

**DESENVOLVIMENTO DE TESTES MOLECULARES PARA A
INVESTIGAÇÃO PRÉ- E PÓS-SINTOMÁTICA DE DOENÇAS
MONOGÊNICAS CAUSADAS POR EXPANSÃO DE REPETIÇÕES
TRINUCLEOTÍDICAS INSTÁVEIS**

LAÍS GOMES SARLO

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY
RIBEIRO
CAMPOS DOS GOYTACAZES - RJ
MAIO DE 2013**

**DESENVOLVIMENTO DE TESTES MOLECULARES PARA A
INVESTIGAÇÃO PRÉ- E PÓS-SINTOMÁTICA DE DOENÇAS
MONOGÊNICAS CAUSADAS POR EXPANSÃO DE REPETIÇÕES
TRINUCLEOTÍDICAS INSTÁVEIS**

LAÍS GOMES SARLO

Dissertação de Mestrado apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro como parte das exigências para obtenção do título de Mestre em Biociências e Biotecnologia, sob orientação do Prof. Dr. Enrique Medina-Acosta.

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY
RIBEIRO
CAMPOS DOS GOYTACAZES - RJ
MAIO DE 2013**

**DESENVOLVIMENTO DE TESTES MOLECULARES PARA A
INVESTIGAÇÃO PRÉ- E PÓS-SINTOMÁTICA DE DOENÇAS
MONOGÊNICAS CAUSADAS POR EXPANSÃO DE REPETIÇÕES
TRINUCLEOTÍDICAS INSTÁVEIS**

LAÍS GOMES SARLO

Dissertação de Mestrado apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro como parte das exigências para obtenção do título de Mestre em Biociências e Biotecnologia, sob orientação do Prof. Dr. Enrique Medina-Acosta

Aprovada em 09 de Maio de 2013.

Comissão Examinadora:

Dra. Carmen Lucia Antão Paiva (Doutora em Genética) - UNIRIO

Dr. Álvaro Fabrício Lopes Rios (Doutor em Genética e Epigenética Animal) - UENF

Dr. Victor Martin Quintana Flores (Doutor em Biociências e Biotecnologia) – UENF

Dr. Enrique Medina-Acosta (Doutor em Parasitologia Médica e Molecular) - UENF
(Orientador)

Obrigada a todos que me acompanharam nesse longo caminho e me ajudaram a chegar aqui. Meus agradecimentos...

- ...primeiramente a Deus por nunca me deixar fraquejar e estar sempre iluminando e me guiando pelos caminhos desta jornada...
- ...aos meus pais por me darem todo o apoio e me escutarem sempre com tanto carinho, me trazendo sempre uma palavra de sabedoria. Sem vocês nada disso seria possível...
- ...aos meus irmãos por acreditarem sempre em mim e por me darem sempre seu apoio...
- ...à minha família por estar sempre ao meu lado, mesmo eu estando longe deles todo esse tempo, e por sempre estarem torcendo por mim. Obrigada por me dar este grande exemplo de honestidade, caráter, e valor...
- ...ao Diogo, meu namorado e meu grande amigo... Obrigada pela paciência, pelo incentivo, pela força e principalmente pelo carinho e por sempre acreditar em mim até quando nem eu mesma acreditava. Valeu a pena toda distância e valeu a pena esperar... Porque colheremos juntos, os frutos dessa espera...
- ...ao Professor Enrique pela paciência na orientação, ensinamentos e incentivo que tornaram possível a conclusão desta dissertação. Agradeço a oportunidade e todo o aprendizado...
- ...ao Filipe pelas por sempre estar disposto a transmitir seus conhecimentos e a ajudar sempre...
- ...a todos os colegas do NUDIM pelo convívio, principalmente à Milena por ser meu braço direito, à Claudia, Graziela, Fabricio, Katrine, e à Priscilla, por terem sempre uma palavra amiga e por estarem sempre dispostos a ajudar...
- ...a minhas grandes amigas Milena e Sâmella pela companhia constante, risadas e conversas infinitas. É com vocês que compartilho angústias, alegrias, felicidades e tantas outras coisas que só amizade faz..
- ...aos eternos amigos de faculdade... Helder, Magda e Viviane... Pelos estudos e apoio...
- ...aos meus mais fiéis amigos de sempre Samily, Maria Caroline, Keli e Gabriel pelo incentivo, força, amizade, carinho e por nunca se esquecerem de mim. Só vocês entendem o meu sumiço, a falta de tempo, o cansaço, a necessidade de isolamento que a escrita exige...
- ...aos membros da banca por aceitarem meu convite...
- ...aos professores da UENF que foram tão importantes na minha vida acadêmica...
- ...à CAPES pela bolsa concedida para que fosse possível à realização desse sonho...
- ...e a todos que contribuíram também indiretamente para a conclusão deste trabalho.

“Não se deve ir atrás de objetivos fáceis. É preciso buscar o que só pode ser alcançado por meio dos maiores esforços.”

(Albert Einstein 1879-1955)

“Em algum lugar, algo incrível está esperando para ser descoberto.”

(Carl Sagan 1934-1996)

ÍNDICE DE FIGURAS E GRÁFICOS

Figura 1. Expansão/contração da sequência repetitiva na fita tardia	2
Figura 2. Relação entre o número de cópias da repetição e o número de imperfeições na sequência repetida.....	3
Figura 3. Modelo para a relação entre o número de cópias herdadas e o tempo necessário entre gerações para que a doença se manifeste.	4
Figura 4. Localização esquemática dos loci das doenças associadas a mutações dinâmicas categorizadas pelo posicionamento em regiões codificadoras e não codificadoras	6
Figura 5. Mutações dinâmicas e instabilidade da repetição	7
Figura 6. Os processos associados com a instabilidade da repetição.	8
Figura 7. Esquema das sequências dos iniciadores 2 e 3 respectivamente.....	28
Figura 8. Detalhes do funcionamento da técnica de <i>triplet repeat primed PCR</i> (TP-PCR).....	31
Figura 9. Informações para a interpretação de um eletroferograma.....	32
Figura 10. Esquema genérico para estimar o número de repetições trinucleotídicas no gene de interesse.....	33
Figura 11. Sequência do produto de amplificação <i>in silico</i> para o marcador D12S112036658	49
Figura 12. Sequência do produto de amplificação <i>in silico</i> do marcador D4S3076583 para a técnica de TP-PCR.....	51
Figura 13. Eletroferogramas representativos dos perfis alélicos de amplificação sob condições otimizadas.	55
Figura 14. Esquema para detecção do número de repetições no gene <i>AR</i> , relacionado com a Doença de Kennedy, utilizando a QF-PCR	56
Figura 15. Esquema para detecção do número de repetições utilizando a versão <i>tetraprimer</i> do ensaio de TP-PCR	57
Figura 16. Esquema para detecção do número de repetições utilizando a versão <i>triprimer</i> do ensaio de TP-PCR	58
Figura 17. Eletroferograma representativo dos perfis dos produtos de amplificação da região CAG do gene <i>ATXN3</i>	59

Figura 18. Eletroferograma representativo dos perfis dos produtos de amplificação da região CAG do gene <i>AR</i>	60
Figura 19. Eletroferograma representativo dos perfis dos produtos de amplificação da região CGG do gene <i>FMR1</i>	61
Figura 20. Sequência do produto de amplificação <i>in silico</i> para o marcador DXS66765056, desenhado para diagnóstico da Doença de Kennedy.....	70
Figura 21. Heredograma do núcleo familiar 226 com uma expansão no alelo 274/280 no gene <i>AR</i>	72
Gráfico 1. Frequência de distribuição das repetições (CAG) _n no gene <i>ATXN3</i> , relacionado à Doença de Machado-Joseph.	68
Gráfico 2. Frequência de distribuição das repetições (CGG) _n no gene <i>FMR1</i> , relacionado à Síndrome do X-frágil.	69
Gráfico 3. Frequência de distribuição das repetições (CAG) _n no gene <i>AR</i> em 100 indivíduos	70

ÍNDICE DE TABELAS

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano	35
Tabela 2. Informações referentes às doenças poliglutamínicas.	44
Tabela 3. Informações referentes às doenças não-poliglutamínicas.	45
Tabela 4. Relação dos iniciadores gene e repetição-específicos desenhados e tamanho dos <i>amplicons</i> dos marcadores para as quinze doenças causadas por expansão de repetições trinucleotídicas.....	47
Tabela 5. Condições otimizadas para a QF-PCR dos cinco marcadores testados <i>in vitro</i>	53
Tabela 6. Número de nucleotídeos não variáveis para os marcadores desenhados para cada uma das quinze doenças causadas por expansão de trinucleotídeos.....	56
Tabela 7. Frequências alélicas para os marcadores D14S92537354, DXS66765056 e DXS146993454 para os genes <i>ATXN3</i> , <i>AR</i> , e <i>FMR1</i> respectivamente	63
Tabela 8. Padrões de combinações entre os alelos do gene <i>ATXN3</i> encontrados na população normal.....	65
Tabela 9. Padrões de combinações entre os alelos do gene <i>AR</i> encontrados na população normal.....	66
Tabela 10. Padrões de combinações entre os alelos do gene <i>FMR1</i> encontrados na população normal.....	67
Tabela 11. Evidência genética a favor da maternidade.	71

LISTA DE ABREVIATURAS

A - Adenina;

AR – *Androgen receptor protein*; Proteína receptora de andrógenos;

BLAST - *Basic local alignment search tool*;

C - Citosina;

CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior;

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico;

DNA - *Deoxyribonucleic acid*; Ácido Desoxirribonucléico;

