde possui papel crucial no remodelamento e reabsorção do osso [5, 9]. Sua atividade collagenolítica é responsável por degradar moléculas de colágeno tipo I e proteínas SPARC (*secreted protein acidic and rich in cysteine*), que são os principais constituintes da matriz orgânica óssea (*apud* 14). A catepsina K também degrada moléculas de colágeno tipo II, principal constituinte das cartilagens [9]; fibronectina, que possui papel na regulação de adipócitos [16]; e elastina, proteína estrutural das paredes de vasos sanguíneos [16].

Devido a essas atividades proteolíticas, a alta expressão da Catepsina K está associada a várias doenças humanas como osteoporose; artrite; invasão do câncer e formação de metástases; doenças cardiovasculares; obesidade e esquizofrenia, por outro lado, sua baixa expressão causa Picnodisostose [9, 16].

1.1.3 Mutações descritas na literatura

Até o momento foram identificadas 50 diferentes mutações no gene *CTSK* associadas ao fenótipo de Picnodisostose [5, 12, 17–24]. Desse total, 31 mutações são do tipo *missense*, cinco são do tipo *nonsense*, oito *frameshift*, quatro mutações ocorrem em região de *splicing*, uma mutação ocorre pela deleção de um códon e uma mutação é do tipo *stop codon*, responsável por abolir um códon de terminação. A Figura 1 e a Tabela 1 resumem a localização gênica e algumas informações sobre essas mutações já descritas na literatura.

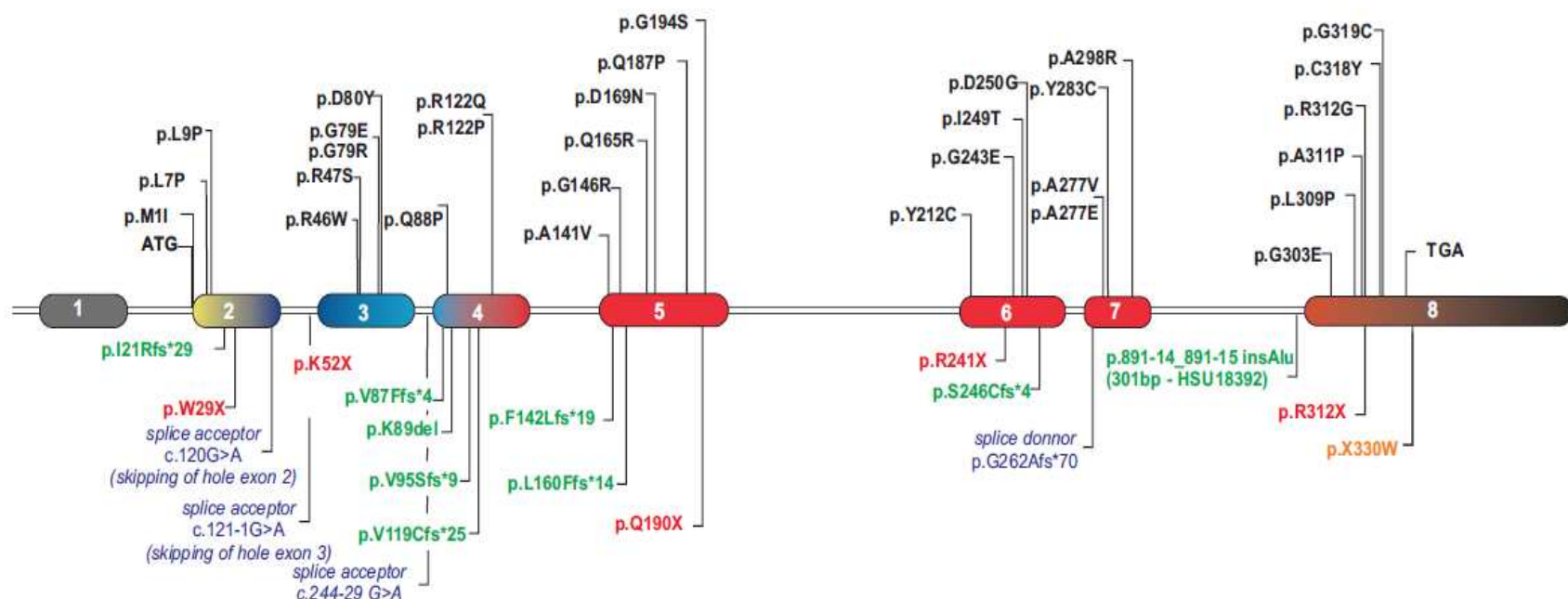


Figura 1. Estrutura do gene *CTSK*. O desenho mostra os 8 éxons e a distribuição das 50 mutações relacionadas à Picnodisostose. As mutações do tipo *missense* estão representadas em negrito na parte superior da figura, mutações do tipo *frameshift* e deleção de códon estão em verde, mutações do tipo *nonsense* em vermelho, mutações de *splicing* em azul e itálico e a mutação que abole o sítio de término da tradução está em laranja. O éxon 1 e a parte terminal do exon 8 não são codificantes (cinza). As regiões exônicas responsáveis pela codificação da pré região (15 aminoácidos), pró região (99 aminoácidos) e região madura (215 aminoácidos) estão representadas em amarelo, azul e vermelho respectivamente. As letras "ATG" no éxon 2 representam o códon de início da tradução e "TGA", o códon de término.

Tabela 1. Resumo das mutações encontradas no gene *CTSK* em pacientes com Picnoidisostose, descritas em artigos publicados até Agosto de 2015.

Localização	Origem do paciente	Referência	Alteração na sequência codificante (NM_000396.3)	Efeito no aminoácido (NP_000387.1)
Missense (31)				
Exon 2	Turquia	Arman, <i>et al.</i> , 2014 [17]	c.3G>A	p. M1I
Exon 2	Itália	Donnarumma, <i>et al.</i> , 2007 [25]	c.20T>C	p.L7P
Exon 2	Japão	Nishi, <i>et al.</i> , 1999 [26]	c.26T>C	p.L9P
Exon 3	Alemanha	Schilling, <i>et al.</i> , 2007 [27]	c.136C>T	p.R46W
Exon 3	Caribe	Pangrazio, <i>et al.</i> , 2014 [12]	c.235G>A	p.R47S
Exon 3	Áustria	Fratzl-Zelman, <i>et al.</i> , 2004 [28]	c.236G>A	p.G79R
Exon 3	Norte da Europa	Hou, <i>et al.</i> , 1999 [29]	c.139C>A	p.G79E
Exon 3	Turquia	Arman, <i>et al.</i> , 2014 [17]	c.238G>T	p.D80Y
Exon 4	Bangladesh	Arman, <i>et al.</i> , 2014 [17]	c.263A>C	p.Q88P
Exon 4	Japão	Pangrazio <i>et al.</i> , 2014 [12]	c.365G>C	p.R122P
Exon 4	China	Matsushita, <i>et al.</i> , 2012 [19]	c.365G>A	p.R122Q
Exon 5	Desconhecido	Zheng, <i>et al.</i> , 2013 [18]	c.422C>T	p.A141V
Exon 5	Marrocos, Am.	Chavassieux, <i>et al.</i> , 2008 [30]	c.436G>C	p.G146R
Exon 5	Hisp.	Gelb <i>et al.</i> , 1996 [15]	c.494A>G	p.Q165R
Exon 5	Itália	Donnarumma, <i>et al.</i> , 2007 [25]	c.505G>A	p.D169N
Exon 5	Turquia	Arman, <i>et al.</i> , 2014 [17]	c.560A>C	p.Q187P
Exon 5	China	Li, <i>et al.</i> , 2009 [31]	c.580G>A	p.G194S
Exon 5	Itália	Donnarumma, <i>et al.</i> , 2007 [25]		
Exon 6	Espanha	Hou, <i>et al.</i> , 1999 [29]	c.635A>G	p.Y212C
Exon 6	Paquistão	Donnarumma, <i>et al.</i> , 2007 [25]	c.728G>A	p.G243E
Exon 6	Espanha	Khan, <i>et al.</i> , 2010 [32]	c.746T>C	p.I249T
Exon 6	Paquistão	Donnarumma, <i>et al.</i> , 2007 [25]	c.749A>G	p.D250G
Exon 7	Bélgica	Donnarumma, <i>et al.</i> , 2007 [25]	c.830C>T	p.A277V
Exon 7	Índia, Portugal	Donnarumma, <i>et al.</i> , 2007 [25]	c.830C>A	p.A277E
Exon 7	Desconhecido	Donnarumma, <i>et al.</i> , 2007 [25]	c.848A>G	p.Y283C
Exon 7	Japão	Donnarumma, <i>et al.</i> , 2007 [25]	c.892T>C	p.W298R
Exon 8	México	Gelb, <i>et al.</i> , 1998 [33]	c.908G>A	p.G303E
Exon 8	Dinamarca	Hou, <i>et al.</i> , 1999 [29]	c.926T>C	p.L309P
Exon 8	Suíça	Xue <i>et al.</i> , 2015 [23]	c.931G>C	p.A311P
Exon 8	Honduras	Nishi, <i>et al.</i> , 1999 [26]	c.934C>G	p.R312G
Exon 8	Brasil	Toral-Lopez <i>et al.</i> , 2010 [34]	c.953G>A	p.C318Y
Exon 8	Paquistão	Haagerup, <i>et al.</i> , 2000 [35]	c.955G>T	p.G319C
		Nishi, <i>et al.</i> , 1999 [26]		
		Hou, <i>et al.</i> , 1999 [29]		
		Bertola <i>et al.</i> , 2010 [36]		
		Donnarumma, <i>et al.</i> , 2007 [25]		
Nonsense (5)				
Exon 2	Desconhecido	Xue <i>et al.</i> , 2015 [23]	c.87G>A	p.W29X
Exon 3	Norte da Europa	Hou, <i>et al.</i> , 1999 [29]	c.154A>T	p.K52X
Exon 5	Norte da Europa	Hou, <i>et al.</i> , 1999 [29]	c.568C>T	p.Q190X

Exon 6	Am. Hisp.	Gelb, <i>et al.</i> , 1996 [15]	c.721C>T	p.R241X
Exon 8	Turquia	Arman, <i>et al.</i> , 2014 [17]	c.934C>T	p.R312X
Frameshift (8)				
Exon 2	Marrocos	Donnarumma, <i>et al.</i> , 2007 [25]	c.60_61dupGA	p.I21Rfs*29
Exon 4	Áustria		c.259delG	p.V87Ffs*4
Exon 4	Paquistão	Fratzl-Zelman, <i>et al.</i> , 2004 [28]	c.282dupA	p.V95Sfs*9
Exon 4	Turquia	Donnarumma, <i>et al.</i> , 2007 [25]	c.354dupT	p.V119Cfs*25
Exon 5	Japão		c.426delT	p.F142Lfs*19
Exon 5	Índia	Özdemir, <i>et al.</i> , 2013 [21]	c.479dupT	p.L160Ffs*14
Exon 6	Caribe	Fujita, <i>et al.</i> , 2000 [37] Singh <i>et al.</i> , 2014 [22]	c.737_738delCT	p.S246Cfs*4
<u>Intron 7</u>	Turquia	Pangrazio, <i>et al.</i> , 2014 [12] Arman <i>et al.</i> , 2014 [17]	c.891-14_891-15 ins HSU18392*	-----
Deleção de Codon (1)				
Exon 4	Paquistão	Pangrazio, <i>et al.</i> , 2014 [12]	c.266_268delAGA	p.K89del
Splice (4)				
Exon 2	Tailândia	Utokpat <i>et al.</i> , 2013 [20]	c.120 G>A (r.-1 120del)	-----
<u>Intron 2</u>	Dinamarca	Haagerup, <i>et al.</i> , 2000 [35]	c.121-1G>A	p.V41_M81del
<u>Intron 3</u>	Arábia Saudita	Alhashem <i>et al.</i> , 2015 [24]	c.244 -29 G>A**	-----
Exon 7	Egito	Donnarumma, <i>et al.</i> , 2007 [25]	c.890G>A (r.785_890del)	p.G262Afs*70
Stop Codon (1)				
Exon 8	Israel	Gelb, <i>et al.</i> , 1996 [15]	c.990A>G	p.X330W

* Inserção em orientação reversa de 301 pares de base da sequência *Alu* representada pelo Código HSU18392.

** Trabalho exposto em conferência. 12th ISDS MEETING ISTANBUL Hilton İstanbul Bosphorus Hotel, Turkey.

Am. Hisp (América Hispânica): Parte da América colonizada por Espanhóis.

NOTA: A tabela informa a localização da mutação no DNA e o efeito no aminoácido, bem como a origem do paciente no qual a mutação foi identificada pela primeira vez e a referência. Em sua confecção foram utilizadas as recomendações da HGVS (*Human Genome Variation Society*) para padronizar a nomenclatura das mutações, por esse motivo, algumas mutações podem estar com notação diferente da que consta nos artigos referência.

Dentre as 50 mutações descritas na literatura [12, 15, 17–38], apenas para 32 delas há descrição do fenótipo dos pacientes. Dessa forma, para 18 mutações (p.L9P, p.I21Rfs*29, p.V41_81Mdel, p.K52X, p.V95Sfs*9, p.F142Lfs*19, p.Q165R, p.Q190X, p.G194S, p.Y212C, p.D250G, p.G262Afs*70, p.A277E, p.W298R, p.L309P, p.A311P, p.G319C e p.X330W) o fenótipo não pode ser confrontado com o genótipo. A falta de descrição fenotípica é justificada pelos autores, pela similaridade

clínica e radiológica de seus pacientes com o quadro típico da Picnodisostose [25, 26, 29, 35, 37, 39].

1.1.4 A prevalência da Picnodisostose

Segundo dados epidemiológicos da União Europeia, doenças raras são aquelas com prevalência inferior a um afetado a cada 2.000 indivíduos nascidos [40]. Já no Brasil, de acordo com o Ministério da Saúde, Artigo 3º do Capítulo 1 da Portaria N°199, uma doença é considerada rara se afeta 1,3 pessoas para cada 2.000 indivíduos ou menos [41]. A Picnodisostose, portanto, é considerada uma doença rara. Baseado nos casos descritos na literatura, menos de 200 relatados desde 1962 até 2009 [42], alguns autores estimaram sua prevalência em torno de 1 a 1,7 por 1.000.000 de indivíduos [43–45].

1.2 DADOS LOCAIS

A partir da visita da Coordenadora do grupo local de Displasia Esquelética ao Hospital Infantil Albert Sabin (Fortaleza, CE) em outubro de 2013 para discussão dos casos locais de displasias esqueléticas, foi observada uma alta frequência de pacientes com Picnodisostose. Na época foram apresentados dados clínico-radiológicos de 22 famílias com pelo menos um afetado por esta doença. A taxa de consanguinidade entre essas famílias foi de 59% [46]. Diante desses dados, a hipótese de possível efeito fundador motivou a presente investigação.

De acordo com o evolucionista Ernst Mayr (*apud* 47), dá-se o nome de efeito fundador a uma pequena população isolada que foi originada a partir de poucos indivíduos, os quais representam apenas uma parcela da população original. Esses indivíduos fundadores podem, ao acaso, portar algum alelo deletério raro que ao invés de ser eliminado por seleção natural, é mantido por várias gerações devido a ocorrências de uniões entre indivíduos consanguíneos.

2. OBJETIVOS

2.1 GERAL

Realizar uma investigação molecular de famílias com pelo menos um indivíduo afetado por Pichnodisostose no Estado do Ceará por meio do sequenciamento do gene *CTSK* e disponibilizar investigação semelhante para famílias de outras regiões.

2.2 ESPECÍFICOS

1. Avaliar tipo(s) e frequência de mutações encontradas nas famílias de portadores de Pichnodisostose;
2. Avaliar possível relação entre genótipo e fenótipo;
3. Analisar a origem das famílias a partir da identificação do local de nascimento dos pais e avós maternos e paternos;
4. Avaliar a taxa de consanguinidade referida pelas famílias e confrontá-la com a taxa real de consanguinidade a partir dos resultados do sequenciamento do *CTSK* das famílias estudadas;
5. Investigar, por meio do estudo de haplótipos, a hipótese de efeito fundador, caso a investigação molecular prévia ratifique tal hipótese.

3. CASUÍSTICA, MATERIAIS E MÉTODOS

3.1 CASUÍSTICA

Neste estudo foram incluídos 33 probandos, dentre os quais, 18 são oriundos do Estado do Ceará e 15 de outras localidades (São Luís, MA; Campina Grande, PB; Formosa, GO; Jaguariúna, SP; Campinas, SP; São Carlos, SP; Guarulhos, SP; São Paulo, SP e Porto Alegre, RS). Os dados clínicos disponíveis, bem como as fotos clínicas e radiografias de todos os pacientes, foram previamente avaliados pela orientadora desse trabalho, antes de ser realizada a investigação molecular.

Esse trabalho foi aprovado pelo Comitê de Ética e Pesquisa da FCM-UNICAMP e pelo Comitê de Ética e Pesquisa do HIAS de Fortaleza (Anexo 1). Todos os participantes da pesquisa foram orientados quanto à proposta de estudo e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) (Anexo 2).

3.2 MÉTODOS

Toda a metodologia deste trabalho foi desenvolvida no Laboratório de Genética Molecular do Departamento de Genética da FCM-UNICAMP.

3.2.1 Desenho e Padronização dos *Primers*

Para a amplificação e sequenciamento dos éxons do gene *CTSK* foram construídos oligonucleotídeos iniciadores (*primers*) que flanqueiam essas regiões com o auxílio dos programas *Gene Runner®* e *Primer 3®*.

Foram utilizados cinco pares de *primers* (Apêndice 1) cujas temperaturas ideias de anelamento foram determinadas através de *TouchDown-PCR* (*TouchDown Polymerase Chain Reaction*), com gradiente de temperatura.

3.2.2 Obtenção do material biológico

As amostras de sangue periférico dos participantes da pesquisa foram coletadas nos locais de atendimento dos pacientes e seus familiares e encaminhadas para o laboratório. A extração do DNA foi realizada utilizando o protocolo manual de extração com fenol-clorofórmio adaptado de Sambrook e Green [48]. O produto da extração foi quantificado em espectrofotômetro EpochTM (BioTek Instruments, Inc.©) e diluído para a concentração final de 50ng/μL.

3.2.3 Amplificação do DNA por reação em cadeia da Polimerase (PCR)

A partir do DNA genômico de cada paciente, as regiões flanqueadas pelos *primers* foram amplificadas pelo método *Touchdown* PCR utilizando-se o Kit Taq DNA Polymerase (Thermo Scientific) nas seguintes condições: 1 μL de tampão 10X com KCl (concentração final de 1X); 1 μL de MgCl (2,5mM); 1 μL de dNTP mix (contendo 0,2 mM de cada dNTP- Thermo Scientific); 0,5 pmol de cada par de *primers* específicos (*forward* e *reverse*); 0,05 μL de Taq DNA polimerase recombinante (0,25 U); 50 ng de DNA controle e água para o volume final de 10μL.

As condições para amplificação foram: desnaturação inicial a 95°C por 5 minutos (1 ciclo); 20 ciclos com desnaturação a 95°C por 30 segundos, anelamento inicial com a temperatura de 64,1°C no primeiro ciclo, seguido de anelamento diminuindo 0,5°C a cada ciclo, até atingir temperatura ideal final de anelamento de 54,1°C, extensão a 72°C por 1 minuto; 15 ciclos com desnaturação a 95°C por 30 segundos, anelamento com temperatura ideal de 54,1°C, extensão a 72°C por 1 minuto e extensão final a 72°C por 10 minutos (1 ciclo).

Para cada éxon amplificado foi adicionada uma amostra controle (sem DNA), com a finalidade de identificar eventuais contaminações durante o processo.

A confirmação da qualidade da PCR foi realizada através da aplicação de 2μL de cada amostra, misturada a 2 μL de solução de azul de bromofenol em gel de agarose 1,5% corado com o fluoróforo *SYBR Safe*® (concentração final de 0,2 μL/mL de gel, Invitrogen), realização da eletroforese (~25 minutos a 110 v) e revelação sob luz ultravioleta. A fim de se identificar os fragmentos amplificados, foi

utilizado um marcador de peso molecular de 100 pares de base (*GeneRuler*[™] DNA Ladder, Thermo Scientific©).

3.2.4 Sequenciamento direto do gene *CTSK* pelo método de Sanger

Para a reação de sequenciamento foi utilizado o kit *ABI Prism*[®] *BigDye*[®] *Terminator V3.1* (*Applied Biosystems*[®]), nas seguintes condições: 2µL da solução tampão fornecida; 0,6 µL de Bigdye[®]Terminator; 0,5 µL de iniciador senso ou antissenso (10pmol/ µL); 1 µL do produto da PCR e água para o volume final de 10 µL. As amostras foram colocadas em termociclador e submetidas a um programa de 35 ciclos, composto pelas etapas de desnaturação à 95°C por 20 segundos, anelamento à temperatura específica de 54,1°C por 15 segundos e extensão à 60°C por 1 minuto.

A purificação da reação foi realizada com EDTA e etanol, como sugerido pelo fabricante (*Applied Biosystems*[®]). Ao final, as amostras foram solubilizadas em 10 µL de formamida Hi-Di[™] (*Applied Biosystems*[®]), e sequenciadas no aparelho ABI 3500 XL (*Applied Biosystems*[®]).