DH - *Huntington Disease*; Doença de Huntington;

DM1 - Distrofia Miotônica tipo 1

DRPLA - *Dentatorubralpallidoluysian Atrophy*; Atrofia dentatorubropalidoluisiana;

FAPERJ - Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro;

FRAXE - *X-linked mental retardation*; Retardo Mental ligado ao Cromossomo X;

FRDA - *Friedreich's Ataxia*; Doença de Friedreich;

FXS - *Fragile X Syndrome*; Síndrome do X-frágil;

G - guanina;

HTT - *Huntintin protein*; Proteína Huntingtina;

iPCR – Reação em Cadeia da Polimerase *in silico*;

Kb - Kilobase

MJD - *Machado-Joseph Disease*; Doença de Machado-Joseph;

mRNA - RNA mensageiro

NUDIM - Núcleo de Diagnóstico e Investigação Molecular;

nt - Nucleotídeo;

OIZ - *Okazaki Initiation Zone*; Zona de Iniciação do Fragmento de Okazaki;

OMIM - *Online Mendelian Inheritance in Man*;

pb - Pares de Base;

PCR - *Polymerase Chain Reaction*; Reação em Cadeia da Polimerase;

Poli-Q - Poliglutamina;

Q - Glutamina

QF-PCR - *Quantitative Fluorescent Polymerase Chain Reaction*; Reação Quantitativa por Fluorescência em Cadeia da Polimerase;

RNA - Ácido Ribonucléico;

SBMA - *Spinobulbar Muscular Atrophy*; Atrofia Muscular Bulbo Espinhal

SCA 1 - *Spinocerebellar Ataxia type 1*; Ataxia Espinocerebelar tipo 1;

SCA 2 - *Spinocerebellar Ataxia type 2*; Ataxia Espinocerebelar tipo 2;

SCA 3 - *Spinocerebellar Ataxia type 3*; Ataxia Espinocerebelar tipo 3;

SCA 6 - *Spinocerebellar Ataxia type 6*; Ataxia Espinocerebelar tipo 6;

SCA 7 - *Spinocerebellar Ataxia type 7*; Ataxia Espinocerebelar tipo 7;

SCA 8 - *Spinocerebellar Ataxia type 8*; Ataxia Espinocerebelar tipo 8;

SCA 12 - *Spinocerebellar Ataxia type 12*; Ataxia Espinocerebelar tipo 12;

SCA 17 - *Spinocerebellar Ataxia type 17*; Ataxia Espinocerebelar tipo 17;

STR - *Short Tandem Repeat*; Repetição Curta em Tandem;

UENF - Universidade Estadual do Norte Fluminense;

T - Timina.

TNRs - *Trinucleotide Repeats*"; Repetições Trinucleotídicas;

TP-PCR - *Triplet Repeat Primed Polymerase Chain Reaction*; Reação em Cadeia da Polimerase com iniciadores de repetições trinucleotídicas;

TRDB - *Tandem Repeats Database*;

UCSC – Universidade da Califórnia em Santa Cruz;

UENF – Universidade Estadual do Norte Fluminense.

SUMÁRIO

ÍNDICE DE FIGURAS E GRÁFICOS	vii
ÍNDICE DE TABELAS.....	ix
LISTA DE ABREVIATURAS	x
RESUMO.....	xv
ABSTRACT.....	xvi
1. INTRODUÇÃO	1
1.1. Mutações dinâmicas	1
1.2. Doenças humanas associadas a mutações dinâmicas.....	4
1.3. Doenças monogênicas associadas à expansão de TNRs	8
1.3.1. Doenças poliglutamínicas	10
1.3.1.1. Ataxia espinocerebelar tipo 1.....	10
1.3.1.2. Ataxia espinocerebelar tipo 2 (Síndrome de Wadia-Swami).....	11
1.3.1.3. Ataxia espinocerebelar tipo 3 (Doença de Machado-Joseph).....	12
1.3.1.4. Ataxia espinocerebelar tipo 6.....	13
1.3.1.5. Ataxia espinocerebelar tipo 7.....	14
1.3.1.6. Ataxia espinocerebelar tipo 12.....	14
1.3.1.7. Ataxia espinocerebelar tipo 17	15
1.3.1.8. Atrofia Dentatorubropalidoluisana.....	15
1.3.1.9. Atrofia Muscular Bulbo Espinhal (Doença de Kennedy)	16
1.3.1.10. Doença de Huntington	16
1.3.2. Doenças não-poliglutamínicas.....	17
1.3.2.1. Ataxia de Friedreich	18
1.3.2.2. Ataxia espinocerebelar tipo 8.....	18
1.3.2.3. Distrofia Miotônica 1	19
1.3.2.4. Síndrome do X-frágil	19
1.3.2.5. Retardo Mental ligado ao Cromossomo X	20

1.4. Diagnóstico molecular de doenças causadas por expansões trinucleotídicas	21
2. OBJETIVOS	24
2.1. Objetivo Geral	24
2.2. Objetivos Específicos	24
3. METODOLOGIA	25
3.1. Material biológico	25
3.2. Sujeitos da pesquisa	25
3.3. Aspectos éticos da pesquisa	25
3.4. Extração de DNA	25
3.5. Quantificação do DNA genômico	26
3.6. Métodos	26
3.6.1. Análise exômica e de regiões reguladoras 5' para busca de repetições trinucleotídicas	26
3.6.2. Desenho dos iniciadores	26
3.6.3. Padronização da QF-PCR convencional	28
3.6.4. PCR com iniciador de repetições trinucleotídicas (TP-PCR)	28
3.6.5. Eletroforese capilar	31
3.6.6. Identificação e caracterização dos perfis alélicos	32
3.6.7. Frequências alélicas e dados populacionais	33
4. RESULTADOS	34
4.1. Varredura exômica e de regiões reguladoras 5' de TNRs	34
4.2. Desenho de marcadores para as quinze doenças causadas por expansão de repetições trinucleotídicas	46
4.3. Padronização dos ensaios para três dos marcadores desenhados	52
4.3.1. Padronização de ensaio para QF-PCR	52
4.3.2. Padronização de ensaio para TP-PCR	57
4.4. Dados populacionais relativos aos marcadores	61
5. DISCUSSÃO	73

6. CONCLUSÕES	77
7. REFERÊNCIAS BIBLIOGRÁFICAS	78
ANEXO.....	89

RESUMO

As doenças neurológicas monogênicas causadas por expansões de repetições trinucleotídicas instáveis constituem em humanos um conjunto clinicamente heterogêneo de neuropatias hereditárias agrupadas em poliglutamínicas e não-poliglutamínicas. Na maioria destas doenças ocorre antecipação genética, determinada pela relação inversa, transgeracional entre o número de cópias da repetição, a idade do início e a severidade dos sintomas. A investigação molecular pré- e pós-sintomática direta destas doenças consiste na enumeração de cópias das repetições trinucleotídicas pelo sequenciamento, considerado teste padrão ouro. A investigação indireta das expansões envolve a hibridação de sondas (*Southern blotting*) ou a amplificação pela Reação em Cadeia da Polimerase (PCR) de longa distância. A PCR Quantitativa por Fluorescência (QF-PCR) tem se mostrado um método de diagnóstico molecular rápido das mutações dinâmicas. No entanto, estes métodos indiretos apenas estimam o número de repetições. Para a enumeração das repetições pode ser também utilizada a PCR com Iniciador de Repetições Trinucleotídicas (TP-PCR) na versão *triprimer* ou *tetraprimer*. O objetivo do presente trabalho foi desenvolver ensaios moleculares da QF-PCR e TP-PCR para a investigação pré- e pós-sintomática de doenças monogênicas causadas por expansão de repetições trinucleotídicas instáveis. Para identificar os *loci* foi realizada a mineração *in silico* de repetições trinucleotídicas por meio de varredura exômica e em regiões reguladoras 5' no genoma humano de referência. Foram identificados 74 *loci* de repetições trinucleotídicas em genes associados a doenças, mas destes, apenas quinze, estão relacionados com doenças monogênicas causadas por expansão de repetições trinucleotídicas. Para estes quinze *loci* foram desenvolvidos marcadores moleculares gene e repetição-específicos visando à genotipagem por QF-PCR e TP-PCR. Foram determinados os parâmetros populacionais para os marcadores relacionados com as doenças de Machado-Joseph e Kennedy e síndrome do X-frágil. As taxas de heterozigose para as repetições trinucleotídicas associadas a essas doenças foram 0,85, 0,84 e 0,29, respectivamente. Para os *loci* contendo repetições CAG nos genes *ATXN3* e *AR* foram observados 21 e 15 alelos, respectivamente. Para o *locus* CGG no gene *FMR1* foram observados 14 alelos. Alelos com pré-mutação foram observados apenas para o gene *AR*. Os núcleos familiares destes indivíduos com número de repetições acima do limiar de normalidade foram investigados, e apenas uma família apresentou expansão, mas não patogênica. Houve expansão transgeracional de duas repetições nesta família. Visando ao diagnóstico genético pré- e pós-sintomático, os marcadores identificados permitiram caracterizar as expansões de repetições trinucleotídicas instáveis.

Palavras-chave: Diagnóstico pré- e pós-sintomático, doenças monogênicas, expansão de repetições trinucleotídicas, mutações dinâmicas, QF-PCR, TP-PCR.

ABSTRACT

A set of fifteen clinically heterogeneous hereditary neurodegenerative disorders caused by expansion of trinucleotide unstable repeats in humans, are grouped into polyglutamine and non-polyglutamine. In most of these disorders genetic anticipation may occur determined by inverse relationship between transgenerational number copy of repeats, the age of onset and severity of symptoms. Pre- and post-symptomatic direct molecular investigation of these diseases consist in enumeration of trinucleotide repeats copies by sequencing, considered the gold standard test. The indirect investigation of expansions involves hybridization probes (Southern blotting) or amplification by long distance Polymerase Chain Reaction (PCR). Quantitative Fluorescence PCR (QF-PCR) has been proved to be an effective tool for a rapid molecular diagnostic method for dynamic mutations. However, these methods only estimate the repeats number. For the repeats enumeration Triplet Primer PCR (TP-PCR) in triprimer or tetraprimer version can also be used . The aim of this study was to develop molecular assays of QF-PCR and TP-PCR to investigate pre- and post-symptomatic of monogenic diseases caused by unstable expansion of trinucleotide repeats. To identify trinucleotide repeats *loci* in silico mining was performed by scanning exomic and 5' regulatory regions in the reference human genome. We identified 74 *loci* containing trinucleotide repeats in genes associated with diseases, but of these diseases, only fifteen were related with monogenic diseases caused by expansion of trinucleotide repeats. For these fifteen *loci*, gene and repeat-specific molecular markers were developed aiming genotyping by QF-PCR and TP-PCR. Population parameters were determined for markers associated with Machado-Joseph and Kennedy disease and Fragile X syndrome. The rates of heterozygosity for trinucleotide repeats associated with these diseases were 0.85, 0.84 and 0.29, respectively. For loci containing CAG, repeats were observed for *ATXN3* and *AR* genes, 21 and 15 alleles were detected, respectively. For CGG at the *FMR1* gene *locus* 14 alleles were observed. Pre-mutation alleles were observed only for the *AR* gene. The other members of families of individuals bearing number of repeats above the threshold of normality were investigated, and only one family showed an expansion, but it was not pathogenic. There were two replications of transgenerational expansion in this family. For pre-and post-symptomatic diagnosis, identified markers allowed the characterization of unstable trinucleotide repeats expansions.

Keywords: Dynamic mutations, monogenic disorders, pre- and post-symptomatic diagnostic, QF-PCR, trinucleotide repeats expansion, TP-PCR.

1. INTRODUÇÃO

1.1. Mutações dinâmicas

Nem todas as mutações causadoras de doenças são transmitidas de forma estável dos genitores aos seus descendentes (La Spada, 1997), por isso foi criado o termo “mutação dinâmica”, para distinguir sequências repetidas de DNA que possuem características únicas de instabilidade e expansão, das outras formas de mutação. (Hori *et al.*, 1999; Mirkin, 2007; Chen *et al.*, 2009; McMurray, 2010).

Ao contrário das mutações estáticas que são retidas nos tecidos somáticos e são transmitidas aos descendentes de maneira estável, as mutações que envolvem repetições são dinâmicas e seus produtos permanecem em processo de mutação nas células dos tecidos e entre as gerações (Pearson *et al.*, 2005).

A mutação dinâmica é na realidade, um processo que envolve múltiplos passos, ao invés de um único evento. O primeiro passo envolve raras e pequenas mudanças no número de cópias ou na composição da repetição, de maneira que um determinado número de repetições instáveis perfeito é alcançado (Hori *et al.*, 1999; Richards, 2001). E uma das características deste tipo de mutação, é que o tamanho da expansão inicial pode determinar a taxa de expansão futura (Mor-Cohen *et al.*, 1997).

O processo de formação deste tipo de mutações é afetado por uma variedade de elementos e fatores bem como aqueles que estão diretamente associados com expansão da repetição (denominados elementos *cis*) e os que possuem interação com a repetição, e apenas contribuem com a instabilidade (denominados fatores *trans*) (Richards, 2001; Chen *et al.*, 2009)).

O mecanismo que permite a formação de estruturas secundárias pode ocorrer durante a replicação, reparo e/ou recombinação do DNA, principalmente no processo que envolve a separação das fitas de DNA, desacoplamento e realinhamento da enzima polimerase, ou seja, no processo de replicação (Richards *et al.*, 1992).

Durante a progressão da forquilha de replicação, uma parte na fita tardia de DNA, chamada “zona de iniciação do fragmento de Okazaki” (OIZ) se mantém temporariamente como fita simples, para garantir a síntese coordenada das novas fitas, líder e tardia. É sugerido, portanto, que as repetições expansíveis formem estruturas secundárias (em forma de grampo, por exemplo) não usuais dentro da OIZ da fita tardia (Cleary *et al.*, 2002; Mirkin, 2006). Caso a estrutura secundária

esteja na fita nascente, significa que haverá expansão da repetição na nova fita. Mas, caso esteja na fita molde de DNA, haverá uma contração da sequência repetitiva (Cleary *et al.*, 2002). Caso o erro não seja reparado na fita nascente de DNA, como consequência haverá o surgimento de novos alelos (Kunkel, 1993) como mostrado na figura 1.

Em geral, alelos com um grande número de cópias livres de interrupções, que são aqueles com repetições perfeitas são mais instáveis que aqueles que possuem um menor número de cópias ou que contém interrupções, ou seja, com repetições imperfeitas (figura 2) (Hori *et al.*, 1999; Kozlowski *et al.*, 2010). Esta interrupção por uma repetição de sequência diferente desempenharia um papel fundamental na estabilidade das repetições do alelo expandido (Zoghbi, 1996).

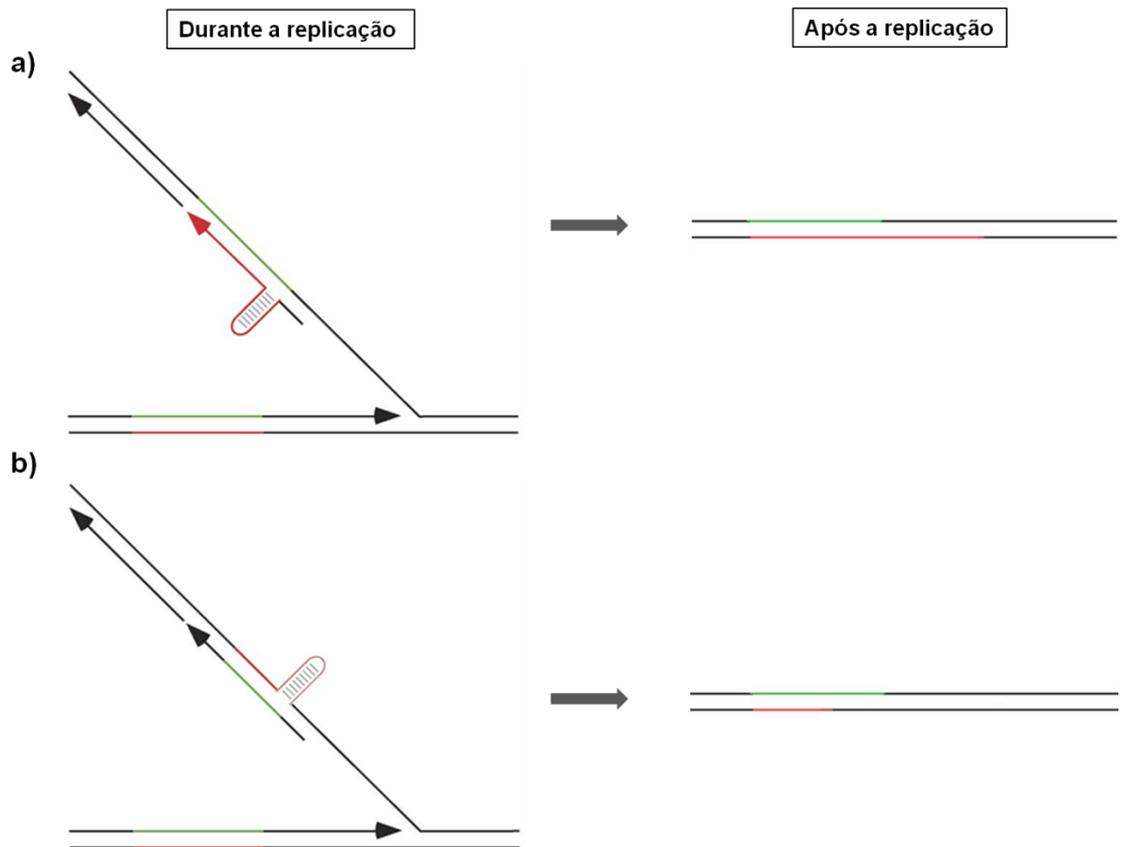


Figura 1. Expansão/contração da sequência repetitiva na fita tardia. a) Quando a estrutura secundária está localizada na fita nascente resulta em expansão da sequência repetitiva. **b)** Quando a estrutura secundária está localizada na fita molde resulta em contração da sequência repetitiva. A sequência em vermelho mostra a estrutura secundária e a sequência em verde, sua fita complementar, ambas as sequências são mostradas durante e após a replicação. Adaptado de Mirkin, 2006).