3.2.5 Análise do Sequenciamento

Os arquivos do sequenciamento foram analisados através do programa *CodonCodeAligner*[®], versão 4.1., cujo papel é o alinhamento das sequências, portanto identificando as mutações a partir das alterações na sequência de nucleotídeos. As alterações encontradas foram comparadas com bancos de dados do *Ensembl*, HGMD[®] e dbSNP, bem como a partir das mutações já descritas em artigos científicos. Para as alterações não descritas na literatura, foram utilizados programas *in silico* a fim de se estimar o efeito na proteína. Os softwares MutPred[®], Polyphen[®], SNP&GO[®], Panther[®] e SIFT[®] avaliam substituições de bases nitrogenadas e o Mutationtaster[®] avalia inserções e deleções.

3.2.6 Estudo de haplótipos

O estudo dos haplótipos foi utilizado para investigar a hipótese de efeito fundador. Para esse estudo, foram selecionados quatro marcadores polimórficos do tipo microssatélite, próximos ao gene *CTSK* (D1S2344, D1S442, D1S498 e D1S2715), os quais já haviam sido utilizados em estudos de genotipagem em pacientes com Picnodisostose [8, 14, 35, 42].

Os *primers* dos marcadores selecionados (Anexo 2) foram submetidos à reação da cadeia polimerase (PCR), segundo protocolo padronizado para os demais *primers* (vide item 3.2.3), porém com uma temperatura ideal de anelamento de 55°C, e não de 54,1°C como previamente utilizada. Para a reação de genotipagem foram utilizados 1 µL do produto da PCR diluído (1:20), 9 µL de formamida Hi-Di™ (*Applied Biosystems*®), e 0,2 µL de padrão de peso molecular 600 LIZ® v 2.0 (*Applied Biosystems*®). As amostras foram genotipadas no aparelho ABI 3500 XL (*Applied Biosystems*®) e os dados gerados foram lidos no programa *GeneMapper* v4.1®.

3.2.7 Revisão de Fenótipos

A revisão dos fenótipos para avaliação de possível correlação genótipo-fenótipo foi realizada a partir da reavaliação de todos os dados clínicos, bem como pela documentação iconográfica (fotos clínicas e radiografias) disponível dos casos. Para essa revisão, construiu-se uma tabela incluindo os seguintes dados: Idade ao ser incluído no presente estudo; baixa estatura; histórico de fraturas; braquidactilia; unhas displásicas; proeminência dos ossos cranianos frontal e parietal; hipoplasia da face média; hipoplasia mandibular; nariz em bico; proptose ocular; esclera azulada; densidade óssea aumentada, osteólise das falanges distais; suturas cranianas e fontanelas abertas; hipoplasia distal da clavícula, além de alguns achados particulares.

4. RESULTADOS

4.1 DESCRIÇÃO DA CASUÍSTICA

Esse estudo teve como objetivo inicial investigar famílias com pelo menos um afetado por Picnodisostose no Estado do Ceará, entretanto, incluiu também famílias provenientes de outros Estados. No total foram analisados 39 indivíduos afetados de 33 famílias, dentre as quais, 18 são provenientes do Ceará e as demais de outras localidades (Tabela 2).

A recorrência em irmãos foi referida em nove famílias, no entanto, apenas seis indivíduos aparentados de seis famílias estavam disponíveis para estudo. Dentre estes indivíduos aparentados, apenas um apresentou sequenciamento negativo (irmão da paciente 12), apesar de possuir fenótipo sugestivo (Tabela 2). Infelizmente, até o momento não foi possível repetir o sequenciamento a partir de nova amostra.

Considerando todos os indivíduos com diagnóstico molecular confirmado, 22 afetados são do sexo masculino e 16 são do sexo feminino. A idade de todo o conjunto de afetados variou de sete meses (0,66 anos) a 61 anos (Tabela 2).

A consanguinidade foi referida em 13 famílias (39,4%). Dessas, três desconheciam o grau de parentesco, portanto foram considerados como tendo um parentesco distante. Avaliando o conjunto de famílias aparentadas da presente casuística, observou-se que as relações de parentesco distante ($F < 1/64$) foram tão comuns quanto as relações de parentesco mais próximas (53,8%) (Tabela 2).

4.2 MUTAÇÕES ENCONTRADAS

O sequenciamento do gene *CTSK* dos pacientes do Ceará identificou cinco diferentes mutações, sendo quatro previamente descritas (p.G146R, p.G194S, p.R241X, p.C318Y) e uma inédita no éxon 2 (p.W29Mfs*10). A maioria dessas mutações foi observada em homozigose (83,3%), apesar de a consanguinidade parental ser referida por apenas 33,3% das famílias (Tabela 2).

No sequenciamento dos pacientes oriundos de outras regiões do Brasil foram encontradas as mesmas mutações observadas nos pacientes do Ceará, exceto pela

paciente 33, do Rio Grande do Sul, na qual se observou a mutação p.A277V. Portanto, no total foram observadas seis diferentes mutações na casuística total (Tabela 2). Dentre o conjunto de 15 probandos provenientes de outras localidades do Brasil, oito apresentaram mutações em homozigose (53,3%), apesar de sete terem referido consanguinidade (Tabela 2).

Uma vez que mutações no gene *CTSK* são consideradas raras, mutações em homozigose sugerem fortemente a presença de ancestral comum. Portanto, a alta frequência de mutações em homozigose na casuística total (69,7%) sugere que a consanguinidade parental pode ser de fato, maior que a referida pelas famílias (Tabela 2).

Por meio do sequenciamento dos genitores dos pacientes foi possível determinar as origens materna e paterna de cada alelo mutado, exceto nos pacientes 25 e 31, os quais são heterozigotos compostos e cujos pais não puderam ser avaliados (Tabela 2).

Tabela 2. Dados clínicos e moleculares das famílias estudadas.

Nº do Caso índice	Sexo	Idade (anos)	F	Rec.	Genótipo do caso índice			
					Alelo materno	éxon	Alelo paterno	éxon
Pacientes do Ceará								
1	M	4	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
2	F	17	0	N	p.W29Mfs*10	2	p.R241X	6
3	M	21	0	N	p.R241X	6	p.R241X	6
4	F	18	1/32	N	p.C318Y	8	p.C318Y	8
5	F	13	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
6	M	18	0	N	p.R241X	6	p.R241X	6
7	F	14	?	*Irmã	p.R241X	6	p.R241X	6
8	M	9	0	Irmã	p.G146R	5	p.G146R	5
9	M	16	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
10	M	14	1/256	N	p.W29Mfs*10	2	p.W29Mfs*10	2
11	M	17	1/16	N	p.R241X	6	p.R241X	6
12	F	12	0	**Irmão	p.R241X	6	p.R241X	6
13	M	18	1/16	N	p.G146R	5	p.G146R	5
14	M	4	0	Irmão	p.G156R	5	p.G194S	5
15	M	11	1/16	N	p.G146R	5	p.G146R	5
16	M	12	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
17	F	24	0	N	p.W29Mfs*10	2	p.R241X	6
18	M	12	0	N	p.G146R	5	p.G146R	5
Pacientes de outros Estados								
19	M	24	?	*Irmã	p.G146R	5	p.G146R	5
20	M	0,66	0	Irmã	p.R241X	6	p.G146R	5
21	F	26	0	*Irmão	p.C318Y	8	p.C318Y	8
22	F	5	0	N	p.C318Y	8	p.R241X	6
23	M	22	?	N	p.R241X	6	p.R241X	6
24	M	6	1/16	N	p.R241X	6	p.R241X	6
25	M	6	0	N	#p.R241X	6	#p.G246R	5
26	F	5	0	Irmã	p.W29Mfs*10	2	p.R241X	6
27	F	28	1/64	N	p.G146R	5	p.G146R	5
28	M	6	1/256	N	p.C318Y	8	p.C318Y	8
29	M	11	0	Irmã	p.R241X	6	p.G146R	5
30	F	8	0	N	p.C318Y	8	p.W29Mfs*10	2
31	M	61	0	N	#p.R241X	6	#p.G146R	5
32	M	10	1/16	N	p.R241X	6	p.R241X	6
33	F	6	1/256	N	p.A277V	7	p.A277V	7

M: Masculino; F: Feminino; S: Sim, ?.: Há consanguinidade, mas o coeficiente de endocruzamento não pode ser calculado, pois a família desconhece o grau de parentesco; Cons.: consanguinidade parental; F: Coeficiente de endocruzamento; Rec.: Recorrência; *Sem DNA disponível para análise, **Não foram encontradas mutações, #Não foi possível determinar a origem, materna ou paterna, dos alelos

A figura 2 ilustra o mapa do Ceará com a distribuição geográfica das cinco mutações encontradas nos 18 pacientes de lá provenientes. Nota-se uma concentração da mutação inédita (p.W29Mfs*10) em pequenas cidades próximas no noroeste do Estado, o que sugere uma origem única desta mutação. Por outro lado, as mutações mais frequentes, p.R241X e p.G146R, aparecem mais dispersas no mapa. As mutações p.G194S e p.C318Y foram as menos frequentes, sendo a primeira observada em um probando, heterozigoto composto, proveniente de Iguatu e a segunda, em homozigose, em um único indivíduo de Fortaleza (Fig. 2).

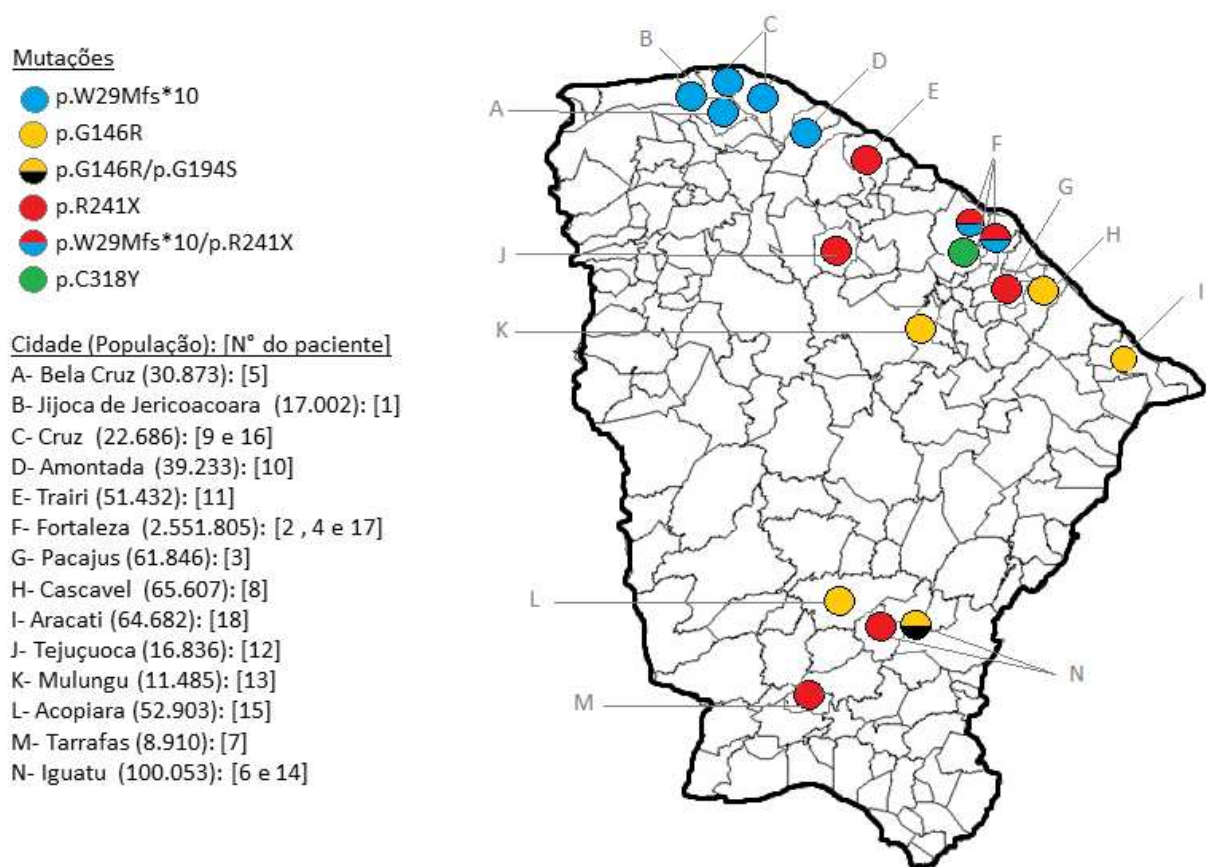


Figura 2. Distribuição das mutações encontradas nos pacientes provenientes do Ceará (1 a 18). Cada círculo representa um paciente que carrega um tipo de mutação (homozigoto) ou dois (heterozigoto composto), de acordo com as cores indicadas na legenda da figura.

Nos pacientes provenientes de outras regiões (paciente 19 ao paciente 33) foram detectadas quatro mutações idênticas às observadas nas famílias do Ceará (p.W29Mfs*10, p.G146R, p.R241X e p.C318Y), e uma diferente (p.A277V), porém já descrita na literatura.

A figura 3 ilustra a distribuição das mutações encontradas nesse grupo de pacientes. As mutações observadas mais frequentemente nos pacientes do Ceará (p.R241X e p.G146R), também foram as mais frequentes nesses pacientes. A mutação inédita, p.W29Mfs*10, foi encontrada em dois pacientes heterozigotos compostos (26 e 30), cujos genitores portadores desta mutação possuem antepassados nordestinos (Tabela 3).

Ao serem analisadas as origens das famílias materna e paterna dos demais pacientes, foi verificado que todos, exceto três (pacientes 26, 31 e 33), referem que ambos os genitores possuem antepassados nordestinos (Tabela 3). A família paterna do paciente 26 é proveniente de Minas Gerais com antepassados portugueses; o paciente 31, de São Carlos, relatou ter antepassados da península Ibérica (portugueses e espanhóis); finalmente, os genitores da paciente 33, de Vacaria no Rio Grande do Sul, relataram antepassados italianos (Tabela 3).

Dentre os pacientes homozigotos desse grupo, apenas o paciente 21 não referiu consanguinidade parental, contudo, seus genitores são naturais de cidades muito próximas no estado da Paraíba, sugerindo consanguinidade desconhecida e/ou distante (Tabela 3).

Mapa das regiões do Brasil

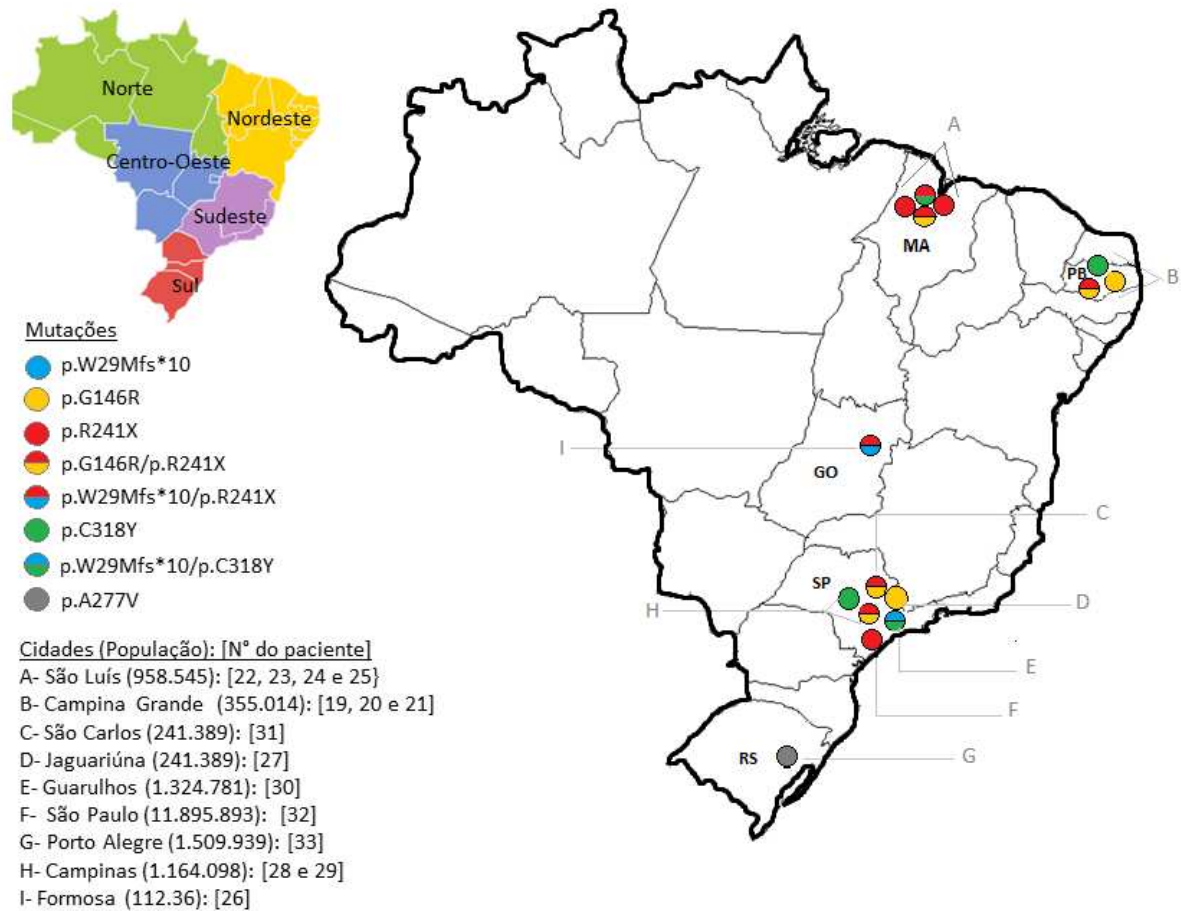


Figura 3. Distribuição das mutações encontradas nos provenientes de outras localidades (19 a 33). Cada círculo representa um paciente que carrega uma mutação (homozigoto) ou duas mutações (heterozigoto composto), de acordo com as cores indicadas na legenda da figura.

Tabela 3. Procedência e origem dos pacientes estudados.

Caso Índice	Endereço atual do paciente	Alelo Materno	Origem da família Materna						Alelo Paterno	Origem da família Paterna						
			Mãe		Avó Materna		Avô Materno			Pai		Avó Paterna		Avô Paterno		
			Cidade	Estado	Cidade	Estado	Cidade	Estado		Cidade	Estado	Cidade	Estado	Cidade	Estado	
Pacientes do Estado do Ceará																
1	J.Jericoacoara	W29Mfs*10	Cruz	CE	J.Jericoacoara	CE	Cruz	CE	W29Mfs*10	J.Jericoacoara	CE	J.Jericoacoara	CE	J. Jericoacoara	CE	
2	Fortaleza	W29Mfs*10	Itarema	CE	Itarema	CE	Itarema	CE	R241X	Itarema	CE	Itarema	CE	Itarema	CE	
3	Pacajus	R241X	Pacajus	CE	Fortaleza	CE	Iguatu	CE	R241X	Jaguaretama	CE	Jaguaretama	CE	Jaguaretama	CE	
4	Fortaleza	C318Y	Fortaleza	CE	-	-	-	-	C318Y	-	CE	-	-	-	-	
5	Bela Cruz	W29Mfs*10	Bela Cruz	CE	Bela Cruz	CE	Bela Cruz	CE	W29Mfs*10	Bela Cruz	CE	Bela Cruz	CE	Bela Cruz	CE	
6	Iguatu	R241X	Iguatu	CE	-	PB	-	PB	R241X	Iguatu	CE	Iguatu	CE	Iguatu	CE	
7	Tarrafas	R241X	-	CE	-	-	-	-	R241X	-	CE	-	-	-	-	
8	Cascavel	G146R	Cascavel	CE	Cascavel	CE	Cascavel	CE	G146R	Mossoró	RN	Cascavel	CE	Mossoró	RN	
9	Cruz	W29Mfs*10	Cruz	CE	Cruz	CE	Cruz	CE	W29Mfs*10	Cruz	CE	-	-	-	-	
10	Amontada	W29Mfs*10	-	CE	-	-	-	-	W29Mfs*10	-	CE	-	-	-	-	
11	Trairi	R241X	-	CE	-	-	-	-	R241X	-	CE	-	-	-	-	
12	Tejuçuoca	R241X	-	CE	-	-	-	-	R241X	-	CE	-	-	-	-	
13	Mulungu	G146R	-	CE	-	-	-	-	G146R	-	CE	-	-	-	-	
14	Iguatu	G146R	-	CE	-	-	-	-	G194S	-	CE	-	-	-	-	
15	Acopiara	G146R	-	CE	-	-	-	-	G146R	-	CE	-	-	-	-	
16	Cruz	W29Mfs*10	Cruz	CE	-	-	-	-	W29Mfs*10	Bela Cruz	CE	-	-	-	-	
17	Fortaleza	W29Mfs*10	Natal	RN	Natal	RN	Natal	RN	R241X	Fortaleza	CE	Trairi	CE	Fortaleza	CE	
18	Aracati	G146R	Aracati	CE	-	-	-	-	G146R	Aracati	CE	-	-	-	-	
Pacientes de outros Estados																
#19	C. Grande	PB	G146R	Jericó	PB	-	-	-	-	G146R	Jericó	PB	-	-	-	-
#20	C. Grande	PB	R241X	Mombaça	CE	-	-	-	-	G146R	Sousa	PB	-	-	-	-
#21	C. Grande	PB	C318Y	Arara	PB	-	-	-	-	C318Y	Casserengue	PB	-	-	-	-
#22	São Luís	MA	C318Y	São Luís	MA	Grajaú	MA	Grajaú	MA	R241X	São Luís	MA	Gov.Eug.Barros	MA	I. Carrapatal	MA
#23	São Luís	MA	R241X	-	MA	-	-	-	-	R241X	-	MA	-	-	-	-
#24	São Luís	MA	R241X	Jenipapolând.	MA	-	-	-	-	R241X	-	MA	-	-	-	-
#25	São Luís	MA	R241X (?)	-	MA	-	-	-	-	G146R(?)	-	MA	-	-	-	-
26	Formosa	GO	W29Mfs*10	Currais Novos	RN	Currais Novos	RN	-	RN	R241X	Uberlândia	MG	Uberlândia	MG	Uberlândia	MG
#27	Jaguaruína	SP	G146R	Acopiara	CE	-	-	-	-	G146R	Acopiara	CE	-	-	-	-
#28	Campinas	SP	C318Y	Arapiraca	AL	-	AL	-	AL	C318Y	Feira Grande	AL	Feira Grande	AL	Feira Grande	AL
#29	Campinas	SP	R241X	R. Pombal	BA	-	BA	-	BA	G146R	R. Pombal	BA	-	BA	-	BA
#30	Guarulhos	SP	C318Y	V. Conquista	BA	V. Conquista	BA	V. Conquista	BA	W29Mfs*10	S. Parnaíba	SP	J. do Norte	CE	-	PR
31	São Carlos	SP	R241X(?)	São Carlos	SP	São Carlos	SP	São Carlos	SP	G146R(?)	São Carlos	SP	São Carlos	SP	São Carlos	SP
#32	São Paulo	SP	R241X	Lagoa	PB	-	-	-	-	R241X	Lagoa	PB	-	-	-	-
33	P. Alegre	RS	A277V	Vacaria	RS	Vacaria	RS	Vacaria	RS	A277V	Vacaria	RS	Vacaria	RS	Vacaria	RS

#Pacientes com antepassados nordestinos. (?): não foi possível determinar origem, materna ou paterna, de cada um dos alelos. J Jericoacoara: Jijoca de Jericoacoara, C. Grande: Campina Grande, Gov. Eugênio de Barros: Governador Eugênio de Barros, Jenipapolând.: Jenipapolândia, R. Pombal: Ribeira do Pombal, V. da Conquista: Vitória da Conquista, S. Parnaíba: Santana do Parnaíba, J. do Norte: Juazeiro do Norte. Todas as mutações, exceto a p.W29Mfs*10 já são descritas na literatura.