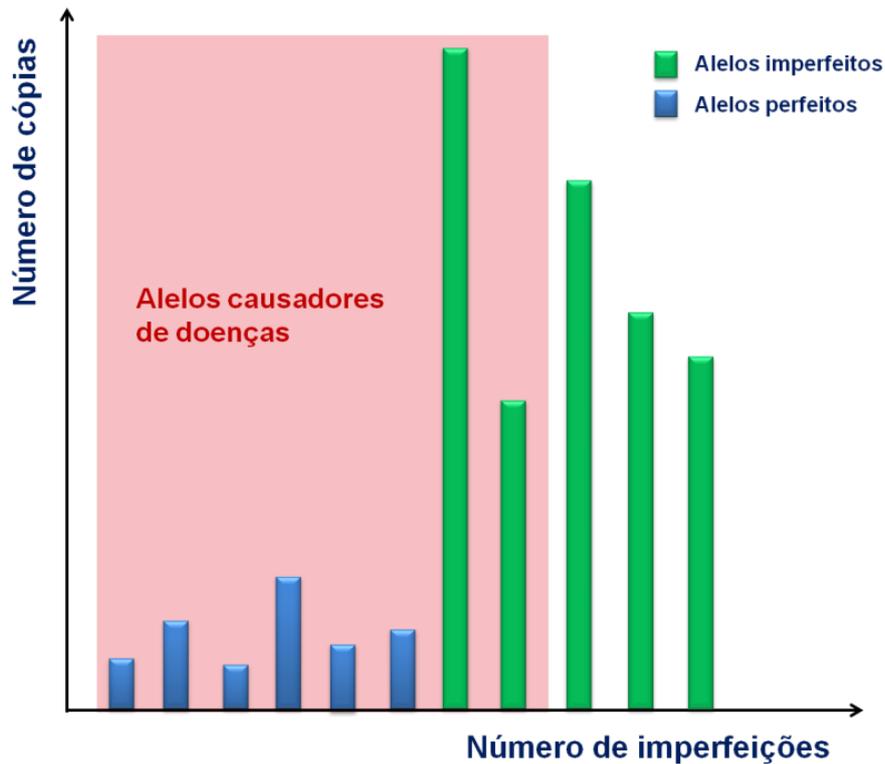


Figura 2. Relação entre o número de cópias da repetição e o número de imperfeições na sequência repetida. Os alelos perfeitos são mais instáveis e por isso mais suscetíveis a causarem doenças relacionadas à expansão de repetições.

As propriedades gerais deste tipo de mutação são: (a) taxa de mutação está relacionada à mudança (geralmente aumento) do número “perfeito” de cópias da repetição; (b) raros eventos únicos levam a um aumento do número perfeito de repetições de um determinado alelo. Estes têm números cópia perto do considerado normal e são um grupo de alelos instáveis com elevado risco de expansão causadora da doença; (c) doenças causadas por expansão no número de cópias da repetição, exibem uma relação entre o número de cópias da repetição e a severidade e idade do início dos sintomas (Richards *et al.*, 1997; Richards, 2001).

Juntas, essas propriedades são responsáveis pelo aumento da gravidade e diminuição da idade de início da doença nas gerações seguintes, em um fenômeno chamado de “antecipação genética” (Richards *et al.*, 1997; Pearson *et al.*, 2005). A penetrância das mutações responsáveis por estas doenças aumenta nas sucessivas gerações (Orr *et al.*, 2007).

Além disso, o tempo necessário, entre as gerações, para que a doença se manifeste pode representar o tempo que leva para o alelo com número de cópias associado à doença (pré-mutado) chegar a um número de cópias superior ao limiar

crítico dentre o número de células afetadas na população celular, como ilustrado na figura 3 (Richards, 2001).

A mudança no número de cópias pode contribuir com a relação observada em algumas doenças neurodegenerativas entre o número de cópias do *locus* e a idade de início dos sintomas. O tempo (t) que leva para a repetição atingir um limiar crítico pelo aumento incremental no número de cópias será determinado pelo número de cópias herdadas (n). Quanto maior for o n , menor é o tempo. O tempo (t) decorrido para levar a um número de cópias causador da doença é inversamente proporcional ao número de cópias iniciais (n), onde $t = k/n$ (k é determinado pelo conteúdo da repetição) (Richards, 2001).

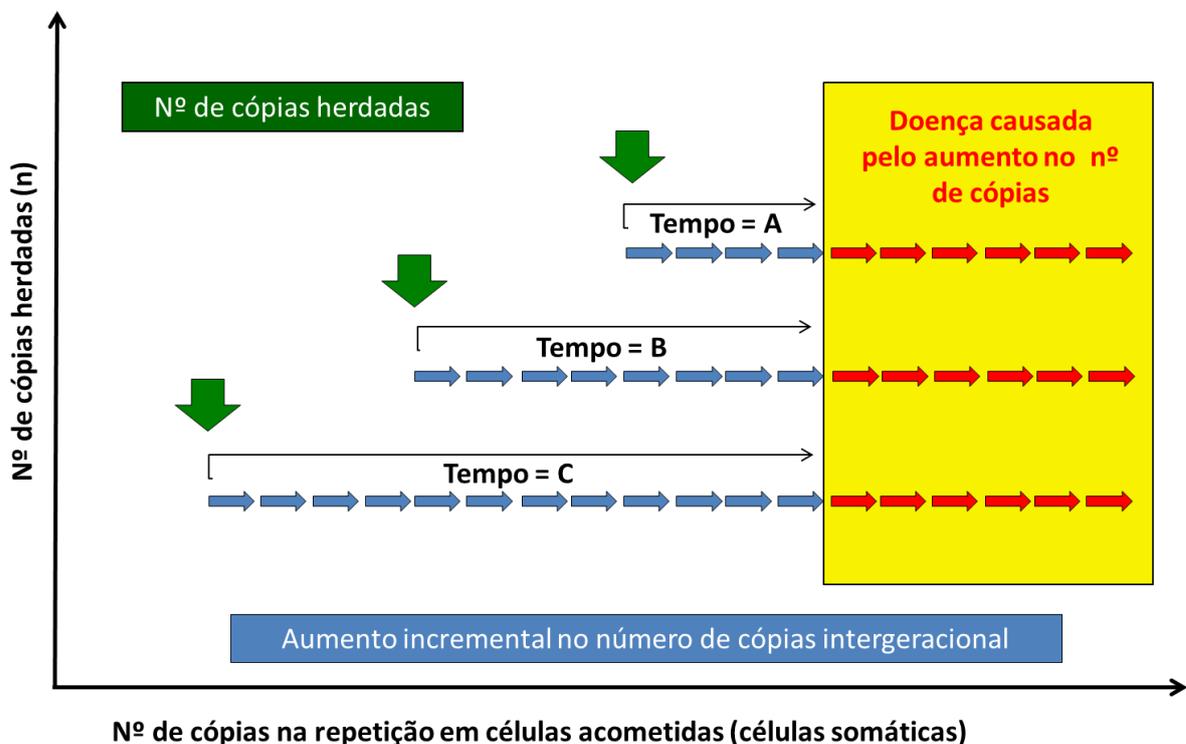


Figura 3. Modelo para a relação entre o número de cópias herdadas e o tempo necessário entre gerações para que a doença se manifeste. (Adaptado de Richards, 2001).

1.2. Doenças humanas associadas a mutações dinâmicas

O genoma humano é formado por uma estrutura molecular altamente compartimentalizada. Essa estrutura apresenta padrões complexos quanto à estrutura dos genes e a composição das sequências repetitivas e seu papel no desenvolvimento humano (Levy *et al.*, 2007).

As regiões codificadoras são definidas por uma variedade de funções relacionadas às sequências de DNA e RNA, como regulação gênica, transcrição do RNA, *splicing* do RNA e metilação do DNA. O sequenciamento do exoma humano (sequência completa de todos os éxons do genoma humano) revelou que há entre 25 a 32 mil genes com suas regiões reguladoras, que correspondem a apenas 2% de todo o genoma. Esses genes estão imersos em um enorme número de regiões não codificadoras, que correspondem a 98% do genoma (Lander *et al.*, 2001; Moreno *et al.*, 2011).

Os microssatélites, sequências curtas de até seis nucleotídeos repetidos em tandem, correspondem a 3% de todo o genoma, onde as maiores regiões de ocorrências destes tipos de repetições ocorrem em regiões intergênicas e intrônicas, mas uma fração destas repetições, predominantemente as repetições trinucleotídicas, aparecem em éxons e podem ser benéficas, neutras ou deletérias (Kozlowski *et al.*, 2010).

No genoma humano, o número de repetições em tandem de qualquer *locus*, varia entre os indivíduos, sendo essa variação um acontecimento normal e que raramente acarretam consequências negativas (Cleary *et al.*, 2005). A exceção está em uma subclasse de microssatélites que incluem certos tipos de repetições do tipo trinucleotídicas (“*trinucleotide repeats*” - TNRs), tetranucleotídeos, pentanucleotídeos, minissatélites (repetições de 7 a 64pb) e megassatélites (repetições com grande número de kb) (Pearson *et al.*, 2005), as quais estão relacionadas a pelo menos 36 doenças associadas a mutações dinâmicas (Cleary *et al.*, 2003; Pearson *et al.*, 2005; Lutz, 2007; Cannon *et al.*, 2012).

Dentre todas as doenças, algumas têm seus *loci* localizados em regiões não-codificadoras (denominadas regiões *non-coding*) e outras em regiões codificadoras (denominadas regiões *coding*) como mostrado na figura 4.

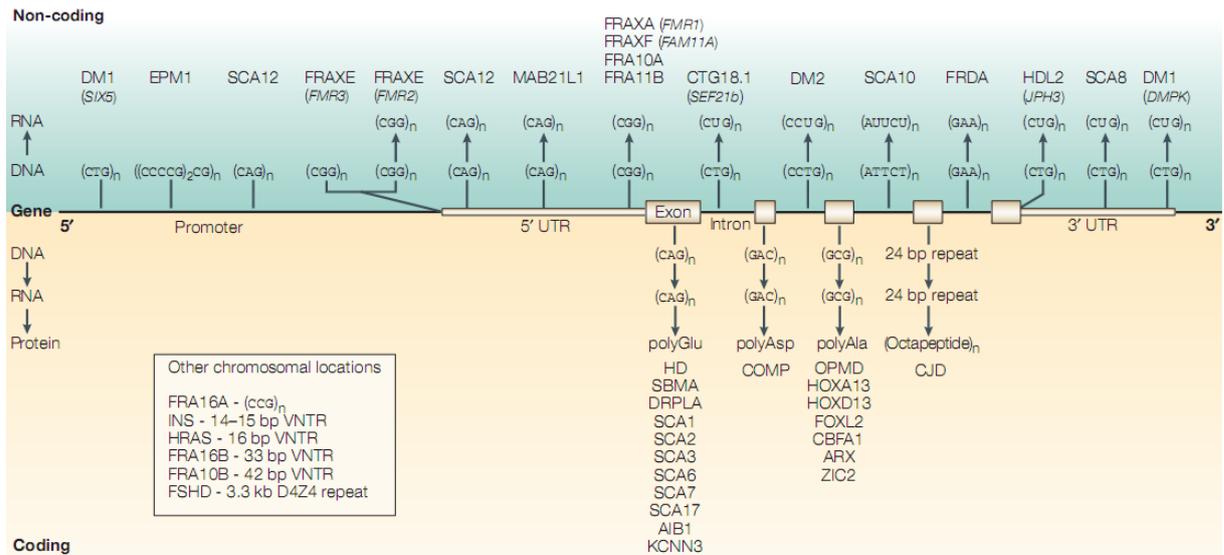


Figura 4. Localização esquemática dos loci das doenças associadas a mutações dinâmicas categorizadas pelo posicionamento em regiões codificadoras e não codificadoras. Na região azul são as doenças que possuem seus *locus* localizados em regiões não codificadoras, e em amarelo, aquelas localizadas em regiões codificadoras. As que possuem outras localizações cromossômicas estão sendo mostradas no quadro em destaque. **AIB1**, Amplificação em Câncer de Mama 1; **ARX**, Fator de transcrição relacionado a Aristaless; **CBFA1**, Fator de ligação nuclear α 1; **CJD**, Doença de Creutzfeld–Jakob; **COMP**, Proteína Oligomérica de Matrix Cartilaginosa; **DM2**, Distrofia Miotônica 2; **DRPLA**, Atrofia Dentatorubropalidoluisana; **EPM1**, Epilepsia Progressiva Miotônica 1; **FOXL2**, Forquilha de transcrição L2; **FRAXA/FRAXE/FRAXF/FRA10A/FRA10B/FRA11B/FRA16A/FRA16B**, Tipos Raros de Sítios Frágeis; **FRDA**, Ataxia de Friedreich; **FSHD**, Distrofia facioescapularhumeral; **HD**, Doença de Huntington; **HDL2**, Doença semelhante a Huntington 2; **HOXA13/HOXD13**, Fatores de transcrição A13 and D13; **HRAS**, Oncogene Viral v-Ha-ras homólogo a sarcoma de rato; **INS**, insulina; **JPH3**, junctofilina3; **KCNN3**, Canais de Potássio Ativados por Cálcio de Condutância Baixa; **MAB21L1**, Semelhante a mab-21 1; **OPMD**, Distrofia Muscular Oculofaríngea; **SBMA**, Atrofia muscular espinhal e bulbar; **SCA**, Ataxias Espinocerebelares; **VNTR**, Número variável de repetições em tandem; **ZIC2**, Família de Zinco membro 2 (Pearson *et al.*, 2005).

Para essas doenças, a instabilidade da repetição presente no *locus* está diretamente ligada com a idade de início dos sintomas, severidade e progressão da doença (Kennedy *et al.*, 2003; Wheeler *et al.*, 2003). No *locus* de cada doença há uma faixa normal do número de cópias acima da qual a repetição torna-se instável, ou seja, o número de cópias pode ser aumentado até onde a repetição manifeste uma doença. É chamada de pré-mutação (alelos intermediários) aqueles alelos que possuem o número de cópias acima do normal, mas que ainda é um número insuficiente para causar a doença (Richards *et al.*, 1997) (Figura 5).

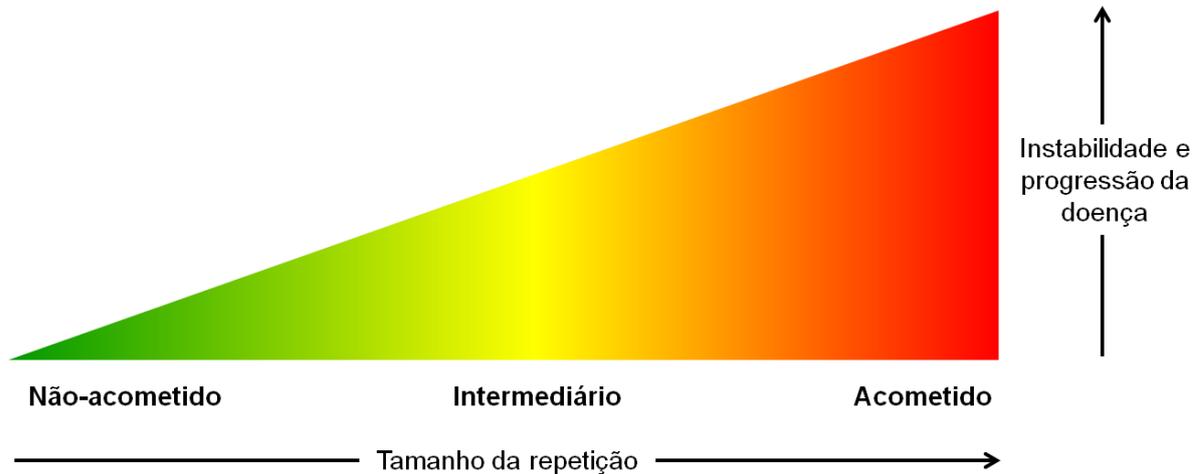


Figura 5. Mutações dinâmicas e instabilidade da repetição. A área amarela corresponde ao limiar do tamanho da repetição, onde há possibilidade de haver instabilidade. Acima desse limiar (área em vermelho), as repetições se tornam altamente instáveis, caracterizando as chamadas “mutações dinâmicas”, onde o tamanho da expansão está associado a início, severidade e progressão da doença. O número dessas repetições varia entre as doenças (Adaptado de Cleary *et al.*, 2005).

A instabilidade observada é limitada ao *locus* de cada doença, sugerindo um complexo mecanismo de mutação. Nos tecidos somáticos e germinativos, é provável que a instabilidade tenha surgido da participação de estruturas alternativas de DNA durante múltiplos processos que envolvem o metabolismo do DNA (replicação, reparo e recombinação), que podem contribuir individualmente ou em combinação, dependendo do tecido e do estágio de desenvolvimento. O esquema é mostrado na figura 6 (Cleary *et al.*, 2005).

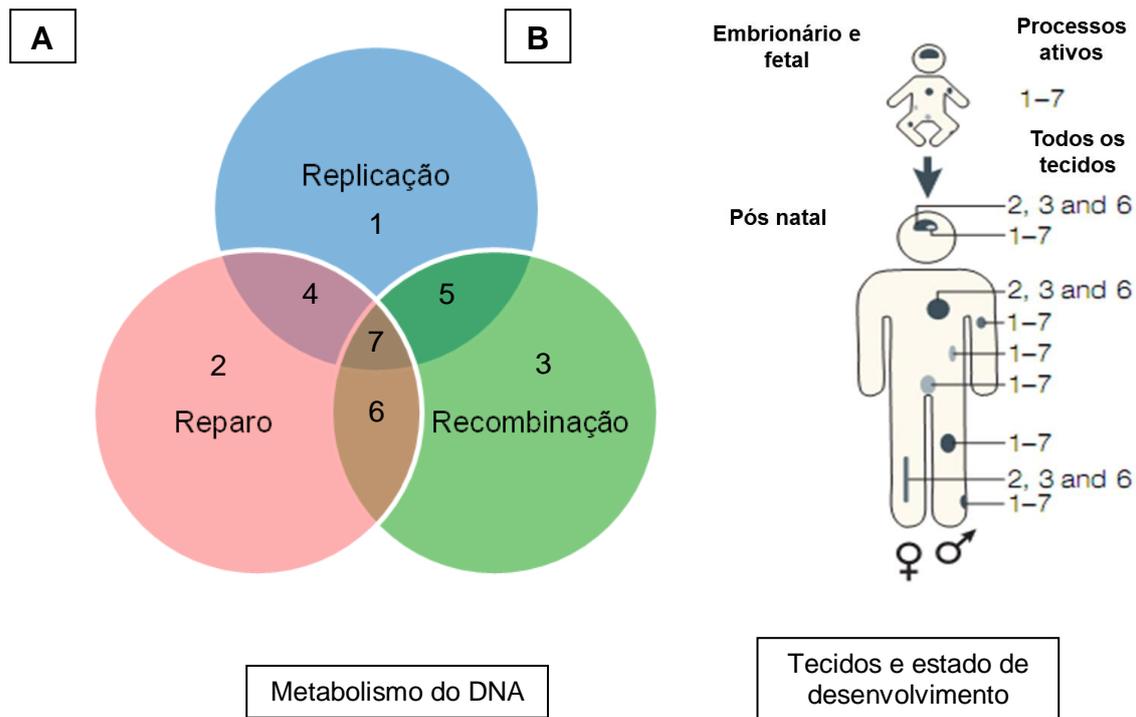


Figura 6. Os processos associados com a instabilidade da repetição. (A) Os processos metabólicos do DNA, replicação, reparo e recombinação, estão associados com a instabilidade da repetição separadamente (1-3) ou em conjunto com outros processos (4-7). (B) O envolvimento em vários processos metabólicos varia dependendo do tecido ou do estágio de desenvolvimento. Os números da figura B, correspondem aos processos da figura A (Adaptado de Pearson *et al.*, 2005).