4.3 Combinação das mutações encontradas em todos os probandos da casuística

Na tabela 4 estão listadas as combinações alélicas maternas e paternas encontradas nos probandos das 33 famílias estudadas. As mutações aparecem em homozigose nos indivíduos afetados de 23 famílias, apesar de apenas 12 terem relatado consanguinidade parental (Tabela 2).

Dentre os probandos homozigotos, oito são portadores da mutação p.R241X, seis a mutação p.G146R, cinco a mutação inédita p.W29Mfs*10, três a mutação p.C318Y e um a mutação p.A277V.

Entre os dez probandos heterozigotos compostos, todos de famílias não consanguíneas, foram identificadas cinco combinações diferentes. A combinação mais frequente foi entre as mutações p.R241X e p.G146R, observada em quatro probandos. Esta primeira mutação também foi vista em combinação com a mutação inédita p.W29Mfs*10 em três probandos e juntamente com a mutação p.C318Y em apenas um indivíduo.

As mutações p.W29Mfs*10 e p.C318Y também foram detectadas combinadas em um paciente. A mutação p.G194S, que aparece uma única vez em toda a coorte analisada, foi encontrada em combinação com a mutação p.G146R em um único probando (Tabela 4).

Tabela 4. Combinação das mutações encontradas nos probandos das 33 famílias estudadas

Alelo materno Alelo paterno	p.W29Mfs*10 éxon 2	p.G146R éxon 5	p.G194S éxon 5	p.R241X éxon 6	p.A277V éxon 7	p.C318Y éxon 8
p.W29Mfs*10 éxon 2	5/33	-	-	-	-	1/33
p.G146R éxon 5	-	6/33	1/33	4/33*	-	-
p.G194S éxon 5	-	-	-	-	-	-
p.R241X éxon 6	3/33	-	-	8/33	-	1/33
p.A277V éxon 7	-	-	-	-	1/33	-
p.C318Y éxon 8	-	-	-	-	-	3/33

*Em dois desses probandos a origem do alelo materno e paterno não pode ser determinada. Os números dentro dos quadrantes são a frequência da combinação encontrada nos probandos. Os homozigotos estão em destaque.

4.4 Breve caracterização da mutação inédita p.W29Mfs*10 (c.83dupT)

A mutação inédita não está descrita em nenhum banco de dados (HGMD, dbSNP, *Ensembl*). A duplicação de uma única base, Timina, na posição 83 do cDNA no éxon 2, altera a matriz de leitura durante a tradução da proteína, gerando um códon de parada dez aminoácidos à frente do primeiro modificado, na posição 29 do peptídeo.

Apesar de ser uma mutação *frameshift* e, portanto, patogênica por definição, foi efetuada análise *in silico* pelo software Mutationtaster®, corroborando sua patogenicidade e possível alteração na estrutura da proteína. A figura 2 mostra o eletroferograma produzido pelo programa *CodonCodeAligner*® a partir do sequenciamento de um paciente homozigoto para essa mutação e de um indivíduo controle.

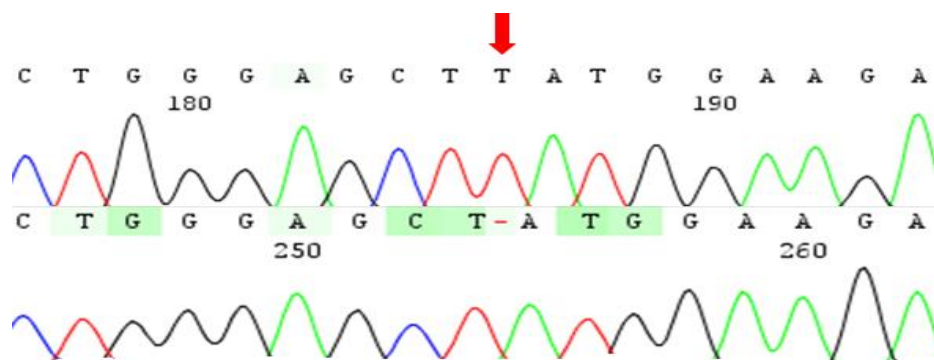


Figura 4. Eletroferograma mostrando a duplicação em homozigose da base timina no éxon 2 (c.83dupT). É possível visualizar que há uma base timina a mais na sequência de DNA do paciente homozigoto (superior) em comparação com a sequência de DNA de um indivíduo controle (inferior).

4.5 Breve caracterização das mutações já descritas na literatura

Na mutação p.G146R, ocorre a transversão da base Guanina para Citosina na posição 436 do cDNA (éxon 5), provocando a substituição do aminoácido Glicina pelo aminoácido Arginina no domínio maduro da proteína, próximo ao sítio ativo [9, 39]. Estudos *in vitro* demonstraram que apesar do nível de expressão do RNAm ser normal nas células mutantes, o nível de expressão da proenzima está baixo e não há detecção de proteína madura nem de atividade enzimática da Catepsina K [29].

A mutação p.G194S, também localizada no éxon 5, é resultado da substituição do aminoácido Glicina por Serina em uma região bastante conservada das proteínas papaínas [25]. Essa troca parece diminuir a flexibilidade conformacional da cadeia polipeptídica, além de gerar um novo sítio de glicolização [25].

No éxon 6, a transição de uma Citosina para uma Timina na posição 721 do cDNA, resulta na troca de uma Arginina por um códon de parada no domínio maduro da enzima (p.R241X). Como consequência dessa troca há produção de uma proteína truncada em que estão ausentes dois resíduos conservados de sítio ativo (resíduos H²⁷⁶ e N²⁹⁶) [9, 39]. A posição 721 do cDNA do gene *CTSK* está localizada em uma região com várias repetições de nucleotídeos C (Citosina) e G (Guanina), chamada de ilhas CpG, a qual é propícia para ocorrência de mutações (23).

Na mutação p.A277V, a troca de uma Citosina, localizada nas ilhas CpG do éxon 7, por uma Timina ocorrendo na posição 830 do cDNA, resulta na troca de aminoácidos em uma região próxima ao sítio ativo H²⁷⁶ e essa substituição parece interferir na atividade catalítica da enzima [9, 29]. Estudos realizados *in vitro* em células mutantes mostraram uma expressão normal do RNAm, porém com baixo nível de expressão da proenzima e nenhuma detecção da proteína madura ou de atividade enzimática [9, 29]. Essa mutação está localizada em uma região propensa a mutações, tanto pelo fato de estar em uma ilha CpG [5], como por estar próxima ao sítio ativo da proteína [16].

Por fim, a transição de uma Guanina por uma Adenina na posição 953 do cDNA (éxon 8) na mutação p.C318Y, ocasiona a troca do aminoácido Cisteína por Tirosina [25]. Essa troca provoca dobramentos anormais na estrutura quaternária da proteína, decorrentes da modificação das ligações de dissulfeto entre os aminoácidos da posição 318 e 269 [25]. A posição 318 da proteína, também está localizada próxima ao sítio ativo da proteína [16].

4.6 Análise de Haplótipos

A partir da sugestão de que a mutação inédita, p.W29Mfs*10, pode ter uma origem única no noroeste do Ceará (Figura 2) e estar se disseminando por meio dos

casamentos consanguíneos frequentes nessa região, a hipótese de efeito fundador foi mantida para essa mutação. Para testar essa hipótese foi realizado o estudo de haplótipos com quatro marcadores microssatélites nos nove pacientes portadores desta mutação e seus respectivos genitores e irmãos afetados.

A figura 5 mostra, de acordo com dados disponíveis na literatura, a localização de alguns marcadores próximos ao gene *CTSK*. Estão apresentadas as distâncias gênicas (em centimorgans-cM), mapeada e disponível no banco de dados *Genethon Genetics Maps*; e a física (em Megabases-Mb), mapeada e disponível pelo banco de dados *Ensembl*. Embora a distância varie de um mapa para outro, a ordem dos marcadores é mesma.

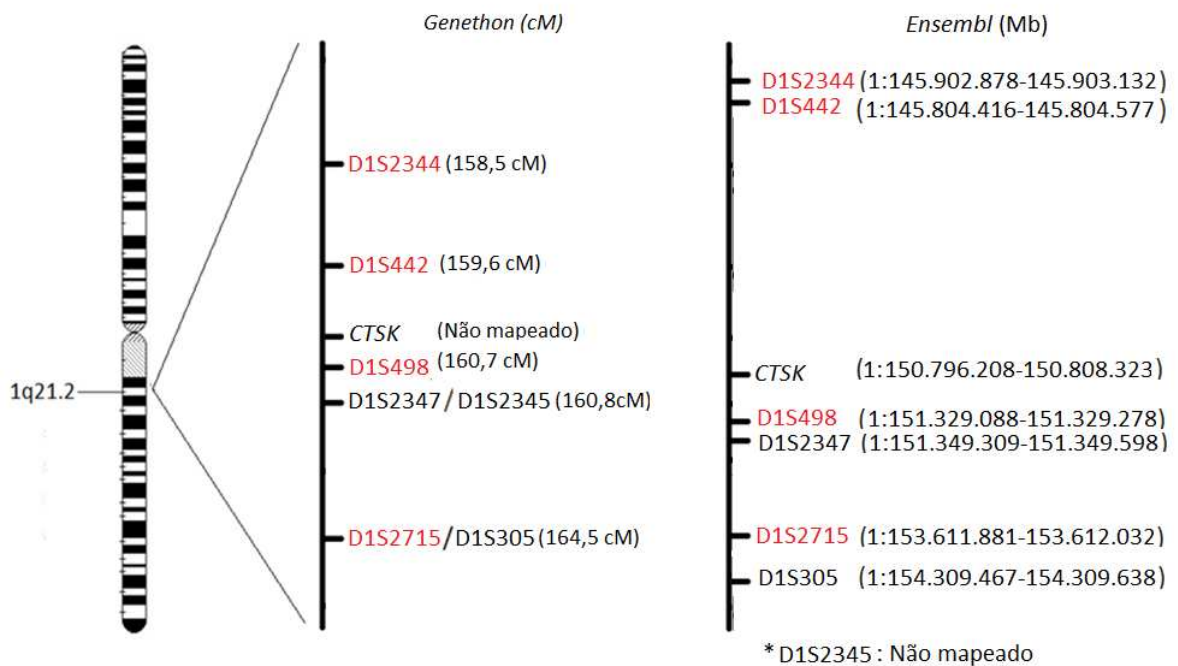


Figura 5. Distribuição de alguns marcadores no cromossomo 1 de acordo com as distâncias gênicas (*Genethon*) e físicas (*Ensembl*). Em vermelho estão os marcadores utilizados no presente estudo de haplótipos.

A análise da genotipagem dos probandos está resumida na figura 6, a qual mostra o tamanho dos fragmentos de cada marcador, distribuídos de acordo com a herança alélica do pai ou da mãe. Apesar de nem todos os pacientes apresentarem o mesmo haplótipo para todos os marcadores como era previsto, foi possível determinar um haplótipo mais frequente, o qual foi detectado em oito dos 14 alelos portadores da mutação inédita.

Dentre os cinco pacientes homozigotos, dois apresentaram ambos os alelos com esse haplótipo mais comum (pacientes 1 e 10) (Figura 6), embora apenas o paciente 10 tenha relatado consanguinidade parental ($F=1/256$) (Tabela 2).

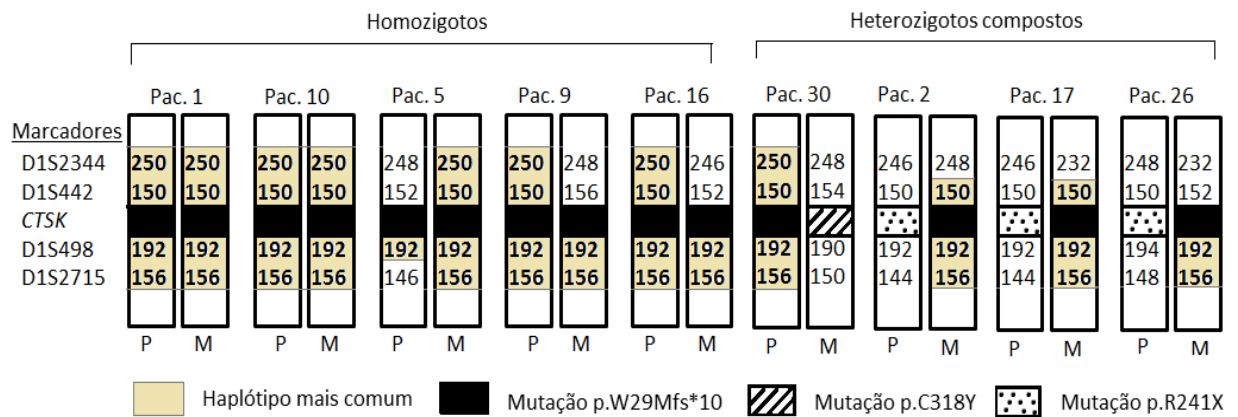


Figura 6. Haplótipo dos nove pacientes portadores da mutação inédita (p.W29Mfs*10). O gene *CTSK* se encontra entre os marcadores D1S442 e D1S498. Pac: Paciente, P: Alelo paterno, M: Alelo materno.

De acordo com a tabela 5, o tamanho mais frequente do marcador D1S2344 foi de 250pb (57,14%) enquanto no marcador D1S442 o tamanho mais frequente foi de 150pb, presente em 71,42% dos alelos com a mutação inédita.

Todos os haplótipos portadores da mutação inédita apresentaram o marcador D1S498 com 192pb, enquanto o marcador D1S2715 apresentou tamanho de 156pb em 13 dos 14 haplótipos portadores da mutação inédita (92,85%).

Tabela 5. Comprimento dos marcadores nos alelos com a mutação inédita.

Comprimento dos Marcadores (pb)	Alelo com mutação Inédita (p.W29Mfs*10)	
D1S2344	Número (Total 14)	Porcentagem (%)
232	2	14,3
246	1	7
248	3	21,42
250*	8	57,14
D1S442		
150*	10	71,42
152	3	21,42
156	1	7
D1S498		
192*	14	100
D1S2715		
146	1	7
156*	13	92,85

*Em destaque estão os comprimentos dos fragmentos de cada marcador encontrado no haplótipo mais comum.

Mesmo não havendo um grupo controle de genotipagem para comparação dos padrões dos marcadores encontrados, a presença dos mesmos comprimentos dos fragmentos, principalmente dos marcadores D1S498 e D1S2715 dentre os indivíduos portadores da mutação inédita fortalece a hipótese de haver um ancestral comum a todos eles.

4.7 Revisão de Fenótipos

A reavaliação dos dados clínicos e radiológicos de todos os pacientes da casuística foi dificultada devido à ausência muitos de dados entre os pacientes de fora. Dos 39 pacientes encaminhados, foram recuperadas todas as documentações e dados clínicos de apenas 12 (30,7%). Dentre os demais pacientes, a documentação radiológica de 17 e documentação fotográfica de 10 estavam incompletas. As informações clínicas 17 indivíduos não foram disponibilizadas. Os dados recuperados e/ou obtidos na reavaliação estão resumidos na tabela 6 abaixo, a qual foi confeccionada com base no artigo de Turan *et al*, 2014 [13] para caracterização clínica e radiológica da Picnodisostose. As alterações odontológicas não estão contempladas na tabela, devido à insuficiência de dados.

Tabela 6. Revisão fenotípica dos pacientes estudados. Essa tabela foi construída baseando-se na reavaliação dos dados clínicos, radiológicos e informações clínicas disponíveis.