1.3. Doenças monogênicas associadas à expansão de TNRs

Definem-se como monogênicas, os traços de doenças genéticas que dependem da variação alélica ou de mutação em nível de um único gene. Esse conceito de variação genética, normal ou patológica é designado como mendeliano (Cavagnari, 2011). No entanto, há a existência de doenças monogênicas que em definição não são mendelianas, por razão de mecanismo molecular da mutação ou da localização da proteína com a mutação (Peltonen *et al.*, 2006).

As mutações patogênicas que afetam o número de repetições nucleotídicas foram primeiramente descritas no início dos anos 90, quando as mutações causadoras da Síndrome do X-frágil (FRAXA) e da Atrofia muscular bulbo espinhal (SBMA) foram identificadas como expansões de repetições trinucleotídicas (Gatchel *et al.*, 2005). Essas doenças, no entanto, não podiam ser explicadas nos termos de genética mendeliana clássica (Mirkin, 2006), pois um dos princípios centrais da genética clássica mendeliana é justamente que mutação herdada é transmitida de maneira estável através das gerações (Mirkin, 2007), o que trazia grande conflito quanto ao diagnóstico destas doenças.

O fenômeno de antecipação genética foi descrito primeiramente em 1918 para a Distrofia miotônica 1, uma doença humana de caráter hereditário. Esta doença parecia ter um início mais precoce e a gravidade da doença aumentava à medida que o gene mutante era transmitido de uma geração para outra (Mirkin, 2006). Um padrão de herança semelhante foi detectado anos depois para outras doenças neurológicas, incluindo a Doença de Huntington, Ataxia de Friedreich, Atrofia muscular bulbo espinhal e diversas ataxias (Rosales-Reynoso *et al.*, 2009).

Essas doenças causadas por expansão de sequências repetidas possuem a característica de possuir um viés de transmissão onde as doenças associadas à transmissão paterna (geralmente as causadas por expansão da sequência CAG) há um maior risco de expansão e as que possuem alelos associados à transmissão materna apresentam fenótipos mais severos relacionados (Cumming e Zoghbi, 2000).

As TNRs correspondem a 0,1% dos microssatélites e são altamente polimórficas em tamanho. Além disso, possuem um elevado potencial de modulação de processos celulares, incluindo processos de transcrição, *splicing* e tradução (Kashi *et al.*, 2006). As TNRs possuem uma alta frequência de mutação (Hurles, 2005), o que pode levar ao aumento do risco de doenças associadas a esse tipo de repetição (Mirkin, 2007; Hannan, 2010).

Estruturas de repetições trinucleotídicas são importantes na formação da estrutura tridimensional do DNA e flexibilidade e curvatura da fita de DNA, que permitem a interação com elementos não-genéticos. Por isso, além de impedir a formação de proteínas, as repetições trinucleotídicas podem interferir na metilação e *imprinting*, degradação celular de proteínas por ubiquitinação, e processamento através do proteossomo, assim como ativação e inativação de genes pela maquinaria extra-nuclear (Lutz, 2007).

As doenças associadas à instabilidade de repetições trinucleotídicas são uma categoria de doenças causadas pelo aumento no número de trinucleotídicas em determinados genes (King, 2011), sendo o tipo mais comum de repetição causadora de doenças neurológicas (Orr *et al.*, 2007). Estas são divididas em duas categorias determinadas pelo tipo de repetição: doenças poliglutamínicas e não-poliglutamínicas. São descritas na literatura quinze doenças monogênicas causadas por expansão de *loci* exônicos e de regiões 5' reguladoras de repetições trinucleotídicas.

1.3.1. Doenças poliglutamínicas

As doenças poliglutamínicas são as mais comuns e estão associadas à expansão do trinucleotídeo CAG, a qual, quando presente na região de codificação de um gene, codifica para o aminoácido glutamina (Q). Portanto, essas doenças são referidas como doenças poliglutamínicas (doenças PoliQ) (King, 2011).

Os peptídeos de poliglutamina estão associados a doenças neurodegenerativas autossômicas dominantes. Um dos efeitos sobre a patogênese dessas doenças está associado ao fato de proteínas mutantes expandidas contendo poliglutaminas em sua composição possuírem efeito de ganho tóxico (Lutz, 2007). Além disso, repetições do tipo CAG estão associadas com transcrição gênica, com vias de interferência na sobrevivência celular, com morte celular programada, e com agregação nuclear e citoplasmática nos neurônios com consequentes alterações no sistema de transporte axonal rápido (Morfini *et al.*, 2005).

Há evidência genética que indica um mecanismo de ganho de função crítico para cada doença poliglutamínica, devido à habilidade dos peptídeos de poliglutamina, quando em grande quantidade, formarem agregados e grandes inclusões (Ross *et al.*, 2004), podendo causar mudança na conformação da proteína, e com isso, alteração na interação normal com outras proteínas celulares. Essas interações alteradas resultam em disfunção neuronal, o que leva a neurodegeneração e perda neuronal (Cummings *et al.*, 2000; Orr *et al.*, 2007).

Das doenças causadas por expansão de glutaminas, localizadas em éxons, nove (SCA1, SCA2, SCA3, SCA6, SCA7, SCA17, DRPLA, SBMA e HD) das dez possuem sua patogênese relacionada ao mecanismo de ganho de função, e apenas uma (SCA12) está associada a mecanismos patogênicos desconhecidos.

1.3.1.1. Ataxia espinocerebelar tipo 1

A ataxia espinocerebelar tipo 1 (SCA1) é uma doença autossômica dominante, neurodegenerativa progressiva. É causada por expansão de trinucleotídeos CAG no éxon 8 do gene *ATXN1*, localizado no braço curto do cromossomo 6 (região 6p22.3), e codifica a proteína ataxina-1 (Trott *et al.*, 2010). O número de repetições varia de 6 a 39 em alelos normais e de 41 a 83 nos alelos mutados, sendo que ambos os casos, tanto os alelos normais como os expandidos, são transcritos (Gatchel *et al.*, 2005).

Os sintomas iniciam-se por volta dos 33 anos de idade, podendo haver uma variação, dependendo do número de repetições CAG presente no alelo com a mutação, havendo uma correlação inversa entre o início da doença e o número de repetições (fenômeno de antecipação) (Trott *et al.*, 2010). Os sintomas dessa doença envolvem ataxia, fala arrastada, espasticidade e dificuldade cognitiva (Gatchel *et al.*, 2005). O processo neurodegenerativo que ocorre na SCA1 acontece devido a erros na transcrição e tradução de sequências expandidas do trinucleotídeo CAG presente na região codificadora do gene, o que resulta na síntese de um produto gênico alterado, com uma sequência de poliglutamina expandida (Orr *et al.*, 1993), e conseqüentemente esse gene irá codificar para uma proteína defectiva. A ataxina-1 é expressa principalmente no sistema nervoso central, no entanto as alterações mais graves e mais frequentes quando esta proteína é super-expressa, estão relacionadas à neurodegeneração devido a interações anormais com outras moléculas nos neurônios envolvidos (Zoghbi, 1995), podendo ocorrer perda de células de Purkinje no córtex cerebelar, assim como a perda de neurônios no núcleo olivar inferior, no núcleo dentado e no núcleo vermelho. Isso mostra que *ATXN1* também é expresso no núcleo dos neurônios (Servadio *et al.*, 1995).

Até 2006, no Brasil, a maior parte dos indivíduos relatados acometidos pela SCA1 eram da região sul do país (Trott *et al.*, 2006).

1.3.1.2. Ataxia espinocerebelar tipo 2 (Síndrome de Wadia-Swami)

A ataxia espinocerebelar tipo 2 (SCA2), também conhecida como Síndrome de Wadia-Swami, é uma doença autossômica dominante, neurodegenerativa, causada por expansão do trinucleotídeo CAG no éxon 1 do gene *ATXN2*, localizado braço longo do cromossomo 12 (região 12q24.12), e codifica a proteína Ataxina-2 (Gispert *et al.*, 1993).

Em indivíduos normais, a repetição CAG não apresenta polimorfismo em tamanho, variando entre 14 e 31 repetições. Esse número normal de repetições, é frequentemente interrompida por uma ou mais repetições CAA, e em contraste, os alelos expandidos, que possuem número de repetições entre 32 e 200, não possuem esse trinucleotídeo em sua sequência (Cancel *et al.*, 1997; Geschwind *et al.*, 1997). Esta interrupção pelo trinucleotídeo CAA desempenharia um papel fundamental na estabilidade das repetições do alelo expandido em SCA2 (Zoghbi, 1996).

A idade de início média da SCA2 se dá em torno dos 30 anos de idade (Trott *et al.*, 2010), onde os casos juvenis da doença estão associados à presença de 130 a 200 repetições CAG no alelo expandido (Mao *et al.*, 2002). Para esta doença, também há a ocorrência do fenômeno de antecipação, sendo a transmissão paterna associada às antecipações mais expressivas (Trott *et al.*, 2010). Os sintomas relacionados à SCA2 são ataxia, polineuropatia, diminuição dos reflexos e variante infantil com retinopatia (Gatchel *et al.*, 2005).