Fenótipo Clínico												Fenótipo Radiológico				Fenótipo atípico
Nº Ref. mapa	Genótipo	Baixa estatura	Fraturas	Braquidactília	Unhas displásicas	Proem. ossos frontal /parietal	Hipoplas Face média (maxila)	Hipoplas Mand.	Nariz em bico	Proptose Ocular	Esclera azulada	Dens. óssea aumentada	Osteólise (falanges distais)	Suturas e fontanelas abertas	Hipoplas. distal da clavícula	Outros achados
1	W29Mfs*10/W29Mfs*10	+	+	+	+	+	+	-	+	+	+	+	+	+	?	Ang mand obtuso
5	W29Mfs*10/W29Mfs*10	+	?	+	+	+	+	+	+	+	-	+	+	+	?	
9	W29Mfs*10/W29Mfs*10	+	-	+	+	+	-	-	+	-	-	+	+	+	?	Craniossinostose
10	W29Mfs*10/W29Mfs*10	+	+	+	+	+	+	+	+	+	-	?	?	?	?	
16	W29Mfs*10/W29Mfs*10	+	+	+	+	+	+	-	+	-	-	+	?	+	?	Ang mand obtuso
2	W29Mfs*10/ R241X	+	-	+	+	+	-	+	+	-	-	+	+	+	-	
17	W29Mfs*10/R241X	+	+	+	+	+	-	-	+	+	-	+	+	+	?	Ang man obtuso
26	W29Mfs*10/R241X	+	-	?	-	+	+	+	+	+	+	?	?	?	?	
26i	W29Mfs*10/R241X	+	+	+	?	+	+	+	+	+	+	?	+	?	?	
30	W29Mfs*10/ C318Y	+	?	+	+	+	+	-	+	+	-	+	+	+	-	Osso wormianos, Ang mand obtuso
3	R241X/ R241X	+	+	+	+	+	+	-	+	+	+	+	+	+	?	Ossos Wormianos, Ang mand obtuso
6	R241X/ R241X	+	+	+	+	+	+	+	-	+	+	+	+	+	?	Muitas fraturas
7	R241X/ R241X	+	+	+	?	?	?	+	?	?	?	+	?	+	?	Anemia, Ang mand obtuso, Hepatoesplenomeg, <i>Pectus carinatum</i>
11	R241X/ R241X	+	+	+	+	+	+	+	+	+	-	?	?	?	?	Cardiopatia, Macroglossia
12	R241X/ R241X	+	+	+	+	+	+	+	+	+	-	+	+	+	?	<i>Pectus carinatum</i>
12i	R241X/ R241X	+	?	+	+	+	+	+	+	+	-	?	?	?	?	
23	R241X/R241X	?	?	?	?	?	?	?	?	?	?	+	+	?	?	
24	R241X/R241X	?	+	+	+	+	+	+	+	+	-	+	+	+	?	Ang mand obtuso
32	R241X/R241X	-	+	+	+	+	+	-	+	+	-	+	+	+	?	Laringite, Apnéia obstrutiva, Ang mand obtuso

Nº Ref. mapa	Genótipo	Baixa estatura	Fraturas	Braquidactilia	Unhas displásicas	Proem. ossos frontal /parietal	Hipoplas Face média (maxila)	Hipoplas Mand.	Nariz em bico	Proptose Ocular	Esclera azulada	Dens. óssea aumentada	Osteólise (falanges distais)	Suturas e fontanelas abertas	Hipoplas. distal da clavícula	Outros achados
20	R241X/G146R	-	-	+	-	+	+	+	+	+	-	?	?	+	?	
20i	R241X/G146R	+	+	+	+	+	+	+	+	+	-	?	?	?	?	
29	R241X/G146R	+	+	+	+	+	+	-	-	+	-	+	+	+	+	Ang mand obtuso
29i	R241X/G146R	+	-	+	-	+	-	-	-	+	-	+	-	+	+	Ang mand obtuso
31	R241X/G146R	+	+	+	+	+	+	+	+	+	-	+	+	+	+	Muitas fraturas, Escoliose, Pneumonia, Cardiopatia, Hipermob. Juntas
25	R241X/G146R	+	?	+	-	+	+	?	?	+	+	+	-	?	?	Atraso escolar
8	G146R/ G146R	+	+	+	?	+	+	-	-	+	+	+	+	+	?	Ang mand obtuso
8i	G146R/ G146R	+	+	+	+	+	-	-	-	+	+	+	+	+	?	Ang mand obtuso
13	G146R/ G146R	?	+	?	?	?	?	?	?	?	?	+	?	?	?	Cardiopatia
15	G146R/ G146R	?	+	+	-	+	+	+	+	-	-	+	+	?	?	
18	G146R/ G146R	+	+	+	-	+	+	+	+	+	+	+	?	+	+	Infec. Resp.
19	G146R/G146R	+	+	+	+	+	+	+	-	+	-	+	+	+	?	Ang mand obtuso
27	G146R/G146R	+	+	+	+	+	+	-	+	+	-	+	+	+	?	Ang mand obtuso
14	G146R/ G194S	+	?	+	+	+	+	-	+	+	+	+	+	+	?	Ang mand obtuso, Macroglossia
14i	G146R/ G194S	?	?	+	-	+	+	+	+	+	-	?	?	?	?	
4	C318Y/ C318Y	+	-	+	+	+	+	-	+	+	-	+	+	+	?	Escoliose
21	C318Y/C318Y	+	?	+	-	+	+	-	+	+	-	?	?	?	?	Ang mand obtuso
28	C318Y/C318Y	+	-	+	+	+	+	-	-	+	-	+	+	+	+	Ossos wormianos
22	C318Y/R241X	+	+	+	+	+	?	?	?	?	+	+	+	+	?	Ang mand obtuso
33	A277V/A277V	+	?	+	?	+	+	+	+	+	-	+	+	+	?	Apneia obst.grave (traqueostomia) , craniossinostose
Média		32/34	23/30	36/36	25/33	36/36	30/35	19/33	27/34	31/35	11/36	29/29	25/27	27/27	5/7	

Ref.: Referência, *Idade em que o paciente/responsável assinou o TCLE para realizar o estudo molecular, Proem.: Proeminência, Hipoplas.:Hipoplasia., Mand.: Madibular, Dens.: Densidade, Ang. Mand. Obtuso: Ângulo mandibular obtuso, apnéia obst.: Apnéia obstrutiva, hepatoesplenomeg.: Hepatoesplenomegalia, hipermob.: Hipermobilidade, Infec. Resp.: Infecções respiratórias recorrentes, Macroglos.:Macroglossia, (+) Característica presente,(-) Característica ausente, (?) Característica não revisada por falta de documentação.

4.8 Produção Científica

Este trabalho deu origem ao artigo “*Molecular analysis of the CTSK gene in a cohort of 28 Brazilian families with Pycnodysostosis showing the spreading of the mutations*”, o qual foi submetido à Revista Científica *Orphanet Journal of Rare Diseases* em Dezembro de 2015. O texto na íntegra encontra-se no Apêndice 3.

5. DISCUSSÃO

5.1 A Prevalência da Picnodisostose é subestimada?

A prevalência da Picnodisostose, estimada na literatura, é de 1 a 1,7 /1.000.000 indivíduos. Essa estimativa está baseada no número de casos descritos na literatura [42, 48] e tem sido apenas repetida por outros autores [5, 43–45]. Não há nenhum estudo com desenho epidemiológico adequado voltado especificamente para a Picnodisostose. Além de muitos pacientes não estarem relatados na literatura, também deve haver subdiagnóstico e casos incorretamente diagnosticados [10, 12]. Em um estudo de investigação molecular foram encontradas mutações até então inéditas no gene *CTSK* em seis pacientes previamente diagnosticados com Osteopetrose[12]. Esses pacientes apresentavam características pouco comuns na Picnodisostose como: anemia, comprometimento visual grave, malformação de *Chiari*, esplenomegalia ou acroosteólise das falanges pouco aparente. Portanto, é possível supor que a prevalência real da Picnodisostose seja, de fato, maior do que a estimada.

5.2 A Picnodisostose no Ceará

A alta frequência de pacientes com Picnodisostose no Ceará, somada à alta frequência de casamentos consanguíneos entre as famílias dos afetados motivou esse estudo, que teve como hipótese inicial de trabalho um possível efeito fundador. Entretanto, nas 18 famílias cearenses estudadas (famílias 1 a 18), foram encontradas cinco mutações diferentes, sendo uma delas inédita (p.W29Mfs*10). Apesar de apenas 33,3% das famílias terem referido consanguinidade, a maioria das mutações foram encontradas em homozigose (83,3%), sugerindo que a real taxa de consanguinidade entre essas famílias está mais próxima de 80 % do que 33%, como relatado pelas famílias. Portanto, a presença de diferentes mutações nessa população estudada

invalida a hipótese de efeito fundador e alta taxa de consanguinidade pode, portanto, explicar a alta frequência da Picnodisostose nessa região.

A elevada taxa de uniões consanguíneas no nordeste brasileiro é bem conhecida há décadas [49]. Enquanto a taxa de consanguinidade estimada para todo o Brasil é de 4,8%, no nordeste esta taxa é três vezes maior e em algumas regiões particulares do nordeste, essa taxa pode atingir 30% ou mais [50].

De acordo com a genética clínica, casamentos consanguíneos são definidos como uniões entre indivíduos que são primos em segundo grau, ou possuem parentesco mais próximo, o que significa que o coeficiente de endogamia “F” deve ser igual ou maior do que 1/64 [51]. Casamentos entre consanguíneos cujos ancestrais comuns estão geneticamente distantes ($F < 1/64$), possuem probabilidade mínima de herdarem uma região homozigota por descendência. Entretanto, alelos deletérios autossômicos recessivos podem ser carregados no estado de heterozigose por gerações sem se manifestarem e então, por meio de uma união entre consanguíneos distantes portadores de um alelo deletério, esse alelos podem se combinar no estado de homozigose, manifestando a doença [52]. Visto que a probabilidade de se encontrar um portador de um alelo deletério raro na população em geral é baixa, esse efeito é ainda mais frequente em doenças raras [52]. Dessa forma, ao se examinar famílias com afetados por doenças raras, é comum encontrar uma taxa elevada de consanguinidade parental [52].

Em um estudo semelhante a este realizado em 2014, Arman *et al.* [17] estudaram 14 famílias de uma população altamente endocruzada na Turquia e encontraram sete mutações distintas, todas em homozigose, sendo cinco delas inéditas [17]. Os pesquisadores concluíram que as mutações iniciais, encontradas na Turquia, provavelmente foram segregadas ao longo dos séculos com a ajuda da alta taxa de casamentos consanguíneos. [17].

Assim, de acordo com nossos dados e de modo semelhante ao descrito na literatura, é possível presumir que os alelos mutados, previamente encontrados em outras populações, certamente foram introduzidos no nordeste brasileiro por meio da imigração europeia e africana [53], sendo mantidos e segregados ao longo das gerações devido à alta taxa de endocruzamento existente nessa região.

Quanto à mutação inédita, p.W29Mfs*10, que se mostrou concentrada no noroeste do Ceará (Fig. 2), presume-se que a mesma deve ter surgido ou sido introduzida nessa região, também sendo mantida e segregada por meio de casamentos consanguíneos.

5.3 Picnodisostose em outras regiões do Brasil

Nas 15 famílias provenientes de outros Estados (famílias 19 a 33), incluídas na casuística desse trabalho, cerca de metade é composta por indivíduos heterozigotos compostos, decorrente provavelmente de uma maior miscigenação vista nessas regiões [50]. De acordo com a tabela 3, em todas as famílias, exceto uma (família 33), em que a mutação aparece em homozigose, ambos os genitores são nordestinos.

É Interessante notar que a quase totalidade das mutações encontradas nesses pacientes foram as mesmas encontradas nos pacientes do Ceará. Exceções são a mutação p.G194S, encontrada apenas no Ceará, e a mutação p.A277V, detectada apenas no Rio Grande do Sul. A detecção das mesmas mutações entre os dois grupos de famílias, juntamente com a observação de que os familiares dos probandos desse segundo grupo são oriundos do nordeste, sugerem que todas as mutações encontradas na região Centro-oeste e Sudeste (Fig 3.) podem de fato terem sido introduzidas nesses locais por migração interna da população nordestina.

A migração interna no Brasil, geralmente é motivada por razões de ordem econômica e se dá principalmente a partir dos jovens, que deixam as áreas rurais ou pequenas cidades em busca de melhores oportunidades de trabalho nos grandes centros urbanos, sobretudo do sudeste brasileiro [54]. Considerando que o nordeste possui o menor índice socioeconômico dentre as cinco regiões do Brasil, a migração interna geralmente ocorre do nordeste para o sudeste, sendo São Paulo o Estado mais procurado pelos migrantes [55]. Esse padrão de migração é compatível com o observado no presente estudo, uma vez que a maioria dos pacientes atualmente residindo em São Paulo possui antepassados nordestinos (Fig 3, Tabela 3).

Essa migração interna do nordeste para o sudeste também favorece o aumento da consanguinidade na região de destino [49, 56], fato observado

neste estudo, em que a taxa de consanguinidade parental referida dentre os pacientes provenientes de fora do Ceará foi de 46,6 % (Tabela 2).

5.4 Mutações encontradas e sua origem geográfica

Dentre as mutações mais frequentemente encontradas na casuística deste trabalho, estão p.G146R e p.R241X, as quais aparecem tanto em homozigose como em combinação entre si ou com outras três mutações (Tabela 4). De acordo com a literatura, essas mesmas mutações, juntamente com a mutação p.A277V são as mutações mais frequentemente encontradas nos pacientes com Picnodisostose [5].

A mutação p.G146R (c.436G>C, éxon 5) foi, juntamente com a mutação p.R241X (c. 721C>T, éxon 6), uma das primeiras a serem descritas no estudo molecular pioneiro realizado por Gelb *et al.* no ano de 1996, logo após a descoberta da associação do gene *CTSK* com a Picnodisostose [15].

Na casuística estudada por Gelb, a mutação p.G146R foi encontrada em homozigose em dois irmãos de uma família consanguínea de origem árabe marroquina e em combinação com a mutação p.R241X em um paciente heterozigoto composto de origem américo-hispânica e família não consanguínea [15].

Posteriormente foram relatados casos da mutação p.G146R em pacientes originários da Tunísia [25, 38], Brasil [36] e México [57], enquanto a mutação p.R241X foi descrita em pacientes de origem italiana [25], portuguesa [25, 29], espanhola [25, 26, 29], canadense [26], mexicana [8] e brasileira [36].

No presente trabalho, todos os pacientes, exceto dois (26 e 31), portadores dessas duas mutações, p.G146R e p.R241X, possuem antepassados nordestinos de ambos os lados, materno e paterno (Figura 3). O paciente 26 que herdou a mutação p.R241X de seu pai, tem antepassados portugueses. O paciente 31, portador de ambas as mutações, p.G146R e p.R241X, tem antepassados portugueses e espanhóis (Tabela 3).

A mutação p.C318Y (c.953G>A, éxon 8), descoberta em pacientes brasileiros [36] e posteriormente relatada em Portugal em um único paciente de origem africana [58], foi detectada nesse estudo em um paciente do Ceará (Figura 2) e quatro pacientes provenientes de diferentes Estados, um do

Maranhão, um da Paraíba e dois de São Paulo (Figura 3). Esses dois pacientes do Estado de São Paulo (31 e 32), possuem antepassados nordestinos (Tabela 3). Assim, é provável que esta mutação tenha sido introduzida no nordeste por meio de imigrantes portugueses e posteriormente tenha se dispersado devido à migração dessas famílias dentro do próprio nordeste e também para o sudeste.

Outras duas mutações p.G194S (c.580G>A, éxon 5) e p.A277V (c.935C>T, éxon 7) foram encontradas uma única vez nesta coorte de pacientes. A primeira foi vista em combinação com a p.G146R no paciente 14 e seu irmão, provenientes de Iguatu, no Ceará (Figura 3). As informações sobre a origem e antepassados dessa família não estavam disponíveis, mas como essa mutação foi descrita apenas em pacientes italianos até o momento [25], é possível que os antepassados paternos tenham origem italiana, uma vez que o paciente herdou esta mutação de seu pai (Tabela 3).

A mutação p.A277V foi encontrada em homozigose apenas na paciente 33, filha de um casal consanguíneo do Rio Grande do Sul com antepassados Italianos (Figura 3). Essa mutação foi primeiramente descrita em homozigose por Gelb *et al.*, 1998 [33] em um indivíduo de origem Belga e de família não consanguínea. Curiosamente nesse caso, a mutação foi herdada por dissomia uniparental parterna [33]. A posteriori, essa mutação foi relatada em indivíduos de origem japonesa [19, 26, 37] e paquistanesa [25, 42]. Tendo vista a detecção aparentemente isolada dessa mutação no sul do país, pode-se especular que ela tenha surgido a partir de uma mutação *de novo* num ancestral comum e passou despercebida por gerações até ser encontrada em homozigose pelo casamento consanguíneo. Ou pode ter sido introduzida por um imigrante italiano há algum tempo atrás, apesar de não ter sido relatado nesse país.

A mutação inédita p.W29Mfs*10 foi observada em homozigose em cinco pacientes do Ceará; em combinação com a mutação mais frequente, p.R241X, em dois pacientes do Ceará e um de Goiás e em combinação com a mutação p.C318Y em um paciente de Guarulhos, este último com antepassados provenientes de Juazeiro do Norte, Ceará (Tabela 3). A concentração dessa mutação, na maioria das vezes em homozigose, no noroeste do Ceará com aparente dispersão para outros estados, reforçou a

hipótese de efeito fundador, justificando a investigação do estudo de haplótipos.

5.5 Análise de haplótipos ligados à mutação inédita

A análise de haplótipos realizada em nove pacientes, homozigotos e heterozigotos compostos para a mutação inédita, demonstrou haver um haplótipo comum, presente em oito dos 14 alelos mutados (Fig. 5), que segrega juntamente com esta mutação, sustentando a hipótese de haver um ancestral comum.

No primeiro estudo de haplótipos realizado em 1995, visando investigar a localização do gene responsável pelo fenótipo de Picnodisostose, Gelb e colaboradores delimitaram uma região homozigota por descendência entre os marcadores D1S442 (159,6 cM) e D1S305 (164,1 cM) no cromossomo 1 [14]. Posteriormente esse *locus* foi modificado para a região entre os marcadores D1S2344 (158,5 cM) e D1S2345 (160,8 cM) pelos mesmos autores [15, 59].

Entretanto, outros pesquisadores como Johnson *et al.*, 1996 [8], Haagerup *et al.*, 2000 [35] e Naeem *et al.*, 2009 [42], mostraram que a provável posição do gene *CTSK* é na primeira região proposta, entre os marcadores D1S442 (159,6 cM) e D1S305/D1S2715 (164,1 cM). De acordo com esses estudos, o marcador D1S2344 estaria menos ligado ao gene do que os outros marcadores citados, por esse motivo teria maior probabilidade de sofrer recombinação do que os demais. De fato, no subgrupo de pacientes estudados nesse trabalho, o marcador D1S2344 foi o que mais apresentou variação dentre os alelos portadores da mutação inédita (Tabela 5). Curiosamente, o padrão de 232pb para este marcador só foi encontrado nos pacientes 17 e 26 (Figura 5), os únicos que não possuem antecedentes Cearenses, mas herdaram a mutação inédita da família materna originária do Rio Grande do Norte (Tabela 3).

Enquanto nas análises de Johnson *et al.*, 1996 [8] o marcador D1S442 apresentou recombinação, não sendo considerado como parte integrante do haplótipo, em outros trabalhos este marcador não apresentou variação entre indivíduos de uma mesma família, sendo considerado informativo [35, 42, 59].

Na nossa avaliação, esse marcador apresentou três tamanhos de fragmentos diferentes, sendo o mais frequente o de comprimento de 150 pb (Tabela 5).

O marcador D1S498 (160,7cM), por outro lado, se mostrou o mais informativo dentre os vários marcadores utilizados em diversos estudos de haplótipos no gene *CTSK*, não apresentando nenhuma recombinação em amostras de indivíduos da mesma família [8, 35, 42, 59], o que sugere que a localização do gene seja bem próxima a esse marcador. Resultado semelhante foi observado na nossa avaliação, pois todos os alelos portadores da mutação inédita apresentaram o mesmo comprimento para este marcador (192pb) (Tabela 5).

Considerando os haplótipos portadores da mutação inédita na presente casuística, apenas um não apresentou o mesmo padrão de comprimento para o marcador D1S2715 (Tabela 5). Dentre os quatro autores citados, todos indicam que este marcador segrega em conjunto com o gene *CTSK* [8, 35, 39, 42], embora esteja mapeado pelo *Genethon* a uma distância gênica maior do gene *CTSK* do que os marcadores D1S2344 e D1S442 (Figura 5).

Ao analisarmos a distância física desses marcadores e do gene *CTSK*, dada pelo *Ensembl*, percebemos que esta se aproxima mais dos resultados obtidos neste estudo (Figura 5). De acordo com o mapeamento físico, os marcadores D1S2344 e D1S442, localizados acima do gene *CTSK*, estão mais distantes do que os marcadores D1S498 e D1S2715, localizados logo abaixo do gene. Assim, sugerimos que a heterogeneidade de comprimento vista nos marcadores D1S2344 e D1S442 muito provavelmente é resultado da ocorrência de recombinações ao longo das gerações descendentes de um ancestral comum e que os marcadores D1S498 e D1S2715 estão segregando em conjunto com a mutação inédita, com uma taxa de recombinação menor.

No conjunto, esses dados fortalecem a hipótese de efeito fundador e ainda sugerem uma relação de parentesco distante entre os indivíduos afetados, reforçando a alta taxa de endocruzamento dessa população e favorecendo a interpretação de parentesco distante desconhecido pelas famílias, uma vez que entre os indivíduos genotipados, apenas o paciente 10 relatou consanguinidade parental ($F=1/256$, Tabela 2).

5.6 Análise fenotípica

A revisão da literatura mostra que muitos autores não especificam o fenótipo clínico-radiológico dos indivíduos com Picnodisostose. Dentre as 50 mutações descritas na literatura [12, 15, 17–38], 18 delas não podem ser avaliadas quanto a uma possível correlação genótipo-fenótipo por falta de caracterização fenotípica desses pacientes [25, 26, 29, 35, 37, 39]. Felizmente, as cinco mutações já descritas encontradas em nossa casuística possuem uma caracterização fenotípica, ainda que breve.

Os pacientes descritos na literatura [13, 19, 33, 36, 37, 42] portadores das mesmas mutações presentes nos pacientes analisados neste trabalho apresentam fenótipo bastante similar entre si (Tabela 6).

Nem todas as características dos pacientes que compõem a presente casuística puderam ser bem avaliadas devido à falta de dados. Contudo, as características típicas da Picnodisostose [13], como baixa estatura, braquidactilia, proeminência dos ossos frontal e parietal, hipoplasia da face média, proptose do olho, densidade óssea aumentada, osteólise das falanges distais e suturas e fontanelas amplas estão presentes em mais de 90% dos pacientes revisados, independentemente do tipo de mutação.

A ocorrência de fraturas, unhas displásicas e nariz em bico também foram características muito frequentes, sendo vistas em mais de 70% dos pacientes, enquanto o ângulo mandibular obtuso foi visto em 63% dos pacientes e hipoplasia da mandíbula em 54% (Tabela 6). Essas frequências são semelhantes às descritas na literatura [5]. Embora a frequência de braquidactilia, osteólise das falanges distais, unhas displásicas e nariz em bico seja menor nos pacientes revisados por Xue *et al.*, 2011 [5], essa diferença é devido à ausência de informações e não à ausência dessas características na casuística.

A esclera azulada, presente em 30% dos pacientes analisados (Tabela 6) é, de fato, uma característica menos comum nos pacientes com Picnodisostose, de acordo com a literatura [10]. Infecções respiratórias, apnéia obstrutiva e hipoplasia distal da clavícula que ocorrem com frequência, segundo a literatura [5, 10, 11], talvez tenham sido subestimadas em nossa casuística devido à falta de informações clínicas e radiografias apropriadas.

Alguns achados atípicos como *Pectus carinatum*, ossos wormianos e macroglossia foram vistos em mais de um paciente, entretanto, não parecem estar ligados a nenhuma mutação específica (Tabela 6).

Ademais, de acordo com a literatura, a grande maioria dos pacientes afetados por Pichodisostose possui complicações odontológicas devido à hipoplasia de face média, micrognatia, palato em ogiva ou ângulo mandibular obtuso [10], entretanto, as alterações dentárias frequentemente não são relatadas nos prontuários médicos.

7. CONCLUSÕES

Em síntese, os resultados obtidos nessa investigação, nos permitem as seguintes conclusões:

- ✓ A alta frequência de Picnodisostose no estado do Ceará se deve principalmente a alta taxa de endocruzamento da população
- ✓ As quatro mutações encontradas no Ceará, e previamente descritas, podem ter sido introduzidas por imigrantes que colonizaram o nordeste, tendo as mesmas se disseminado pela alta taxa de endocruzamento local.
- ✓ A maior parte das mutações observadas nas outras regiões do Brasil, também observadas no CE, podem ter se dispersado pelo processo de migração interna da população (sobretudo pela migração nordeste-sudeste), sendo reforçadas pelo endocruzamento entre esses migrantes.
- ✓ É possível que a mutação inédita, p.W29Mfs*10, tenha uma origem única no estado do CE e a sua disseminação esteja ocorrendo pela alta taxa de endocruzamento.
- ✓ O estudo de haplótipos de indivíduos com a mutação inédita, p.W29Mfs*10, sugere, de fato, a existência de um ancestral comum e, portanto, de um efeito fundador para essa mutação.
- ✓ A alta frequência das mutações em homozigose entre os afetados por Picnodisostose oriundos do Ceará sugere que a taxa real de consanguinidade seja bem maior que a referida pelas famílias – 80% ao invés de 33%.
- ✓ Os fenótipos dos indivíduos com Picnodisostose são bastante similares, independentemente do tipo de mutação no gene *CTSK*.

8. REFERÊNCIAS BIBLIOGRÁFICAS

1. Barbosa-Buck O, Orioli M, Grac M, Lopez-camelo J, Castilla E, Cavalcanti D: **Clinical Epidemiology of Skeletal Dysplasias in South America Selection and Classification of Patients.** *Am J Med Genet A* 2012:1–8.
2. Stevenson DA, Carey JC, Byrne JLB, Srisukhumbowornchai S, Feldkamp ML: **Analysis of skeletal dysplasias in the Utah population.** *Am J Med Genet A* 2012, **158A**:1046–54.
3. Bonafe L, Cormier-Daire V HC, Lachman R, Mortier G MS, Nishimura G, Sangiorgi L SR, Sillence D, Spranger J S-FA, Warman M US: **Nosology and Classification of Genetic Skeletal Disorders: 2015 Revision.** *Am J Med Genet Part A* 2015, **9999A**:1–24.
4. Warman ML, Cormier-Daire V, Hall C, Krakow D, Lachman R, LeMerrer M, Mortier GM, Mundlos S, Nishimura G, Rimoin DL, Robertson S, Savarirayan R, Sillence D, Spranger J, Unger S, Zabel B S-FA: **Nosology and classification of genetic skeletal disorders: 2010 revision.** *Am J Med Genet A* 2010, **155**:943–968.
5. Xue Y, Cai T, Shi S, Wang W, Zhang Y, Mao T, Duan X: **Clinical and animal research findings in pycnodysostosis and gene mutations of cathepsin K from 1996 to 2011.** *Orphanet J Rare Dis* 2011, **6**:20.
6. Elmore SM; Richmond MD: **Pycnodysostosis : A Review.** *J Bone Jt Surg Am* 1967, **49**:153–162.
7. Hodder A, Huntley C, Aronson JK, Ramachandran M: **Pycnodysostosis and the making of an artist.** *Gene* 2015, **555**:59–62.
8. Johnson M, Polymeropoulos M, Vos H, Ortiz de Luna R, Francomano C: **A nonsense mutation in the cathepsin K gene observed in a family with pycnodysostosis.** *Genome Res* 1996:1051–1055.
9. Motyckova G, Fisher DE: **Pycnodysostosis: role and regulation of cathepsin K in osteoclast function and human disease.** *Curr Mol Med* 2002,

2:407–21.

10. Alves N, Cantín M: **Clinical and radiographic maxillofacial features of pycnodysostosis.** *Int J Clin Exp Med* 2014, **7**:492–496.
11. Testani E, Scarano E, Leoni C, Dittoni S, Losurdo A, Colicchio S, Gnoni V, Vollono C, Zampino G, Paludetti G, Della Marca G: **Upper airway surgery of obstructive sleep apnea in pycnodysostosis: Case report and literature review.** *Am J Med Genet A* 2014, **164**:2029–35.
12. Pangrazio A, Puddu A, Oppo M: **Exome sequencing identifies CTSK mutations in patients originally diagnosed as intermediate osteopetrosis.** *Bone* 2014, **59**:122–126.
13. Turan S: **Current research on pycnodysostosis.** *Intractable rare Dis Res* 2014, **3**:91–3.
14. Gelb B, Edelson J, Desnick R: **Linkage of pycnodysostosis to chromosome 1q21 by homozygosity mapping.** *Nat Genet* 1995, **173**:1236–1238.
15. Gelb B, Shi G, Chapman H, Desnick R: **Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency.** *Science* 1996, **273**:1236–8.
16. Novinec M, Lenarčič B: **Cathepsin K: a unique collagenolytic cysteine peptidase.** *Biol Chem* 2013, **394**:1163–79.
17. Arman A, Bereket A, Coker A, Kiper PÖS, Güran T, Ozkan B, Atay Z, Akçay T, Haliloglu B, Boduroglu K, Alanay Y, Turan S: **Cathepsin K analysis in a pycnodysostosis cohort: demographic, genotypic and phenotypic features.** *Orphanet J Rare Dis* 2014, **9**:60.
18. Zheng H, Zhang Z, He JW, Fu WZ, Zhang ZL: **A novel mutation (R122Q) in the cathepsin K gene in a Chinese child with Pyknodysostosis.** *Gene* 2013, **521**:176–179.
19. Matsushita M, Kitoh H, Kaneko H, Mishima K, Itoh Y, Hattori T, Ishiguro N: **Novel Compound Heterozygous Mutations in the Cathepsin K Gene in Japanese Female Siblings with Pyknodysostosis.** *Mol Syndromol* 2012, **8550**:254–258.

20. Utokpat P, Panmontha W, Tongkobpetch S, Suphapeetiporn K, Shotelersuk V: **Novel CTSK mutation resulting in an entire exon 2 skipping in a Thai girl with pycnodysostosis.** *Pediatr Int* 2013, **55**:651–5.
21. Ozdemir T, Atik T, Karaca E, Onay H, Ozkinay F, Cogulu O: **A novel mutation in two families with pycnodysostosis.** *Clin Dirmophol* 2013, **22**:102–5.
22. Singh A, Cuevas-covarrubias S, Pradhan G, Gautam VK: **Novel Mutation and White Matter Involvement in an Indian Child with Pycnodysostosis.** *Indian J Pediatr* 2014.
23. Xue Y, Wang L, Xia D, Li Q, Gao S, Dong M, Cai T, Shi S, He L, Hu K, Mao T, Duan X: **Dental Abnormalities Caused by Novel Compound Heterozygous CTSK Mutations.** *J Dent Res* 2015, **94**:674–681.
24. Alhashem A, Alswaid A bin MA: **Pycnodysostosis with novel gene mutation in a saudi family.** In *12th International Skeletal Dysplasia Society Meeting*. Istanbul, Turkey; 2015.
25. Donnarumma M, Regis S, Tappino B, Rosano C, Assereto S, Corsolini F, Di Rocco M, Filocamo M: **Molecular analysis and characterization of nine novel CTSK mutations in twelve patients affected by pycnodysostosis. Mutation in brief 961.** *Hum Mutat* 2007, **28**:524.
26. Nishi Y, Atley L, Eyre DE, Edelson JG, Superti-Furga a, Yasuda T, Desnick RJ, Gelb BD: **Determination of bone markers in pycnodysostosis: effects of cathepsin K deficiency on bone matrix degradation.** *J Bone Miner Res* 1999, **14**:1902–1908.
27. Schilling a F, Mülhausen C, Lehmann W, Santer R, Schinke T, Rueger JM, Amling M: **High bone mineral density in pycnodysostotic patients with a novel mutation in the propeptide of cathepsin K.** *Osteoporos Int* 2007, **18**:659–69.
28. Fratzl-Zelman N, Valenta A, Roschger P, Nader A, Gelb BD, Fratzl P, Klaushofer K: **Decreased bone turnover and deterioration of bone structure in two cases of pycnodysostosis.** *J Clin Endocrinol Metab* 2004, **89**:1538–47.
29. Hou WS, Brömme D, Zhao Y, Mehler E, Dushey C, Weinstein H, Miranda

CS, Fraga C, Greig F, Carey J, Rimoin DL, Desnick RJ, Gelb BD:

Characterization of novel cathepsin K mutations in the pro and mature polypeptide regions causing pycnodysostosis. *J Clin Invest* 1999, **103**:731–738.

30. Chavassieux P, Asser Karsdal M, Segovia-Silvestre T, Neutzsky-Wulff A V, Chapurlat R, Boivin G, Delmas PD: **Mechanisms of the anabolic effects of teriparatide on bone: insight from the treatment of a patient with pycnodysostosis.** *J Bone Miner Res* 2008, **23**:1076–1083.

31. Li H-Y, Ma H-W, Wang H-Q, Ma W-H: **Molecular Analysis of a Novel Cathepsin K Gene Mutation in a Chinese Child with Pycnodysostosis.** *J Int Med Res* 2009, **37**:264–269.

32. Khan B, Ahmed Z, Ahmad W: **A novel missense mutation in cathepsin K (CTSK) gene in a consanguineous Pakistani family with pycnodysostosis.** *J Investig Med* 2010, **58**:720–4.

33. Gelb B, Willner J, Dunn T: **Paternal uniparental disomy for chromosome 1 revealed by molecular analysis of a patient with pycnodysostosis.** *Am J Hum Genet* 1998, **62**:848–854.

34. Toral-López J, Gonzalez-Huerta LM, Sosa B, Orozco S, González HP, Cuevas-Covarrubias SA: **Familial pycnodysostosis: identification of a novel mutation in the CTSK gene (cathepsin K).** *J Investig Med* 2011, **59**:277–80.

35. Haagerup a, Hertz JM, Christensen MF, Binderup H, Kruse T a: **Cathepsin K gene mutations and 1q21 haplotypes in at patients with pycnodysostosis in an outbred population.** *Eur J Hum Genet* 2000, **8**:431–436.

36. Bertola D, Amaral C, Kim C, Albano L, Agüena M, Passos-Bueno MR: **Craniosynostosis in pycnodysostosis: broadening the spectrum of the cranial flat bone abnormalities.** *Am J Med Genet A* 2010, **152A**:2599–603.

37. Fujita Y, Nakata K, Yasui N, Matsui Y, Kataoka E, Hiroshima K, Shiba RI, Ochi T: **Novel mutations of the cathepsin K gene in patients with pycnodysostosis and their characterization.** *J Clin Endocrinol Metab* 2000, **85**:425–31.

38. Tinsa F, Hamouda S, Bellalah M, Bousnina D, Karbout L, Boussetta K, S B: **Unusual feature of pycnodysostosis: pectus carinatum.** *La Tunisi Médicale* 2014, **92**:180–181.
39. Gelb B, Moissoglu K, Zhang J, Martignetti J, Brömme D, Desnick R: **Cathepsin K: isolation and characterization of the murine cDNA and genomic sequence, the homologue of the human pycnodysostosis gene.** *Biochem Mol Med* 1996, **59**:200–6.
40. <http://www.orpha.net>, último acesso em 05/12/2015.
41. http://bvsms.saude.gov.br/bvs/saudelegis/gm/2014/prt0199_30_01_2014.html, último acesso em 05/12/2015
42. Naeem M, Sheikh S, Ahmad W: **A mutation in CTSK gene in an autosomal recessive pycnodysostosis family of Pakistani origin.** *BMC Med Genet* 2009, **10**:76.
43. Mujawar Q, Naganoor R, Patil H, Thobbi AN, Ukkali S, Malagi N: **Pycnodysostosis with unusual findings: a case report.** *Cases J* 2009, **2**:6544.
44. Hernández-Alfaro F, Arenaz Búa J, Serra Serrat M, Mareque Bueno J: **Orthognathic surgery in pycnodysostosis: A case report.** *Int J Oral Maxillofac Surg* 2011, **40**:110–113.
45. Karakurt L, Yilmaz E, Belhan O, Serin E: **Pycnodysostosis associated with bilateral congenital pseudarthrosis of the clavicle.** *Arch Orthop Trauma Surg* 2003, **123**:125–7.
46. Araujo TF, Ribeiro EM, Borges AM CD: **Picnosisostose – Concentração de famílias com alta taxa de consanguinidade no Estado do Ceará sugere Efeito Fundador.** In *XXVI Congresso Brasileiro de Genética Médica*. Fortaleza, Ceará; 2014.
47. Provine WB: **Ernst Mayr: Genetics and speciation.** *Genetics* 2004, **167**:1041–1046.
48. Sambrook J GM: *Molecular Cloning: A Laboratoty Manual*. 4th edition. New

York: Cold Spring Harbor Laboratory Press; 2012.

49. Freire-Maia N: **Genetic effects in Brazilian populations due to consanguineous marriages**. *Am J Med Genet* 1990, **35**:115–117.
50. Santos S, Kok F, Weller M: **Inbreeding levels in Northeast Brazil: strategies for the prospecting of new genetic disorders**. *Genet Mol Biol* 2010, **223**:220–223.
51. Bittles A: **Consanguinity and its relevance to clinical genetics**. *Clin Genet* 2001, **60**:89–98.
52. Jaouad IC, Elalaoui SC, Sbiti A, Elkerh F, Belmahi L, Sefiani A: **Consanguineous marriages in Morocco and the consequence for the incidence of autosomal recessive disorders**. *J Biosoc Sci* 2009, **41**:575–581.
53. **colonial brazil** [<http://thebrazilbusiness.com/article/history-of-colonial-brazil>]
54. Sahota GS: **An Economic Analysis of Internternal migration in Brazil is usually motivated by economic issues in which young people leave rural areas and small cities in search of job opportunities in major citienal Migration in Brazil**. *J Polit Econ* 1968, **76**:218 – 245.
55.
http://www.abep.nepo.unicamp.br/docs/anais/outros/5EncNacSobreMigracao/public_mig_div_bra.pdf, último acesso em 10/11/2015.
56. Araujo CM CD: **Consanguinidade parental em uma amostra de recém-nascidos do CAISM**. In *21° Congresso Interno de Iniciação Científica da Unicamp*. Campinas; 2003.
57. Valdes-flores M, Hidalgo-bravo A, Chima-galan C, Hazan-lasri EJ, Pineda-gomez E, Lopez-estrada D, Zenteno JC: **Molecular and clinical analysis in a series of patients with Pyknodysostosis reveals some uncommon phenotypic findings**. 2014, **7**:3915–3923.
58. Girbal I, Nunes T, Medeira a., Bandeira T: **Pycnodysostosis with novel gene mutation and severe obstructive sleep apnoea: management of a complex case**. *Case Reports* 2013, **2013**:bcr2013200590–bcr2013200590.
59. Gelb BD, Spencer E, Obad S, Edelson GJ, Faure S, Weissenbach J,

Desnick RJ: **Pycnodysostosis: refined linkage and radiation hybrid analyses reduce the critical region to 2 cM at 1q21 and map two candidate genes.** *Hum Genet* 1996, **98**:141–4.

9. APÊNDICES

Apêndice 1: Tabela 7. Primers construídos para estudo molecular do gene *CTSK*.

<i>Primer</i>	Sequência de bases	Tamanho do <i>primer</i>	Tamanho do amplicon
2F	GTCCTTGGAACCAGATGTAC	20pb	417pb
2R	CATGAGTTAGGGAAGAGGG	19pb	-
34F	CTCTTCTTGCAGGCTCTTAATTC	23pb	636pb
34R	GATGTACCTTAATTCCTTGCCC	22pb	-
5F	CAACAGATATTAAGCGGAAGAG	22pb	518pb
5R	CTTCTTGGCCTTCTGTGG	18pb	-
67F	CTGCTGCCTCTGTTAGTTCAC	21pb	744pb
67R	GGGACAGAGAAAGGAATATCG	21pb	-
8F	CTTCCCTCCTTTGATACCTAC	22pb	448pb
8R	GGAAGGATCATTGAAGCAC	20pb	-

A tabela acima mostra os *primers* construídos, bem como seu tamanho e tamanho do amplicon gerado. O éxon 1 não é codificante, por esse motivo não foram construídos *primers* que o flanqueassem. Para os éxons 3 e 4 bem como para os éxons 6 e 7 foram construídos apenas um par de *primers*, uma vez que estes são formados por sequências pequenas e puderam ser unidos em um único fragmento, menor do que 750 pares de base. Todos os pares de primers foram padronizados com tampão KCl e à temperatura ideal de anelamento de 54,1°C.

Apêndice 2. Tabela 8. Primers referentes aos marcadores utilizados no estudo de haplótipos.

Marcador	Sigla sinônima	Artigo referência	Localização no Ensembl	Primer F	Primer R	Tamanho do amplicon (pb)
D1S2344	Z52885	Gelb et al., 1998	1:145.902.878 -145.903.132	TCATGGGACTCTCCATCA	AAATACTCAGGAAATGGCCTA	231-257
D1S442	Z23730	Naeem et al., 2009	1:145.804.416 -145.804.577	AACAAAGCTGGACTGGTAATC	CAGTGTCACACAACTGGTTG	162
Gene CTSK			1:150.796.208 -150.808.323			
D1S498	Z24441	Naeem et al., 2009	1:150.796.208 -150.808.323	TTGCTGAAGGGACATAGTG	TGCTGGGTTATATCCAATATC	183-205
D1S2715	Z53023	Johnson et al., 2015	1:153.611.881 -153.612.032	CACAGGATTCTGCGTCTAACT	TGCTCCAAGAACTGAAGTGA	150-168

]

Apêndice 3. Manuscrito submetido à Revista *Orphanet Journal of rare diseases*

Molecular analysis of the *CTSK* gene in a cohort of 28 Brazilian families with Pycnodysostosis showing the spreading of the mutations from a highly inbreeding region

Thaís Fenz Araujo¹, Erlane Marques Ribeiro^{2,3}, Anderson Pontes Arruda³, Carolina Araujo Moreno^{1,4}, Paula Frassinetti Vasconcelos de Medeiros⁵, Renata Moldenhauer Minillo⁶, Débora Gusmão Melo⁷, Chong Ae Kim⁸, Maria Juliana Rodovalho Doriqui⁹, Têmis Maria Felix¹⁰, Rodrigo Ambrosio Fock¹¹, Denise Pontes Cavalcanti^{1,4*}

* Corresponding author: denisepcavalcanti@gmail.com

¹ Skeletal Dysplasia Group, Department of Medical Genetics, Faculty of Medical Sciences, State University of Campinas, Campinas, SP, Brazil,

Abstract

Background: Pycnodysostosis is an AR sclerosing skeletal dysplasia. Despite its estimated low prevalence (1 per million), in the last years we have found 27 affected individuals from 22 families in the Ceará State a region from the Brazilian Northeast, with a population of 8.904.459 inhabitants, giving a local prevalence of 3 per million. In addition to this high prevalence, 13 families referred consanguinity, suggesting a possible founder effect. This hypothesis prompted us to perform a molecular investigation of these families to test it.

Methods: The coding exons and exon/intron boundaries of *CTSK* gene from patients and their parents were amplified and sequenced by the Sanger method. Besides the families from the Ceará State, we also included 12 families from other regions and searched the origin of each family from the birth locality of the parents and grandparents.

Results: We have studied 34 patients from 28 families with Pycnodysostosis, being 19 patients from 16 families from Ceará State, and 15 patients from 12 families from the other regions. We identified six mutations, five previously described (G146R, G194S, R241X, A277V and C318Y), and a novel mutation (W29Mfs*10). This frameshift variant seems to be originated in a small area and begin to spread. The most of the mutations were found in homozygosity, showing a real consanguinity rate of 87.5% in the Ceará State and of 50% for the other regions. The prevalence of Pycnodysostosis was demonstrated to be a consequence of the high inbreeding rather than a founder effect. We also have demonstrated that the majority of the mutation found in the families the other regions were coming from Ceará State or from other Brazilian Northeast regions.

Conclusions: The high frequency of Pycnodysostosis in the Ceará State is a consequence of the high inbreeding in that region. The novel mutation seems to spread as founder effect. We believe the novel mutation arose in Northeast Brazil, while the other ones should be introduced in the same region, a long time ago, and then, spread due to consanguineous marriages and internal population migration.

Keywords: Cathepsin K, Pycnodysostosis, Inbreeding, novel mutation

Background

Pycnodysostosis (OMIM 265800) is a rare autosomal recessive disorder, classified into the skeletal dysplasias with increase of the bone density [1]. Beyond the altered bone density, this skeletal dysplasia is characterized by disproportioned short stature associated with a typical clinical phenotype that includes a peculiar facial appearance and hand brachydactyly with short distal phalanges [2, 3]. In addition to the increased bone density, the radiological findings include parietal and frontal bossing, delayed closure of sutures and fontanelles, obtuse mandibular angle, susceptibility to fractures and acroosteolysis of the terminal phalanges [2, 3]. Hypoplasia of the maxilla, dental crowding, groove palate, nails hypoplasia, hypo/aplasia of clavicle and recurrent respiratory infections consist other signs seen in many patients [4, 5]. Lifespan, intelligence and sexual development are often normal [3].