O gene *ATXN2* tem sido associado a diversas funções biológicas, como apoptose celular, sinalização mediada por receptor, formação de filamentos de actina, e secreção (Huynh *et al.*, 1999). A ataxina-2 é encontrada no retículo endoplasmático, complexo de golgi, na membrana plasmática, em interação com proteínas de tradução do mRNA e de endocitose. As principais células afetadas pela presença do alelo expandido, são as células de Purkinje (Lastres-Becker *et al.*, 2008).

A SCA2 possui sua maior frequência relatada no leste da Índia (Cellini *et al.*, 2001). No Brasil, diversas famílias com acometidos, foram encontradas no sul do país (Trott *et al.*, 2006).

1.3.1.3. Ataxia espinocerebelar tipo 3 (Doença de Machado-Joseph)

A ataxia espinocerebelar tipo 3 (SCA3) ou Doença de Machado-Joseph (MJD) é o tipo de ataxia com maior frequência na população mundial, com uma prevalência mundial de aproximadamente 30% dentre todas as ataxias espinocerebelares (Schols *et al.*, 2004). A SCA3 é uma doença neurodegenerativa autossômica, dominante e está relacionada ao éxon 8 do gene *ATXN3* ou *MJD1*, localizado na região 14q32.12, e codifica a proteína ataxina-3. (Burk *et al.*, 1996; Schmitt *et al.*, 2003).

ATXN3 possui uma região codificante que apresenta repetições do tipo CAG. Expansões de CAG podem causar esta doença, onde a presença de 12-40 repetições caracteriza alelos normais, não expandidos, enquanto a presença de 52-86 repetições está relacionada à ocorrência de SCA3 (Gatchel *et al.*, 2005). Para SCA3 também ocorre o fenômeno da antecipação, sendo que a instabilidade das repetições expandidas é observada principalmente na transmissão paterna (Lokkegaard *et al.*, 1998).

A repetição instável CAG na região codificadora, quando expandida, causa a alongação de resíduos de poliglutamina próxima a região carboxi-terminal da ataxina-3 (Kawaguchi *et al.*, 1994). A agregação de proteínas intranucleares, pode alterar a interação entre proteínas e causar alteração na expressão gênica (Schmitt *et al.*, 2003). A produção de uma proteína ataxina-3 super expressa leva a morte neuronal preferencialmente os da região subcortical, núcleo dentado, núcleo subtalâmico e da medula espinhal (Rosenberg, 1992).

A SCA3 é caracterizada clinicamente por ataxia cerebelar, parkinsonismo e espasticidade, e os sintomas iniciam-se por volta dos 40 anos de idade, sendo que os primeiros apresentados são desequilíbrio e alterações da marcha (Lokkegaard *et al.*, 1998). A morte pela doença se dá em até 15 anos em média a partir do início dos sintomas (Trott *et al.*, 2006).

1.3.1.4. Ataxia espinocerebelar tipo 6

A ataxia espinocerebelar tipo 6 (SCA6) é uma doença neurodegenerativa de herança dominante autossômica, caracterizada por uma ataxia progressiva, disartria e falta de coordenação motora, devido à atrofia do cerebelo, devido principalmente a perda neuronal limitada exclusivamente às células de Purkinje (Restituito *et al.*, 2000). Algumas famílias com acometidos pela SCA6 foram encontradas no Brasil, apresentando ataxia grave dos membros (Trott *et al.*, 2006).

A SCA6 é causada por expansão da sequência CAG no éxon 47 do gene *CACNA1A*, localizado na região 19p13.2 que codifica a subunidade alfa 1A do canal de cálcio neuronal (Zhuchenko *et al.*, 1997), que está presente em todo o encéfalo, mas expresso principalmente no cerebelo (Restituito *et al.*, 2000). O alelo normal deste gene possui menos que 18 repetições e entre 20 e 33 repetições CAG o alelo é considerado expandido (Gatchel *et al.*, 2005). A expansão de glutaminas na terminação carboxila da proteína levaria a uma disfunção contínua do canal de cálcio, levando a uma doença progressiva (Zhuchenko *et al.*, 1997).

Há a ocorrência do fenômeno de antecipação em SCA6, mas em menos frequência quando comparado às outras ataxias espinocerebelares, devido a uma estabilidade transgeracional observada exclusivamente para esta ataxia (Zhuchenko *et al.*, 1997).

1.3.1.5. Ataxia espinocerebelar tipo 7

A ataxia espinocerebelar tipo 7 (SCA7) caracteriza-se fundamentalmente pela associação de uma ataxia cerebelar e uma distrofia muscular, ambas progressivas. A doença afeta o cerebelo e leva a uma degeneração da retina, mas também acomete muitas outras estruturas do sistema nervoso central com o progresso da doença.

O gene da SCA7, o *ATXN7* está localizado no braço curto do cromossomo 3, no *locus* 3p12-p12.1 apresentando uma repetição CAG que está expandida em indivíduos afetados. Os alelos expandidos apresentam de 36 a 460 repetições e os normais variam de 4 a 17 repetições. A expansão é traduzida em um trato de poliglutamina na proteína ataxina 7, produto do gene da SCA7. O alelo expandido é bastante instável na meiose, havendo forte correlação negativa entre o tamanho da repetição CAG e a idade de início. Há também uma correlação significativa entre o tamanho da expansão e a taxa de progressão da doença. A idade de início média para SCA7 fica em torno dos 30 anos, embora varie dos 3 meses de vida aos 65 anos. O sintoma inicial é a ataxia cerebelar no caso da doença surgir após os 30 anos.

A sequência CAG repetida é particularmente instável na SCA7, e mutações de novo podem ocorrer na transmissão de alelos intermediários, ou pré-mutações (18 a 35 repetições CAG), o que explica a persistência de SCA7, apesar da grave antecipação que poderia levar à extinção da doença (Freund *et al.*, 2009).

1.3.1.6. Ataxia espinocerebelar tipo 12

Indivíduos acometidos pela ataxia espinocerebelar tipo 12 (SCA12) possuem 66 a 78 repetições CAG na região promotora do gene *PPP2R2B* que se localiza em 5q32. Alelos normais, possuem de 7 a 28 repetições (Bahl *et al.*, 2005). Essa expansão do número de repetições normais causa uma desregulação da atividade de uma proteína denominada PP2A, uma enzima crucial das células de Purkinje. O mau funcionamento dessas células levam a uma neurodegeneração progressiva (Duenas *et al.*, 2006).

A idade de início dos sintomas da doença varia de 8 a 55 anos, dependendo do número de repetições CAG no gene *PPP2R2B*. Mas na maior parte dos indivíduos acometidos, os sintomas iniciam-se por volta da quarta década. Esses sintomas são os sintomas clássicos da maioria das ataxias espinocerebelares, que

incluem, tremor na extremidade superior, progredindo ao longo de várias décadas para tremores na cabeça, dificuldade na marcha, dismetria (ausência de noção temporal e espacial), hiper-reflexia, dificuldade de movimento devido aos tremores, movimentos anormais dos olhos, e posteriormente leva a demência (Holmes *et al.*, 1999).

1.3.1.7. Ataxia espinocerebelar tipo 17

A ataxia espinocerebelar tipo 17 (SCA17) é causada por uma expansão do trinucleotídeo CAG no éxon 3 do gene *TBP*, localizado em 6q27, que codifica a proteína de ligação a TATA *box* (é uma sequência de DNA que marca o início da transcrição gênica. Está localizada na região 5' da grande maioria dos genes). Indivíduos que possuem alelos entre 25 e 44 repetições CAG não possuem a doença (Bruni *et al.*, 2004). A SCA17 apenas se manifesta em indivíduos que possuem alelos com 47 ou mais repetições. Uma reduzida penetrância é observada em indivíduos que possuem entre 45 e 46 repetições CAG (Duenas *et al.*, 2006).

Esta é uma doença neurológica autossômica dominante caracterizada por ataxia, disfunção cognitiva, psicoses, convulsões. Seu fenótipo clínico e padrão de herança é similar a Doença de Huntington (Gao *et al.*, 2008).

1.3.1.8. Atrofia Dentatorubropalidoluisana

A Atrofia Dentatorubropalidoluisana (DRPLA) é uma doença progressiva neurodegenerativa causada pela expansão de resíduos de poliglutamina na proteína atrofina 1. O gene associado a esta doença, o *ATN1* que codifica esta proteína localiza-se no éxon 5 do cromossomo 12, em 12p13.31 (Nucifora *et al.*, 2003).

O início dos sintomas geralmente ocorre por volta dos 30 anos de idade, mas podem ocorrer ainda na infância. As manifestações clínicas de DRPLA diferem entre crianças e adultos. Quando os sintomas da doença surgem antes dos 20 anos, envolve movimentos musculares involuntários, mudanças de comportamento, convulsões, problemas de equilíbrio e coordenação (ataxia), e deficiência intelectual. Já quando os sintomas começam a aparecer após os 20 anos, os sinais e sintomas mais frequentes são movimentos incontroláveis dos membros (coreoatetose), ataxia, e demência (Orr *et al.*, 2007).

1.3.1.9. Atrofia Muscular Bulbo Espinhal (Doença de Kennedy)

A Atrofia Muscular Bulbo Espinhal (SBMA) é forma rara de doença do neurônio motor caracterizada por expansão de repetições de trinucleotídeos CAG na região codificadora do gene do receptor androgênico (gene *AR*, também chamado de *HUMARA*) localizado no braço longo do cromossomo X, no *locus* Xq11.12 (Kaimen-Maciel *et al.*, 1998). Alelos normais possuem de 13 a 31 repetições CAG, e alelos expandidos, possuem acima de 40 repetições (Pearson *et al.*, 2005).