Pycnodysostosis was first described in 1962 by Maroteaux and Lamy [2] but its molecular basis was only elucidated in 1996 by Gelb and collaborators [6]. Studying a large consanguineous Israeli Arab family, these last authors found a common genomic region homozygous-by-descent for all 16 affected individuals by Pycnodysostosis at the chromosome 1 (1q21)[6]. Among the genes located in this region, the lysosomal cysteine cathepsin K (*CTSK*) was considered a strong candidate because the encoded protein – lysosomal cysteine protease, is highly expressed in osteoclasts [6]. Sequencing of the *CTSK* gene in the great majority of individuals with a suggestive phenotype of pycnodysostosis has evidenced mutation in this gene [6–26].

The *CTSK* gene has approximately 12 kb (Transcript ID: ENST00000271651.7) and contains 8 exons (GenBank NM_000396.3). The codon for the translation initiator methionine (ATG) is located in exon 2, whereas the termination codon (TGA) is located in middle of exon 8 [2] (Figure 1). This gene encoding a powerful peptidase called lysosomal cysteine Cathepsin K, that is involved in many physiological processes [3]. One of these processes stands out by its role in the bone remodeling due to the organic bone matrix degradation through the potent collagenolytic action in the collagen type I, the major constituent of the bone matrix [2, 3]. It is

also known that the low expression of Cathepsin K is correlated with a defective bone resorption leading to the increase of the bone density and abnormal skeletal growth in patients with Pycnodysostosis [6, 27].

Despite the absence of an specific epidemiologic study, some authors have estimated the Pycnodysostosis prevalence between 1 to 1.7 per million based on the number of cases described in the literature [2, 28–30].

In spite of the rarity of the Pycnodysostosis, in the last years we have found 27 affected individuals from 22 families (59% consanguineous), all of them living in the Ceará State, a region with a population of 8,904,459 inhabitants located in the Brazilian Northeast, giving a local prevalence of 3 per million. This high frequency of Pycnodysostosis, prompted us to perform a molecular investigation of the *CTSK* gene in these families, under the work hypothesis of a possible founder effect. For this survey we were able to study a total of 19 patients from Ceará State and also 15 patients from other Brazilian regions. We tried to identify the precise geographic area from where each family came from. In addition, we performed the sequencing of the *CTSK* gene of all affected individuals and their respective parents. We have found five previously described mutations (G146R, G194S, R241X, A277V, C318Y) and one novel mutation (W29Mfs*10). Rather than a founder effect in the Ceará State, the high rate of inbreeding was found as the cause of the high frequency of Pycnodysostosis in that region. This study also suggests that the majority of the *CTSK* mutations found in the other Brazilian regions were originated from those found in the region northeast. Except for the novel mutation, the other already known mutations should be introduced by the European and African immigrants at long time ago.

Methods

Patients

According to the clinical and radiological criteria suggestive of Pycnodysostosis and revised by one of us (Cavalcanti DP), a total of 34 patients from 28 families were studied by the direct bidirectional sequencing by the Sanger method. Nineteen individuals (16 families) were from Ceará State, and 15 individuals (12 families) were from other Brazilian regions. The molecular analyses included the respective parents and also eventual affected siblings. In addition to the clinical and radiological data from each family, information regarding the birth locality of the respective parents and their respective grandparents were investigated.

This study had the approbation of the Ethic Committee and the written informed consent was obtained from each patient and/or of their respective families.

Molecular genetic analysis

The genomic DNA was extracted by the standard phenol-chloroform protocol from peripheral blood samples obtained from patients and their relatives. The exons and exon-intron boundaries were amplified by polymerase chain reaction (PCR) using the primers described in the table S1 of the supplementary information. PCR reactions were performed in a presence of 50 ng of genomic DNA, 1X KCl buffer, 0,15U taq polymerase, 200 μ M each deoxynucleotide triphosphate (dNTP- Thermo Scientific, 0,5 pmol each primer and water to final volume of 10 μ L. The PCR products were visualized on 1,5% agarose gels. The amplicons produced by PCR were purified and were submitted to the bidirectional direct sequencing using the BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.1 (Applied Biosystems, Foster City, CA, USA), and were analyzed with an ABI 3500 XL automatic sequencer (Applied Biosystems®) using standard methods.

Sequencing data were analyzed through the CodonCodeAligner version 4.1.1 software, and the found mutations were compared with online data banks (Ensembl, HGMD®, NCBI and dbSNP). The novel mutation was valued by the online prediction mutation software

(<http://www.mutationtaster.org>). In addition, 100 alleles from control individuals were investigated for the novel mutation.

Results

The molecular analyses from all the 28 affected individuals, their parents, and also six affected sibs are summarized in the Table 1. All the parental alleles from probands in compound heterozygous status, except one (case 26), could be identified. Recurrence occurred in nine families, but three were not available for studying. The sex distribution considering all the affected individuals was 20 males and 17 females (Table 1).

Parental consanguinity was referred in eleven families (39%) and was more frequent in the families from Ceará State. In two families (cases 7 and 17), the parents recognized they are related but they unaware their degree of kinship, therefore they were interpreted as having a distant consanguinity. In the present casuistic the relationship more distant ($F \leq 1/32$) were more frequent (7 to 4 families) (Table 1).

Rather than a single, five different mutations were found in the Ceará State, four previously described (G146R, G194S, R241X, and C318Y), and a novel one (W29Mfs*10) (Table 1). Therefore, these data suggest that the high frequency of Pycnodysostosis in the Ceará State is the outcome of the high rate of inbreeding in the region. Surprisingly, all these same mutations, but one (A277V), were also observed in the other Brazilian regions, making a total of six different mutations detected in the complete survey. The A277V mutation was observed in a single family from the South region (Table 1). A set of all described mutations are shown in the Figure 1. The mutations detected in this study are highlighted in this figure.

The novel mutation (W29Mfs*10), is an insertion of a thymine between the bases in position 82 and 83 in the exon 2 of the cDNA (Figure 2). This frameshift mutation disrupts 10 amino acids ahead until to find a stop codon in the end of the exon 2. As expected, analysis in

silico predicted this variation as a pathogenic mutation, and therefore, it was absent in 100 alleles from control individuals.

Table 9. Studied patients listed by sex, inbreeding coefficient, familiar recurrence and the found genotype by paternal and maternal alleles.

Case number	Gender	F	Rec.	Index case Genotype			
				Maternal allele	Exon	Paternal allele	Exon
Ceará State							
1	M	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
2	F	0	N	p.W29Mfs*10	2	p.R241X	6
3	M	0	N	p.R241X	6	p.R241X	6
4	F	1/32	N	p.C318Y	8	p.C318Y	8
5	F	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
6	M	0	N	p.R241X	6	p.R241X	6
7	F	?	Sister*	p.R241X	6	p.R241X	6
8	M	0	Sister	p.G146R	5	p.G146R	5
9	M	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
10	M	1/256	N	p.W29Mfs*10	2	p.W29Mfs*10	2
11	M	1/16	N	p.R241X	6	p.R241X	6
12	F	0	Brother	p.R241X	6	p.R241X	6
13	M	1/16	N	p.G146R	5	p.G146R	5
14	M	0	Brother	p.G146R	5	p.G194S	5
15	M	1/16	N	p.G146R	5	p.G146R	5
16	M	0	N	p.W29Mfs*10	2	p.W29Mfs*10	2
Other States							
17	M	?	Sister*	p.G146R	5	p.G146R	5
18	M	0	Sister	p.R241X	6	p.G146R	5
19	F	0	Brother*	p.C318Y	8	p.C318Y	8
20	F	0	N	p.C318Y	8	p.R241X	6
21	F	0	Sister	p.W29Mfs*10	2	p.R241X	6
22	F	1/64	N	p.G146R	5	p.G146R	5
23	M	1/256	N	p.C318Y	8	p.C318Y	8
24	M	0	Sister	p.R241X	6	p.G146R	5
25	F	0	N	p.C318Y	8	p.W29Mfs*10	2
26	M	0	N	#p.R241X	5	#p.G146R	6
27	M	1/16	N	p.R241X	6	p.R241X	6
28	F	1/256	N	p.A277V	7	p.A277V	7

M: Male, F: Female, Y: Yes, N: No, ?: Parents are consanguineous but unaware the family relationship, F:

Inbreeding Coefficient, Rec. Recurrence. *DNA not available for analysis, #Origin of each allele was not

proved. Patients 1 to 16 are from Ceará State (Brazilian northeast), Patients 17 to 19 are from Paraíba

State (Brazilian northeast), Patient 20 is from Maranhão State (Brazilian northeast), Patient 21 is from

Goiás State (Brazilian Central-west), Patients 22 to 27 are from São Paulo State (Brazilian Southeast) and

Patient 28 are from Rio Grande do Sul State (Brazilian South).

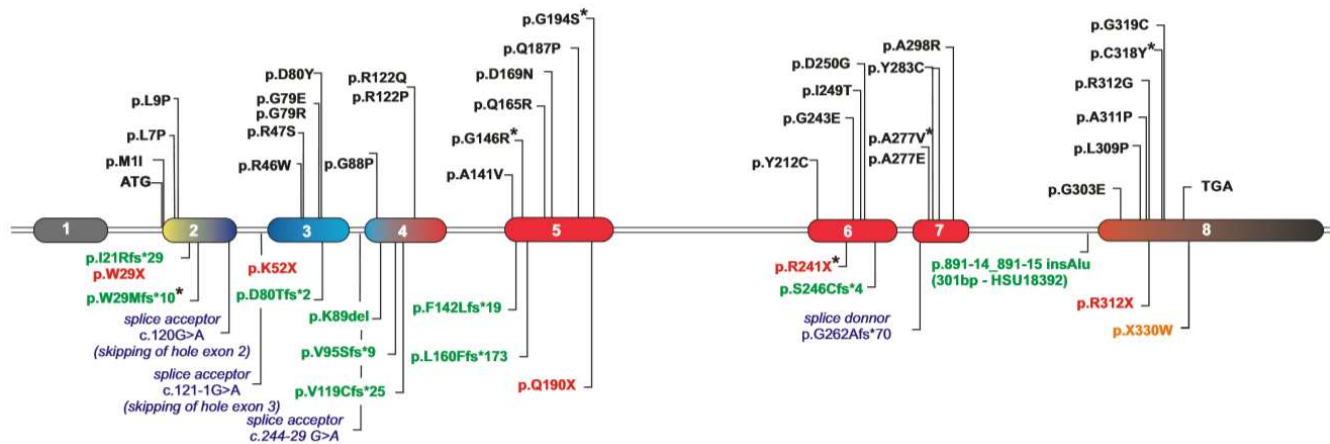


Figure 1. Structure of CTSK gene showing the distribution of the 51 reported mutations. There are 31 Missenses mutations (black), 9 frameshifts and 1 codon deletion mutations (green), 5 nonsense mutations (red), 4 splice mutations (Blue) and 1 stop codon mutation (orange). Exon 1 and the final portion of exon 8 are non-coding (black). The exons are painted according to region of protein they encode. Yellow represents the Pre region (15 amino acids), blue indicates the Pro region (99 amino acids) and red represents the mature protein region (215 amino acids). *Mutations found in this study. For more details about the described mutations see table S2 in supplementary information.

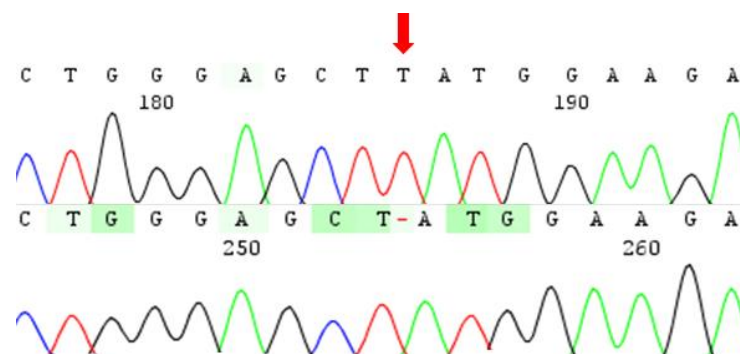


Figure 2. Electropherogram showing a novel frameshift mutation c.82_83insT (p.W29Mfs*10) found in Brazilians patients. The red arrow indicates an insertion of a Thymine in homozygous found in a patient with Pycnodysostosis and below is the sequencing of a control individual.

The most frequent mutation in the whole casuistic was the R241X, followed by the G146R and the novel mutation W29Mfs*10. These mutations, together with the mutation C318Y, were found in both, homozygous and heterozygous status (Table 2). Together, the A277V and the G194S mutations were the less frequent in the present casuistic, both found in a single family. However, while the first was found in homozygosity, the last one was observed

in compound heterozygous status (G194S + G146R). The Table 2 shows the frequency of all combinations of the mutated alleles found in the index cases.

The set of mutations here detected were found in homozygous status in 20 patients (71,4%) (Table 2), especially in the patients from Ceará State, where homozygous mutations were observed in 14 patients (87,5%) (Table 1). Among the 12 patients from the other States the mutations in homozygous status were seen in a half (six) of the patients (Table 1). Since the mutations in the *CTSK* gene is rare, the homozygosity strongly suggest a common ancestor, and therefore a possible consanguineous marriage of the parents, even if this is a distant consanguinity. Based on this consideration, the real consanguinity rate among these families is 71,4% (20 in 28), reaching to 87,5% (14 in 16) among the families from Ceará State. These data suggest that, rather a founder effect, the high consanguinity rate must be the cause of the high frequency of Pycnodysostosis in the Ceará State.

Table 2. Frequency of combination of mutated alleles in the 28 index cases studied. *In one of these patients the origin of maternal and paternal allele wasn't determined. The highlighted ones represent the homozygous patients.

Maternal allele Paternal allele	p.W29Mfs*10 exon 2	p.G146R exon 5	p.G194S exon 5	p.R241X exon 6	p.A277V exon 7	p.C318Y exon 8
p.W29Mfs*10 exon 2	5/28	-	-	-	-	1/28
p.G146R exon 5	-	5/28	1/28	3/28*	-	-
p.G194S exon 5	-	-	-	-	-	-
p.R241X exon 6	2/28	-	-	6/28	-	1/28
p.A277V exon 7	-	-	-	-	1/28	-
p.C318Y exon 8	-	-	-	-	-	3/28

The distribution of the found mutations in the patients from Ceará State shows that the majority are concentrated in some small cities (Figure 3), especially the novel mutation, that seems to have originated, or should have been introduced, in a small region in the Ceará's Northwest coast, and then, began to spread due to inbreeding. For this mutation, we believe it is possible to hypothesizing a possible founder effect.

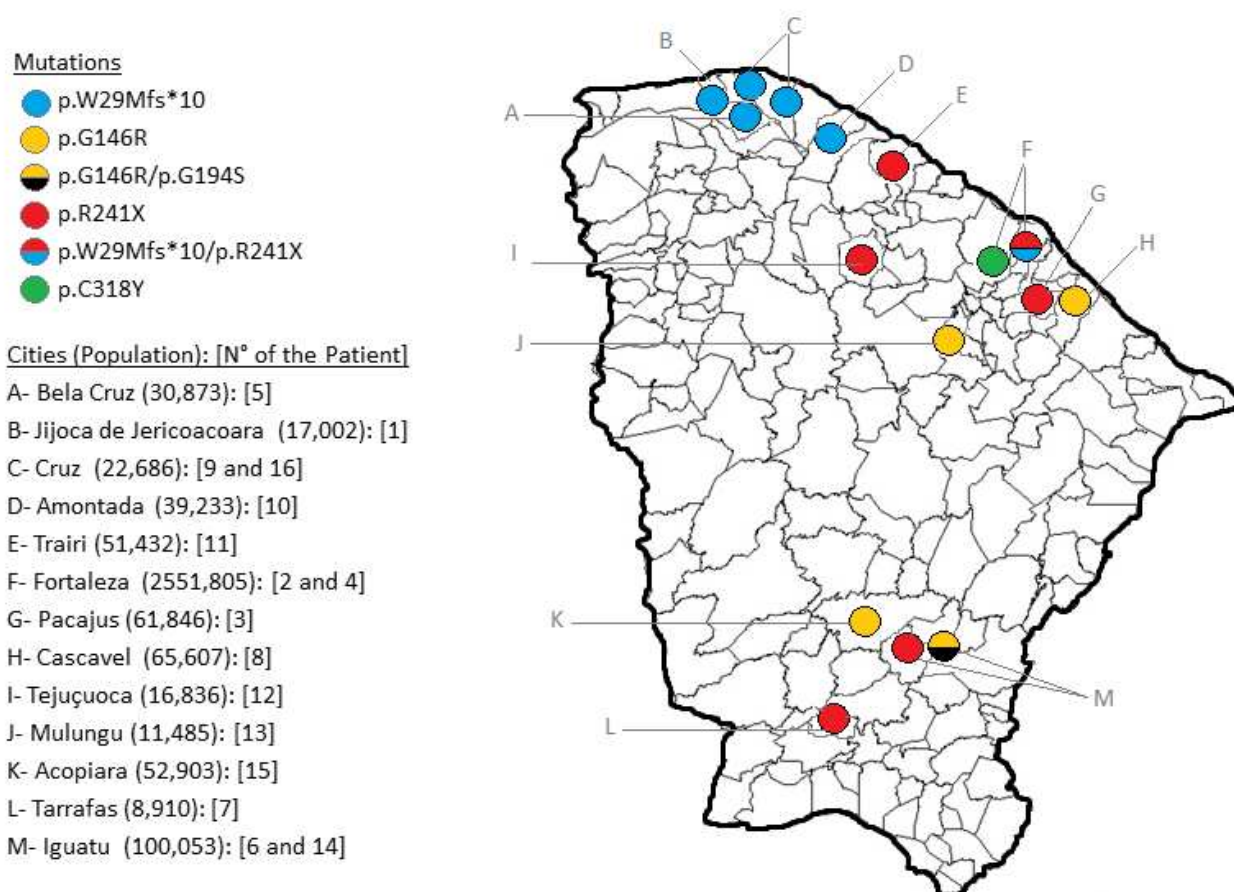


Figure 3. Distribution of mutations found in patients from Ceará (1 to 16). Each circle represents a patient carrying one specific mutation (homozygous) or two (compound heterozygous), according to the colors in the legend.

Among the patients from other States and regions the distribution is as follow: four were from other Northeast States (patients 17, 18, 19 came from Paraíba State, and patient 20 came from Maranhão State); one were from the Midwest region (patient 21 came from Goiás State); six were from the Southeast region (patients 22 - 27 came from São Paulo State); and just one were from the South region (patient 28) (Table 3, Figure 4). When we analyze the familiar background regarding the parents and grandparents birth locality, we found that all, but two families (26 and 28), have some ancestor coming from the Northeast region, near from the Ceará State (Table 3). For the two patients (26 and 28), the patient 26, is from the inside of the São Paulo State and just recognizes local and Latin European ancestor in his

Table 3. The geographic origin of the families, including the parents locality of birth and the respective grandparents, according the found mutations. # Family 's origin in the Northeast, (?)Origin of the parental alleles not proved.

Case number	Current patients' address		Maternal allele	Maternal family origin						Paternal allele	Paternal family origin					
				Mother		Maternal Grandmother		Maternal Grandfather			Father		Paternal Grandmother		Paternal Grandfather	
				City	State	City	State	City	State		City	State	City	State	City	State
Ceará State																
1	J.Jericoacoara		W29Mfs*10	Cruz	CE	Cruz	CE	Cruz	CE	W29Mfs*10	J.Jericoacoara	CE	J.Jericoacoara	CE	J. Jericoacoara	CE
2	Fortaleza		W29Mfs*10	Itarema	CE	Itarema	CE	Itarema	CE	R241X	Itarema	CE	Itarema	CE	Itarema	CE
3	Pacajus		R241X	Pacajus	CE	Fortaleza	CE	Iguatu	CE	R241X	Jaguaretama	CE	Jaguaretama	CE	Jaguaretama	CE
4	Fortaleza		C318Y	Fortaleza	CE	-	-	-	-	C318Y	-	CE	-	-	-	-
5	Bela Cruz		W29Mfs*10	Bela Cruz	CE	Bela Cruz	CE	Bela Cruz	CE	W29Mfs*10	Bela Cruz	CE	Bela Cruz	CE	Bela Cruz	CE
6	Iguatu		R241X	Iguatu	CE	-	PB	-	PB	R241X	Iguatu	CE	Iguatu	CE	Iguatu	CE
7	Tarrafas		R241X	-	CE	-	-	-	-	R241X	-	CE	-	-	-	-
8	Cascavel		G146R	Cascavel	CE	Cascavel	CE	Cascavel	CE	G146R	Mossoró	RN	Cascavel	CE	Mossoró	RN
9	Cruz		W29Mfs*10	Cruz	CE	Cruz	CE	Cruz	CE	W29Mfs*10	Cruz	CE	-	-	-	-
10	Amontada		W29Mfs*10	-	CE	-	-	-	-	W29Mfs*10	-	CE	-	-	-	-
11	Trairi		R241X	-	CE	-	-	-	-	R241X	-	CE	-	-	-	-
12	Tejuçuoca		R241X	-	CE	-	-	-	-	R241X	-	CE	-	-	-	-
13	Mulungu		G146R	-	CE	-	-	-	-	G146R	-	CE	-	-	-	-
14	Iguatu		G146R	-	CE	-	-	-	-	G194S	-	CE	-	-	-	-
15	Acopiara		G146R	-	CE	-	-	-	-	G146R	-	CE	-	-	-	-
16	Cruz		W29Mfs*10	Cruz	CE	-	-	-	-	W29Mfs*10	Bela Cruz	CE	-	-	-	-
Other States																
#17	C. Grande	PB	G146R	Jericó	PB	-	-	-	-	G146R	Jericó	PB	-	-	-	-
#18	C. Grande	PB	R241X	Mombaça	CE	-	-	-	-	G146R	Sousa	PB	-	-	-	-
#19	C. Grande	PB	C318Y	Arara	PB	-	-	-	-	C318Y	Casserengue	PB	-	-	-	-
#20	São Luís	MA	C318Y	São Luís	MA	Grajaú	MA	Grajaú	MA	R241X	São Luís	MA	Gov.Eug.Barros	MA	I. Carrapatal	MA
#21	Formosa	GO	W29Mfs*10	Currais Novos	RN	Currais Novos	RN	-	RN	R241X	Uberlândia	MG	Uberlândia	MG	Uberlândia	MG
#22	Jaguariúna	SP	G146R	Acopiara	CE	-	-	-	-	G146R	Acopiara	CE	-	-	-	-
#23	Campinas	SP	C318Y	Arapiraca	AL	-	AL	-	AL	C318Y	Feira Grande	AL	Feira Grande	AL	Feira Grande	AL
#24	Campinas	SP	R241X	R. Pombal	BA	-	BA	-	BA	G146R	R. Pombal	BA	-	BA	-	BA
#25	Guarulhos	SP	C318Y	V. Conquista	BA	V. Conquista	BA	V. Conquista	BA	W29Mfs*10	S. Parnaíba	SP	J. do Norte	CE	-	PR
26	São Carlos	SP	R241X(?)	São Carlos	SP	São Carlos	SP	São Carlos	SP	G146R(?)	São Carlos	SP	São Carlos	SP	São Carlos	SP
#27	São Paulo	SP	R241X	Lagoa	PB	-	-	-	-	R241X	Lagoa	PB	-	-	-	-
28	P. Alegre	RS	A277V	Vacaria	RS	Vacaria	RS	Vacaria	RS	A277V	Vacaria	RS	Vacaria	RS	Vacaria	RS

Discussion

Pycnodysostosis is a rare autosomal recessive condition classified among the skeletal dysplasias with increase of the bone density. The prevalence is estimated between 1 to 1.7 per million, although none epidemiological study give support to this estimative [2, 9-11]. However, even in the absence of a reliable epidemiological study regarding Pycnodysostosis, the clinical experience supports it as a rare skeletal dysplasia.

In the present survey we have studied, by molecular sequencing, a total of 34 (28 probands) patients with Pycnodysostosis, among a total of 37 affected - 20 males and 17 females. Three related individuals were not available for study (Table 1).

Due to the observed high frequency of Pycnodysostosis in the Ceará State, this survey firstly focused to the molecular study of the 16 patients coming from that region. We have found five different mutations in these patients, four previously described (G146R, G194S, R241X, C318Y) and a novel one (W29Mfs*10) (Table 1, Figure 3). Although 37.5% of the families had referred consanguinity, most of mutations (87.5%) were found in homozygosity (Table 1), which strongly suggests that this is the real consanguinity rate among these families, probably reflecting a distant consanguinity. Therefore, the initial hypothesis of a founder effect was moved away, and we propose the high consanguinity rate as the main cause of the high frequency of Pycnodysostosis in that region. Although the consanguinity rate in the whole Brazilian Northeast is well recognized [31], some small and particular areas in the inside of the Brazilian Northeast reveal consanguinity rate as high as 32% [32].

In clinical genetics a consanguineous marriage is defined as a union between individuals who are related as second cousins or closer, which means a “F” equal or higher than 1/64 [33]. Consanguineous marriages with the common ancestor(s) placed more distant, at least in theory, would result to an insignificant probability of homozygosity by descent. However, deleterious autosomal recessive alleles may remain hidden in heterozygous state within some family for many

generations, and then consanguineous marriages between mutation carriers, even in distant consanguinity, could bring them to come the surface [34]. Since the probability that a carrier finds a partner in the general population who bears the same mutation is very small, this effect is more astonishing for rare diseases [34]. Thus, if we examine a group of families with affected by proven autosomal recessive diseases we must find a higher consanguinity rate than the original population from where these families are coming from [34]. In a similar way, it is known that the consanguinity rate in the Brazilian Northeast is high [31], the consanguinity referred by the families with Pycnodysostosis coming from Ceará State in this study is of 37.5%, however, when we study the molecular basis of these patients we, finally arrive on the real consanguinity rate, that is, in this case, 87.5%.

Since the novel mutation was found to be concentrated in the Northwest coast of the Ceará State (Figure 3) we hypothesize this mutation could have been originated or introduced in some of those cities and then spread due to consanguineous marriages as a founder effect. Unfortunately, we were unable to confirm this hypothesis so far.

The most frequent mutations R241X and G146R were found more disperse in the Ceará State (Figure 3), probably suggesting that their dispersion began earlier. The R241X mutation have already been reported in patients from Portugal [8, 13], Spain [8, 9, 13], Italy [13], Canada [9], and Mexico [35]. The G146R mutation, however, was identified in patients from Tunisia [25] and Morocco [6]. Moreover, both mutations were already described in others Brazilian patients previously reported, but unfortunately the origin of the patients were not informed [16].

The mutation C318Y, observed in just one family from Ceará State was firstly reported in three Brazilian patients- two in homozygosity and one in case as a compound heterozygous [16]. Recently it was detected in an African patient [36]. Another mutation described once in a patient living in region in the South of the Ceará State, the G194S, was detected in an Italian patient so far [13]. We believe that, unlike of the novel mutation W29Mfs*10, never identified in other patients

out from Brazil, the other mutations (G146R, R241X and C318Y), should be introduced in the Northeast Brazil a long time ago, and then spread to other regions due to consanguineous marriages.

For the patients living in other regions outside from Ceará State, five different mutations were identified, four were the same found in the Ceará's patients (W29Mfs*10, G146R, R241X and C318Y) and the other one was the missense mutation A277V, observed in homozygosity in a single patient and of course in heterozygosity in her consanguineous parents (Table 1, Figure 4). This family has Italian ancestors and is from the Brazilian South region, but this mutation was described in a patient with Belgium and Algerian background [7] and then was found in Japanese [9, 10, 19] and Pakistani patients [13, 30]. The distant consanguinity ($F = 1/256$) in this family support the idea that this mutation was segregated hidden and, a consanguineous marriage, put it in evidence as it is expected to happen for the rare genes [34]. Related to this mutation It is also worthy that it is described as a *hot spot* because of its localization in a CpG dinucleotides regions [2].

We believe that, except for the mutation A277V above discussed, the other mutations found in the patients currently living outside from Ceará State came from Northeast, since most of these patients have ancestors in the Northeast region (Table 3). This is particularly interesting when we look at to the novel mutation W29Mfs*10, which was found in compound heterozygous state in two patients living outside from Ceará State (patients 21 and 25). In the patient 21 the novel mutation was present in the maternal allele, whose ancestor is from Currais Novos, a city located in the Rio Grande do Norte, a neighbor State of Ceará. In the patient 25 the novel mutation was in the paternal allele, whose ancestor is from Ceará State (Table 3).

These results suggest that the majority of the mutations found in the families outside from Ceará State are been spread for the other regions by internal migration of the population. The internal migration in Brazil is usually motivated by economic reasons, in which young people leave the rural areas or small towns in search of job opportunities in urban centers or larger cities [37]. Considering that the Northeast region has the lowest socioeconomic level among the five Brazilian regions, the internal migration usually follows the sense from Northeast (with a negative migration

number) to Southeast (with positive migration), being the São Paulo State the destiny of most of migrants [38].

Among the families that are still living in the Northeast, the examination of the Figure 4 suggest that the most migrated from small towns to the bigger cities (Table 3).

This intern migration from Northeast to Southeast also favor to the increase of consanguinity rate in the latter region [31, 39]. In this study we have observed that the real consanguinity rate in the studied families outside from Ceará State was 50%.

Among the mutations found in our complete casuistic, the most frequent were the R241X and the G136R (Table 2). These mutations, together with the A277V are described as the most frequent mutations among the patients affected by Pycnodysostosis worldwide [2]. The third mutation more frequent found in this study was the novel variant W29Mfs*10, for which we hypothesize it might be arisen in the Brazilian Northeast, more precisely in the Northwest coast from Ceará State.

Conclusions

The results of the present study showed that the high consanguinity rate (87.5%) must be the main cause of high frequency of Pycnodysostosis in the Ceará State. We also found a novel mutation (W29Mfs*10), that seems to have been originated in the Northwest coast from Ceará State, and its spread should be producing a founder effect. Finally, by investigation of the origin of the parental families, we have demonstrated that the mutations found in the patients with Pycnodysostosis outside from Ceará State are derived from that found in the Ceará State by the process of internal migration. In these patients the consanguinity is also high (50%).

Supporting data

Additional file 1: Table S1. Nucleotide sequence of each primer, its length and PCR product length.

The primers were named according to exon flanked.

Additional file 2: Table S2. This table summarizes all described mutations in CTSK gene until August 2015, its location in DNA and the effect on amino acid. It is also shown the origin of the first patient described as carrying the mutation and the corresponding author who reported the case.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

Araujo TF participated in the design of the study, carried out the molecular genetic studies, and analyzed the data and writing the manuscript. Ribeiro EM and Arruda AP recruited the patients from Ceará. Medeiros PFV, Minillo RM, Melo DG, Kim CA, Doriqui MJR, Felix TM and Fock RA recruited patients from other regions of Brazil. Moreno CA participated in discussion of the results. Cavalcanti DP coordinated the study, participated in the revision of all patients, the molecular results, and in the discussion of the results and revision of the manuscript. All authors read and approved the final manuscript.

Author's information

¹ Skeletal Dysplasia Group, Department of Medical Genetics, Faculty of Medical Sciences, State University of Campinas, Campinas, SP, Brazil

² Children's Hospital Albert Sabin, Fortaleza, CE, Brazil,

³ Medical Sciences Faculty of Juazeiro do Norte (FMJ), CE, Brazil,

⁴ Perinatal Genetics Program, Department of Medical Genetics, Faculty of Medical Sciences, State University of Campinas, Campinas, SP, Brazil,

⁵ Federal University of Campina Grande, PB, Brazil

⁶ Children's Clinic, City Hall of Guarulhos, SP, Brazil

⁷ Medical Department, Federal University of de São Carlos (UFSCAR), SP, Brazil

⁸ Medical Genetics Unit, Children's Institute, Medical Sciences Faculty, University of São Paulo (FCMUSP), SP, Brazil

⁹ Children's Hospital Juvêncio Mattos, São Luís, MA, Brazil

¹⁰ Medical Genetics Service, Clinical Hospital of Porto Alegre, RS, Brazil

¹¹ Centro de Genética Médica da Universidade Federal de São Paulo, SP, Brazil

Acknowledgments

We thank all the patients and their families. We are also grateful to Thatiane Y Kanazawa, Cynthia Silveira, Karina C Silveira and Luciana Bonadia for the laboratorial support, Maria Dora Lacarrubba Flores for the clinical support, and Maynara Cavalcante for the help in the recovering of the patients' data from Ceará.

Funding

This study was supported by the following grant sponsors: CNPq 402008/2010-3 and 590148/2011-7. CAPES 3300017023p6

References

1. Bonafe L, Cormier-Daire V HC, Lachman R, Mortier G MS, Nishimura G, Sangiorgi L SR, Sillence D, Spranger J S-FA, Warman M US: **Nosology and Classification of Genetic Skeletal Disorders: 2015 Revision**. *Am J Med Genet Part A* 2015, **9999A**:1–24.
2. Xue Y, Cai T, Shi S, Wang W, Zhang Y, Mao T, Duan X: **Clinical and animal research findings in pycnodysostosis and gene mutations of cathepsin K from 1996 to 2011**. *Orphanet J Rare Dis* 2011, **6**:20.
3. Motyckova G, Fisher DE: **Pycnodysostosis: role and regulation of cathepsin K in osteoclast function and human disease**. *Curr Mol Med* 2002, **2**:407–21.
4. Alves N, Cantín M: **Clinical and radiographic maxillofacial features of pycnodysostosis**. *Int J Clin Exp Med* 2014, **7**:492–496.
5. Testani E, Scarano E, Leoni C, Dittoni S, Losurdo A, Colicchio S, Gnoni V, Vollono C, Zampino G, Paludetti G, Della Marca G: **Upper airway surgery of obstructive sleep apnea in pycnodysostosis: Case report and literature review**. *Am J Med Genet A* 2014, **164**:2029–35.
6. Gelb B, Shi G, Chapman H, Desnick R: **Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency**. *Science* 1996, **273**:1236–8.
7. Gelb B, Willner J, Dunn T: **Paternal uniparental disomy for chromosome 1 revealed by molecular analysis of a patient with pycnodysostosis**. *Am J Hum Genet* 1998, **62**:848–854.
8. Hou WS, Brömme D, Zhao Y, Mehler E, Dushey C, Weinstein H, Miranda CS, Fraga C, Greig F, Carey J, Rimoin DL, Desnick RJ, Gelb BD: **Characterization of novel cathepsin K mutations in the pro and mature polypeptide regions causing pycnodysostosis**. *J Clin Invest* 1999, **103**:731–738.
9. Nishi Y, Atley L, Eyre DE, Edelson JG, Superti-Furga a, Yasuda T, Desnick RJ, Gelb BD: **Determination of bone markers in pycnodysostosis: effects of cathepsin K deficiency on bone matrix degradation**. *J Bone Miner Res* 1999, **14**:1902–1908.

10. Fujita Y, Nakata K, Yasui N, Matsui Y, Kataoka E, Hiroshima K, Shiba RI, Ochi T: **Novel mutations of the cathepsin K gene in patients with pycnodysostosis and their characterization.** *J Clin Endocrinol Metab* 2000, **85**:425–31.
11. Haagerup a, Hertz JM, Christensen MF, Binderup H, Kruse T a: **Cathepsin K gene mutations and 1q21 haplotypes in at patients with pycnodysostosis in an outbred population.** *Eur J Hum Genet* 2000, **8**:431–436.
12. Fratzl-Zelman N, Valenta A, Roschger P, Nader A, Gelb BD, Fratzl P, Klaushofer K: **Decreased bone turnover and deterioration of bone structure in two cases of pycnodysostosis.** *J Clin Endocrinol Metab* 2004, **89**:1538–47.
13. Donnarumma M, Regis S, Tappino B, Rosano C, Assereto S, Corsolini F, Di Rocco M, Filocamo M: **Molecular analysis and characterization of nine novel CTSK mutations in twelve patients affected by pycnodysostosis. Mutation in brief #961. Online.** *Hum Mutat* 2007, **28**:524.
14. Schilling a F, Mülhausen C, Lehmann W, Santer R, Schinke T, Rueger JM, Amling M: **High bone mineral density in pycnodysostotic patients with a novel mutation in the propeptide of cathepsin K.** *Osteoporos Int* 2007, **18**:659–69.
15. Li H-Y, Ma H-W, Wang H-Q, Ma W-H: **Molecular Analysis of a Novel Cathepsin K Gene Mutation in a Chinese Child with Pycnodysostosis.** *J Int Med Res* 2009, **37**:264–269.
16. Bertola D, Amaral C, Kim C, Albano L, Aguená M, Passos-Bueno MR: **Craniosynostosis in pycnodysostosis: broadening the spectrum of the cranial flat bone abnormalities.** *Am J Med Genet A* 2010, **152A**:2599–603.
17. Khan B, Ahmed Z, Ahmad W: **A novel missense mutation in cathepsin K (CTSK) gene in a consanguineous Pakistani family with pycnodysostosis.** *J Investig Med* 2010, **58**:720–4.
18. Toral-López J, Gonzalez-Huerta LM, Sosa B, Orozco S, González HP, Cuevas-Covarrubias SA: **Familial pycnodysostosis: identification of a novel mutation in the CTSK gene (cathepsin K).** *J Investig Med* 2011, **59**:277–80.
19. Matsushita M, Kitoh H, Kaneko H, Mishima K, Itoh Y, Hattori T, Ishiguro N: **Novel Compound Heterozygous Mutations in the Cathepsin K Gene in Japanese Female Siblings with Pyknodysostosis.** *Mol Syndromol* 2012, **8550**:254–258.
20. Utokpat P, Panmontha W, Tongkobpetch S, Suphapeetiporn K, Shotelersuk V: **Novel CTSK mutation resulting in an entire exon 2 skipping in a Thai girl with pycnodysostosis.** *Pediatr Int* 2013, **55**:651–5.
21. Zheng H, Zhang Z, He JW, Fu WZ, Zhang ZL: **A novel mutation (R122Q) in the cathepsin K gene in a Chinese child with Pyknodysostosis.** *Gene* 2013, **521**:176–179.
22. Pangrazio A, Puddu A, Oppo M: **Exome sequencing identifies CTSK mutations in patients originally diagnosed as intermediate osteopetrosis.** *Bone* 2014, **59**:122–126.
23. Arman A, Bereket A, Coker A, Kiper PÖS, Güran T, Ozkan B, Atay Z, Akçay T, Haliloglu B, Boduroglu K, Alanay Y, Turan S: **Cathepsin K analysis in a pycnodysostosis cohort: demographic, genotypic and**

phenotypic features. *Orphanet J Rare Dis* 2014, **9**:60.

24. Ozdemir T, Atik T, Karaca E, Onay H, Ozkinay F, Cogulu O: **A novel mutation in two families with pycnodysostosis.** *Clin Dirmophol* 2013, **22**:102–5.
25. Tinsa F, Hamouda S, Bellalah M, Bousnina D, Karbout L, Boussetta K, S B: **Unusual feature of pycnodysostosis: pectus carinatum.** *La Tunisi Médicale* 2014, **92**:180–181.
26. Xue Y, Wang L, Xia D, Li Q, Gao S, Dong M, Cai T, Shi S, He L, Hu K, Mao T, Duan X: **Dental Abnormalities Caused by Novel Compound Heterozygous CTSK Mutations.** *J Dent Res* 2015, **94**:674–681.
27. Novinec M, Lenarčič B: **Cathepsin K: a unique collagenolytic cysteine peptidase.** *Biol Chem* 2013, **394**:1163–79.
28. Mujawar Q, Naganoor R, Patil H, Thobbi AN, Ukkali S, Malagi N: **Pycnodysostosis with unusual findings: a case report.** *Cases J* 2009, **2**:6544.
29. Karakurt L, Yilmaz E, Belhan O, Serin E: **Pycnodysostosis associated with bilateral congenital pseudarthrosis of the clavicle.** *Arch Orthop Trauma Surg* 2003, **123**:125–7.
30. Naeem M, Sheikh S, Ahmad W: **A mutation in CTSK gene in an autosomal recessive pycnodysostosis family of Pakistani origin.** *BMC Med Genet* 2009, **10**:76.
31. Freire-Maia N: **Genetic effects in Brazilian populations due to consanguineous marriages.** *Am J Med Genet* 1990, **35**:115–117.
32. Santos S, Kok F, Weller M: **Inbreeding levels in Northeast Brazil: strategies for the prospecting of new genetic disorders.** *Genet Mol Biol* 2010, **223**:220–223.
33. Bittles A: **Consanguinity and its relevance to clinical genetics.** *Clin Genet* 2001, **60**:89–98.
34. Jaouad IC, Elalaoui SC, Sbiti A, Elkerh F, Belmahi L, Sefiani A: **Consanguineous marriages in Morocco and the consequence for the incidence of autosomal recessive disorders.** *J Biosoc Sci* 2009, **41**:575–581.
35. Johnson M, Polymeropoulos M, Vos H, Ortiz de Luna R, Francomano C: **A nonsense mutation in the cathepsin K gene observed in a family with pycnodysostosis.** *Genome Res* 2015:1051–1055.
36. Girbal I, Nunes T, Medeira a., Bandeira T: **Pycnodysostosis with novel gene mutation and severe obstructive sleep apnoea: management of a complex case.** *Case Reports* 2013, **2013**:bcr2013200590–bcr2013200590.
37. Sahota GS: **An Economic Analysis of Internternal migration in Brazil is usually motivated by economic issues in which young people leave rural areas and small cities in search of job opportunities in major citienal Migration in Brazil.** *J Polit Econ* 1968, **76**:218 – 245.
38. Marques D, Golgher A: **Migration diversity in Brazil : where are the poor people ? .**
39. Araujo CM CD: **Consanguinidade parental em uma amostra de recém-nascidos do CAISM.** In *21° Congresso Interno de Iniciação Científica da Unicamp*. Campinas; 2003.

Additional files

Table S1. Nucleotide sequence of each primer, its length and PCR product length. The primers were named according to exon flanked. F: *primer forward* and R: *primer reverse*. The primers were designed using the transcript ID ENST00000271651 (Ensembl) and the following softwares: Gene Runner, Primer 3 and Primer-BLAST.

Primer	Nucleotide sequence (5'-3')	Primer length	Amplicon length
2F	GTCCTTGGAACCAGATGTAC	20bp	417 bp
2R	CATGAGTTAGGGAAGAGGG	19bp	-
34F	CTCTTCTTGCAGGCTCTTAATTC	23 bp	636 bp
34R	GATGTACCTTAATTCCTTGCCC	22 bp	-
5F	CAACAGATATTAAGCGGAAGAG	22 bp	518 bp
5R	CTTCTTGGCCTTCTGTGG	18 bp	-
67F	CTGCTGCCTCTGTTAGTTCAC	21 bp	744 bp
67R	GGGACAGAGAAAGGAATATCG	21 bp	-
8F	CTTCCCTCCTTTGATACCTAC	22 bp	448 bp
8R	GGAAGGATCATTTGAAGCAC	20 bp	-

Table S2. This table summarizes all described mutations in CTSK gene until August 2015, its location in DNA and the effect on aminoacid. It is also shown the origin of the first patient described as carrying the mutation and the corresponding author who reported the case.

Location in DNA	Patient origin	Reference article	Coding DNA variant sequence [RefSeq:NM_000396.3]	Effect on aminoacid [RefSeq:NP_000387.1]
Missense Mutations (31)				
Exon 2	Turkey	Arman, <i>et al.</i> , 2014	c.3G>A	p. M1I
Exon 2	Italy	Donnarumma, <i>et al.</i> , 2007	c.20T>C	p.L7P
Exon 2	Japan	Nishi, <i>et al.</i> , 1999	c.26T>C	p.L9P
Exon 3	Germany	Schilling, <i>et al.</i> , 2007	c.136C>T	p.R46W
Exon 3	Caribbean	Pangrazio, <i>et al.</i> , 2014	c.235G>A	p.R47S
Exon 3	Austria	Fratzl-Zelman, <i>et al.</i> , 2004	c.236G>A	p.G79R
Exon 3	Northern Europe	Hou, <i>et al.</i> , 1999	c.139C> A	p.G79E
Exon 3	Turkey	Arman, <i>et al.</i> , 2014	c.238G>T	p.D80Y
Exon 4	Bangladesh	Pangrazio <i>et al.</i> , 2014	c.263A>C	p.G88P
Exon 4	Japan	Matsushita, <i>et al.</i> , 2011	c.365G>C	p.R122P
Exon 4	China	Zheng, <i>et al.</i> , 2013	c.365G>A	p.R122Q
Exon 5	Unknown	Chavassieux, <i>et al.</i> , 2008	c.422C>T	p.A141V
Exon 5	Morocco, Hispanic Am.	Gelb <i>et al.</i> , 1996	c.436G>C	p.G146R
Exon 5	Italy	Donnarumma, <i>et al.</i> , 2007	c.494A>G	p.Q165R
Exon 5	Turkey	Arman, <i>et al.</i> , 2014	c.505G>A	p.D169N
Exon 5	China	Li, <i>et al.</i> , 2009	c.560A>C	p.Q187P
Exon 5	Italy	Donnarumma, <i>et al.</i> , 2007	c.580G>A	p.G194S
Exon 6	Spain	Hou, <i>et al.</i> , 1999	c.635A>G	p.Y212C
Exon 6	Pakistan	Khan, <i>et al.</i> , 2010	c.728G>A	p.G243E
Exon 6	Spain	Donnarumma, <i>et al.</i> , 2007	c.746T>C	p.I249T
Exon 6	Pakistan	Donnarumma, <i>et al.</i> , 2007	c.749A>G	p.D250G
Exon 7	Belgium	Gelb, <i>et al.</i> , 1998	c.830C>T	p.A277V
Exon 7	India, Portugal	Hou, <i>et al.</i> , 1999	c.830C>A	p.A277E
Exon 7	Unknown	Xue <i>et al.</i> , 2015	c.848A>G	p.Y283C
Exon 7	Japan	Nishi, <i>et al.</i> , 1999	c.892T>C	p.W298R
Exon 8	Mexico	Toral-Lopez <i>et al.</i> , 2010	c.908G>A	p.G303E
Exon 8	Denmark	Haagerup, <i>et al.</i> , 2000	c.926T>C	p.L309P
Exon 8	Swiss	Nishi, <i>et al.</i> , 1999	c.931G>C	p.A311P
Exon 8	Honduras	Hou, <i>et al.</i> , 1999	c.934C>G	p.R312G
Exon 8	Brazil	Bertola <i>et al.</i> , 2010	c.953G>A	p.C318Y
Exon 8	Pakistan	Donnarumma, <i>et al.</i> , 2007	c.955G>T	p.G319C
Nonsense Mutations (5)				
Exon 2	Unknown	Xue <i>et al.</i> , 2015	c.87G>A	p.W29X
Exon 3	Northern Europe	Hou, <i>et al.</i> , 1999	c.154A>T	p.K52X
Exon 5	Northern Europe	Hou, <i>et al.</i> , 1999	c.568C>T	p.Q190X
Exon 6	Hispanic America	Gelb, <i>et al.</i> , 1996	c.721C>T	p.R241X

Exon 8	Turkey	Arman, <i>et al.</i> ,2014	c.934C>T	p.R312X
Frameshift Mutations (9)				
Exon 2	Morocco	Donnarumma, <i>et al.</i> ,2007	c.60_61dupGA	p.I21Rfs*29
Exon 2	Brazil	Araujo <i>et al.</i> , 2015*	c.82_83insT	p.W29Mfs*10
Exon 3	Austria	Fratzl-Zelman, <i>et al.</i> ,2004	c.238delG	p.D80Tfs*2
Exon 4	Pakistan	Donnarumma, <i>et al.</i> ,2007	c.282dupA	p.V95Sfs*9
Exon 4	Turkey	Özdemir, <i>et al.</i> , 2013	c.354_355insT	p.V119Cfs*25
Exon 5	Japan	Fujita, <i>et al.</i> , 2000	c.426delT	p.F142Leufs*19
Exon 5	India	Singh <i>et al.</i> , 2014	c.480_481insT	p.L160Ffs*173
Exon 6	Caribbean	Pangrazio, <i>et al.</i> , 2014	c.737_738delCT	p. S246Cfs*4
<u>Intron 7</u>	Turkey	Arman <i>et al.</i> ,2014	c. 891-14_891-15 ins HSU18392**	-----
Codon deletion (1)				
Exon 4	Pakistan	Pangrazio, <i>et al.</i> , 2014	c.265_267delAAG	p. K89del
Splicing Mutations (4)				
Exon 2	Thailand	Utokpat <i>et al.</i> , 2013	c.120 G>A	p.M1_40Kdel
<u>Intron2</u>	Denmark	Haagerup, <i>et al.</i> ,2000	c.121-1G>A	p. V41_M81del
<u>Intron 3</u>	Saudi Arabia	Alhashem <i>et al.</i> , 2015	c.244 -29 G>A***	-----
Exon 7	Egypt	Donnarumma, <i>et al.</i> ,2007	c.890G>A	p.G262Afs*70
Stop Codon Mutation (1)				
Exon 8	Israel	Gelb, <i>et al.</i> , 1996	c.990A>G	p.X330W

Hispanic Am. (Hispanic America): America colonized by Spaniards, *Present study, ** It was reported that the insertion in intron 7 is a Alu sequence of 301bp inserted in reverse orientation that introduces a new potential splice acceptor site, ***There is no functional study of this mutation so far, but was described that the mutation seems to affect the splice acceptor site in intron 3.

10. ANEXOS

Anexo 1: Pareceres dos Comitês de Ética

FACULDADE DE CIÊNCIAS
MÉDICAS - UNICAMP
(CAMPUS CAMPINAS)



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: As bases moleculares da Picnodisostose em uma região brasileira com alta frequência dessa condição

Pesquisador: Thaís Fenz Araujo

Área Temática:

Versão: 3

CAAE: 25129214.0.0000.5404

Instituição Proponente: Faculdade de Ciências Médicas - UNICAMP

Patrocinador Principal: MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

DADOS DO PARECER

Número do Parecer: 671.589

Data da Relatoria: 02/06/2014

FACULDADE DE CIENCIAS
MEDICAS - UNICAMP
(CAMPUS CAMPINAS)



Continuação do Parecer: 671.589

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

- 1- Cabe ao pesquisador desenvolver o projeto conforme apresentado nesta plataforma, elaborar e apresentar os relatórios parciais e final, bem como encaminhar os resultados para publicação com os devidos créditos aos pesquisadores associados e ao pessoal técnico participante do projeto (Resolução 466/2012 CNS/MS).
- 2- O TCLE deverá ser elaborado em duas vias, sendo uma retida pelo sujeito da pesquisa ou por seu representante legal e uma arquivada pelo pesquisador (resolução 466/2012).
- 3-Se o TCLE tiver mais de uma página, o sujeito de pesquisa ou seu representante, quando for o caso, e o pesquisador responsável deverão rubricar todas as folhas desse documento, apondo suas assinaturas na última página do referido termo (Carta Circular nº 003/2011/CONEP/CNS).

CAMPINAS, 03 de Junho de 2014

Assinado por:
Fátima Aparecida Bottcher Luiz
(Coordenador)

HOSPITAL INFANTIL ALBERT
SABIN - CE



PARECER CONSUBSTANCIADO DO CEP

Elaborado pela Instituição Coparticipante

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: As bases moleculares da Picnodisostose em uma região brasileira com alta frequência dessa condição

Pesquisador: Thaís Fenz Araujo

Área Temática:

Versão: 1

CAAE: 25129214.0.3001.5042

Instituição Proponente: Faculdade de Ciências Médicas - UNICAMP

Patrocinador Principal: MINISTERIO DA CIENCIA, TECNOLOGIA E INOVACAO

DADOS DO PARECER

Número do Parecer: 739.579

Data da Relatoria: 11/06/2014

HOSPITAL INFANTIL ALBERT SABIN - CE



Continuação do Parecer: 739.579

Considerações sobre os Termos de apresentação obrigatória:

No TCLE apresentado consta como título projeto "Projeto de investigação clínico-molecular sobre a displasias esqueléticas" diferente do título apresentado neste projeto que é "As bases moleculares d Picnodisostose em uma região brasileira com alta frequência dessa condição"

Não foi apresentado termo de assentimento.

Não foi apresentado orçamento detalhado.

Recomendações:

Não se aplica

Conclusões ou Pendências e Lista de Inadequações:

As inadequações foram corrigidas. Projeto apresenta condições éticas para ser realizado.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

Aprovado em reunião do colegiado.

FORTALEZA, 05 de Agosto de 2014

Assinado por:
Regina Lúcia Ribeiro Moreno
(Coordenador)



Anexo 2: Modelo do Termo de Consentimento Livre e Esclarecido utilizado

Faculdade de Ciências Médicas – Departamento de Genética Médica

Folha de Informação do Termo de Consentimento Livre e Esclarecido (TCLE)

Projeto de investigação clínico-molecular sobre as displasias esqueléticas
--

Pesquisador principal:

Denise Pontes Cavalcanti

Departamento de Genética Médica FCM – UNICAMP, Campinas, SP

Tel/Fax: (19) 3521-0385

Email: denisepc@unicamp.br

denisepcavalcanti@gmail.com

Pesquisador responsável pelo Projeto “As bases moleculares da Picnodisostose numa região brasileira com alta frequência dessa condição”

Thaís Fenz Araujo

Departamento de Genética Médica FCM-UNICAMP, Campinas, SP

Tel: 019 3521-8902

Email: thaisfa@fcm.unicamp.br

thais.fenz@gmail.com

Colaborador do Projeto “As bases moleculares da Picnodisostose numa região brasileira com alta frequência dessa condição”

Erlane Marques Ribeiro

Hospital Infantil Albert Sabin, Fortaleza, CE

Tel: 085 9998-6577

Email: erlaneribeiro@yahoo.com.br

INTRODUÇÃO / OBJETIVO DA PESQUISA:

A pesquisa médica, em geral, ajuda no entendimento sobre as causas das doenças bem como na tentativa de encontrar terapias para melhorar a vida das pessoas afetadas. Nós estamos lhe convidando para participar desse estudo porque você ou algum membro de sua família tem problemas com seus ossos e/ou suas articulações. Acreditamos que a doença óssea na sua família pode ser causada por alteração de algum gene (pequeno erro na sua informação genética) relacionado com a formação/desenvolvimento do esqueleto.

O objetivo desta investigação é o de identificar a origem da displasia esquelética em questão e aprofundar no estudo dos mecanismos envolvidos com a mesma. As radiografias, fotos e amostras de sangue do paciente, bem como as informações médicas a respeito da família, poderão ser compartilhadas com outros pesquisadores que trabalham com displasias esqueléticas.

PROCEDIMENTO:

Concordando em participar deste estudo, os pesquisadores terão necessidade de obter alguns dados da história clínica e familiar do paciente, assim como haverá necessidade de dados do exame clínico, fotos clínicas e radiografias para o estabelecimento ou confirmação do diagnóstico clínico. O sangue para extração de DNA poderá ser coletado em qualquer momento e, normalmente a coleta é feita a partir de punção de uma veia periférica. Em várias situações DNA dos pais, e eventualmente de outros familiares, também será necessário para conclusão dos estudos. Normalmente se coleta de 4 a 10 ml de sangue (quantidade equivalente a uma ou duas colheres de sopa).

RISCOS:

A participação na presente pesquisa não acarreta riscos adicionais aos que eventualmente um paciente está exposto quando consulta um médico (ex: ser submetido a exames de raios X, coleta de sangue para exames de sangue). Com a coleta de sangue o paciente pode sentir um breve desconforto e o procedimento pode deixar uma pequena marca no local da punção. Esse procedimento dura uns poucos segundos.

VANTAGENS / RESULTADOS:

Os resultados desta pesquisa devem fornecer um maior conhecimento sobre a doença em questão, embora não seja esperado nenhum benefício direto além do resultado do(s) exame(s) realizado(s). A participação nesta pesquisa não deverá acarretar mudanças no atendimento médico nem no aconselhamento genético imediato. Os resultados da pesquisa serão fornecidos a você e à sua família pelo seu médico assistente ou pelos pesquisadores responsáveis e estarão também referidos no prontuário médico hospitalar.

SIGILO:

Todas as informações médicas, assim como os resultados dos exames e dos testes genéticos realizados nessa pesquisa, farão parte do prontuário médico e dos bancos de dados do projeto de displasias esqueléticas e serão submetidos aos regulamentos do Comitê de Ética em Pesquisa da FCM-UNICAMP.

O sigilo será mantido em todos os estudos colaborativos que possam advir dessa pesquisa por meio da utilização de um número de código para a identificação dos indivíduos participantes. Os resultados ou informações fornecidos devem ser utilizados para fins de publicação científica, no entanto, nenhum nome será utilizado.

FORNECENDO INFORMAÇÃO ADICIONAL:

A qualquer momento é possível obter informações adicionais relativas a presente pesquisa. A Dra. Denise P Cavalcanti e os seus colaboradores estarão disponíveis para responder às suas questões e/ou preocupações pelo Tel: (19) 3521-0385, Campinas, SP. Pacientes podem, se

preferirem, entrar em contato com a colaboradora do projeto Dra. Erlane Marques Ribeiro pelo Tel: (85) 9998-6577, Fortaleza, CE. Em caso de denúncias ou reclamações sobre sua participação no estudo, relacionadas aos aspectos éticos da pesquisa você pode entrar em contato com a secretaria do Comitê de Ética em Pesquisa (CEP): Rua: Tessália Vieira de Camargo, 126; CEP 13083-887 Campinas – SP; telefone (19) 3521-8936; fax (19) 3521-7187; e-mail: cep@fcm.unicamp.br.

RECUSA OU DESCONTINUAÇÃO DA PARTICIPAÇÃO:

A sua participação, bem como dos seus familiares, na presente pesquisa é voluntária. Sendo assim, a qualquer momento você e/ou os seus familiares poderá(ão) se recusar a participar ou retirar o consentimento e interromper a participação na presente pesquisa sem comprometimento dos cuidados médicos que recebe(m) atualmente ou receberá(ão) no futuro.

De forma semelhante, os pesquisadores responsáveis podem interromper a participação de qualquer indivíduo da família nessa pesquisa a qualquer momento em que julgarem apropriado.

COLETA, FUTURAS PESQUISAS E ARMAZENAMENTO DO MATERIAL GENÉTICO:

O DNA, material genético obtido a partir de uma amostra de sangue [seu e/ou de outro(s) membro(s) de sua família] será(ão) usado(s) para esse estudo, se você concordar. Nenhum outro teste ou estudo poderá ser feito com as amostras de DNA, a menos que você concorde com isso. Você pode concordar com o uso da sua amostra de sangue apenas para esse estudo (opção 1 do termo de consentimento), ou também para outros estudos (opção 2 e 3). Você pode escolher ser consultado no futuro sobre sua permissão para uso de sua amostra em um outro estudo (opção 2) ou você pode dar sua permissão agora, sem ter que ser solicitado novamente (opção 3). Em qualquer caso, seu nome será conhecido apenas pelo investigador do estudo e ninguém mais. O que os médicos fazem em pesquisa médicas deve sempre ser aceito pelo Comitê de Ética, um grupo de pessoas (médicos, advogados e outros) que assegurará que o estudo tenha um bom propósito e seja realizado para atender o interesse das pessoas doentes.

Você tem o direito de se retirar desse estudo a qualquer momento. Se você escolher se retirar, nenhum outro estudo será conduzido com sua amostra.

Se você concordar em participar desse estudo, sua amostra será armazenada no Laboratório de Genética Molecular do Depto. de Genética Médica, FCM, UNICAMP sob a responsabilidade do investigador principal. Essas amostras serão usadas exclusivamente para os propósitos científicos descritos aqui. Você pode requerer a destruição da amostra em qualquer momento. Uma cópia dessa folha de informação será fornecida a você.



Faculdade de Ciências Médicas – Departamento de Genética Médica

Assinatura do Termo de Consentimento Livre e Esclarecido (TCLE)

Projeto de investigação clínico-molecular sobre displasias esqueléticas

Pesquisador Principal: Denise P Cavalcanti

A assinatura do(s) item(ns) abaixo indica que você foi informado tanto oralmente como pela folha de Informação sobre a pesquisa mencionada acima. Dessa forma, você confirma que leu e entendeu o conteúdo da folha de Informação e que voluntariamente concorda em participar desse estudo. Você também entendeu que pode questionar qualquer coisa sobre a pesquisa ao seu médico e/ou ao pesquisador e que a informação que lhe foi fornecida permitiu-lhe tomar uma decisão livre e completamente informada sobre a sua participação nesse estudo. Você tem o direito de sair desse estudo em qualquer momento sem qualquer consequência para você. Uma cópia desse termo de consentimento será providenciada para você.

- I. Escolhendo um dos itens abaixo você estará concordando, ou não, com a utilização do seu DNA e/ou o de seu(ua) filho(a) em outras pesquisas ou apenas na atual pesquisa. Em quaisquer das opções abaixo sua identidade será mantida confidencial (ASSINALE APENAS UM DOS ITENS ABAIXO):
- 1 ☐ Sua amostra de sangue e/ou de seu(ua) filho(a) pode ser usada apenas nesse projeto de pesquisa. Essas amostras não podem ser usadas em nenhum outro projeto. Você não quer ser contatado para outras permissões no futuro.
 - 2 ☐ Sua amostra de sangue e/ou de seu(ua) filho(a) pode ser usada apenas nesse projeto, e, para outros projetos com sua permissão. Se essas amostras puderem ser usadas em um outro projeto, você quer ser contatado para que solicitem sua permissão. Uma nova aprovação do Comitê de Ética é necessária para usar sua amostra em um outro estudo.
 - 3 ☐ Sua amostra de sangue e/ou de seu(ua) filho(a) pode ser usada para qualquer outro propósito científico envolvendo este ou qualquer outro projeto. Você não quer ser contatado para permissão. Uma nova aprovação do Comitê de Ética é necessária para que a sua amostra seja utilizada em um outro estudo.
- II. Assinalando os itens abaixo você estará concordando, ou não, com a utilização de fotos clínicas.
- ☐ Utilização de fotos clínicas apenas para estudo
- ☐ Utilização de fotos clínicas para eventuais apresentações em publicações científicas

Nome:

Data de nascimento:

Assinatura:

Se o participante for menor de idade ou um adulto legalmente incapaz:

Assinatura do(s) pai(s) ou responsável legal:

Relação com o participante:

Local e data:

Médico que encaminha

A assinatura abaixo indica que o médico que encaminhou, forneceu uma explicação apropriada sobre a pesquisa, chamou atenção sobre as limitações assinaladas na folha de informação e respondeu as questões dos participantes.

Nome completo:**Assinatura:****Local e data:****Carimbo médico**

Pesquisador principal:

Denise P Cavalcanti
Departamento de Genética Médica
FCM – UNICAMP
R. Tessália V. De Camargo, 126
13083-887 Campinas, SP

Assinatura:**Local e data:**