Típico de receptores nucleares, o gene *AR* possui codifica para uma proteína com um domínio conservado de ligação com uma sequência específica do DNA e um domínio menos conservado de ativação, na região N-terminal, onde os resíduos de poliglutamina na proteína receptora de andrógenos estão localizados. Após ligação do andrógeno no citoplasma, o complexo receptor-ligante é transportado ao núcleo, onde será ativada a transcrição de genes responsivos hormonais (Orr *et al.*, 2007).

É uma doença de herança recessiva ligada ao X, afetando 1 a cada 50000 homens (Kaimen-Maciel *et al.*, 1998). Homens acometidos podem sofrer com câibras musculares durante anos e apenas na quarta ou quinta década começar a desenvolver fraqueza muscular proximal. Reflexos dos tendões são normalmente ausentes. Ginecomastia é comum e hipogonadismo tardio. Há uma progressiva perda da libido, dificuldades em manter ereção, e esterilidade tardia (Orr *et al.*, 2007).

1.3.1.10. Doença de Huntington

A mutação responsável pela Doença de Huntington (HD) consiste na expansão da repetição CAG localizada no éxon 1 do gene *HTT*, na região 4p16.3 Essa região polimórfica é traduzida na porção N-términal da proteína huntingtina, uma proteína de 348 kDa (Dorsman *et al.*, 1999) que é expressa desde a embriogênese e ao longo da vida do organismo, sendo amplamente expressa no Sistema Nervoso Central e em tecidos periféricos (Bhide *et al.*, 1996). A mutação responsável pela HD consiste em uma expansão instável acima de 38 repetições. Alelos normais possuem entre 6 e 29 repetições CAG (Orr *et al.*, 2007).

A huntingtina possui entre as suas funções, a de regulação da transcrição gênica, transporte intracelular e envolvimento na via endossomo-lisossomo. Por

esses motivos, mutações no gene *HTT* causam problemas no transporte intracelular, transmissão sináptica e na função mitocondrial (Landles *et al.*, 2004).

A HD é uma doença de herança autossômica dominante geralmente se manifesta por volta de 35 a 50 anos, mas casos na infância e na velhice também foram descritos. A coreia é geralmente o primeiro sintoma motor em pacientes adultos. Os movimentos tendem a ser espasmáticos e repentinos e acabam afetando drástica e dramaticamente a atividade diária do acometido. Os déficits de memória, alterações na personalidade e depressão também são comuns. A HD é fatal entre 10 a 15 anos após o início dos sintomas, devido à disfunção bulbar e complicações adicionais (Vonsattel *et al.*, 1998).

1.3.2. Doenças não-poliglutamínicas

As doenças relacionadas às repetições trinucleotídicas são aquelas que não estão envolvidas com a repetição CAG (King, 2011).

As repetições expandidas relacionadas a estes tipos de doenças incluem trinucleotídeos do tipo CGG, GCC, GAA, CTG, sendo que, o tipo e localização da repetição dentro do gene irão definir a fisiologia de cada doença envolvida com essas repetições (Di Prospero *et al.*, 2005). Ao contrário das doenças poliglutamínicas, estas estão associadas principalmente à herança materna da repetição expandida. Além disso, as expansões não estão localizadas na região codificante do seu respectivo gene, causando uma alteração no padrão da expressão gênica. Todas as doenças deste grupo são multissistêmicas e não-progressivas (Cummings *et al.*, 2000).

Destas doenças, as causadas por perda de função gênica (ou haploinsuficiência) ocorrem devido à ausência ou à interferência no processo de transcrição pela presença de sequências expandidas na região (Richards, 2001). Três são as associadas a essa causa: doença de Friedreich (FRDA), Síndrome do X-frágil (FRAXA) e Retardo Mental ligado ao Cromossomo X (FRAXE) (Gatchel *et al.*, 2005). A distrofia miotônica 1 (DM1), por outro lado, ocorre devido à presença da repetição expandida, que causa durante a transcrição do RNA, a compartimentalização inadequada do transcrito e como consequência, a produção de uma proteína com função alterada (Miller, J. W. *et al.*, 2000). E a ataxia espinocerebelar 8 (SCA8) é causada por expansão da repetição CTG na região

3'UTR do transcrito, no entanto é ligada a mecanismos patogênicos não conhecidos (Stevanin *et al.*, 2000).

1.3.2.1. Ataxia de Friedreich

A Ataxia de Friedreich (FRDA) é uma doença neurodegenerativa autossômica recessiva causada pela expansão do trinucleotídeo GAA no éxon 1 do gene frataxina (*FXN*), localizado em 9q21.11 (Ciotti *et al.*, 2004). Indivíduos com 5 a 30 repetições GAA são considerados normais, e aqueles que possuem acima de 70 mutações são indivíduos acometidos. A maior parte dos pacientes possuem dois alelos expandidos, mas alguns foram identificados sendo portadores de apenas um alelo expandido e uma mutação de ponto no segundo alelo (Orr *et al.*, 2007).

Sintomas da Ataxia de Friedreich se iniciam por volta do 10 anos de idade, os pacientes geralmente apresentam uma progressiva marcha de ataxia, geralmente devido a perda de propriocepção. Posteriormente a ataxia, as extremidades se tornam descoordenadas, os tremores se tornam aparentes, e os músculos distais começam a enfraquecer. A maior parte dos pacientes desenvolvem uma cardiomiopatia, e a medida que a doença avança, arritmias e insuficiência cardíaca congestiva se tornam mais frequentes (Orr *et al.*, 2007).

1.3.2.2. Ataxia espinocerebelar tipo 8

A Ataxia espinocerebelar tipo 8 (SCA8) está associada com a expansão do trinucleotídeo CTG na região não transcrita 3' terminal do gene *ATXN8OS* no *locus* 13q21. O produto funcional deste gene é um RNA não codificador. A região genômica de *ATXN8OS* possui 32Kb e contém 6 éxons que podem sofrer *splicing* alternativo (Mutsuddi *et al.*, 2005). Aqueles que possuem alelos com 15 a 34 repetições CTG são considerados indivíduos normais, e os que possuem acima de 89 repetições, acometidos (Todd *et al.*, 2010).

Pacientes acometidos por SCA8 manifestam sintomas clínicos que incluem disartria, dificuldade na respiração e instabilidade na marcha, ataxia nos membros e com isso, a percepção de vibração não vai sendo mais detectada. A idade e início dos sintomas varia entre 18 e 65 anos (Mutsuddi *et al.*, 2005).

1.3.2.3. Distrofia Miotônica 1

Uma expansão instável do trinucleotídeo CTG na região 3' terminal do gene *DMPK* (localizado em 19q13.32) é o erro genético causador da Distrofia Miotônica 1, a doença neuromuscular autossômica dominante mais comum em adultos, com uma incidência de 1 em 8000 entre os caucasianos. Enquanto alelos normais apresentam de 5 a 37 repetições CTG, alelos expandidos possuem acima de 50 repetições, que podem expandir até centenas de repetições (Falk *et al.*, 2006).

DM1 é caracterizada por hipotonia, diplegia facial e retardo mental. Há disfunção no sistema nervoso central. O padrão de fraqueza muscular é distal. Nessa doença há um aumento na espessura da fibra muscular, núcleo central e atrofia muscular (Orr *et al.*, 2007).

1.3.2.4. Síndrome do X-frágil

A síndrome do X-frágil (FRAXA) é responsável por cerca de metade dos casos de retardo mental com herança ligada ao cromossomo X e é a segunda causa mais comum de deficiência mental depois da Síndrome de Down. A FRAXA acomete 1 em cada 4000 homens e 1 em cada 7000 mulheres nascidos vivos (Cirillo *et al.*, 2012). Está associado com a expansão da repetição do trinucleotídeo CGG, localizado na região 5' não traduzida do éxon 1 do gene *FMR1* mapeado no *locus* Xq27.3, que coincide com o sítio frágil sensível a folato (Bagni *et al.*, 2012). O número de repetições CGG neste gene permite identificar o perfil dos indivíduos analisados em: normal, que possui entre 6 e 39 repetições; intermediário, com 40 a 59 repetições; pré-mutado, de 60 a 200 repetições; e com mutação completa, acima de 200 repetições (Bagni *et al.*, 2012).

A chance de uma pré-mutação se tornar uma mutação completa na próxima geração, depende do tamanho da repetição CGG, onde essa possibilidade é perto de 100% para alelos com mais de 99 repetições CGG (Yrigollen *et al.*, 2012). Nesta região pode haver uma interrupção de apenas uma repetição pelo trinucleotídeo AGG entre as repetições de CGG o que contribui para a estabilização do alelo, fazendo com que indivíduos com essa interrupção mas que possuam mutação, possam não apresentar os sintomas da doença (Felix-Lopez *et al.*, 2006; Lyon *et al.*, 2010; Bagni *et al.*, 2012). Além disso, o número de interrupções por AGG é inversamente proporcional ao risco de uma pré-mutação, se tornar uma mutação completa nas próximas gerações (Yrigollen *et al.*, 2012).

A base molecular da síndrome é a expansão e subsequente metilação do trinucleotídeo CGG na região 5' do gene *FMR1* (Thyagarajan *et al.*, 2008). A hipermetilação leva ao silenciamento gênico, e a ausência do produto do gene *FMR1*, uma proteína ligante de RNA (FMRP) envolvida na transcrição do RNA ribossômico em diversos tipos celulares. Em neurônios, FMRP é abundante e participa da remodelação dendrítica e na plasticidade celular, relacionada à capacidade cognitiva (Felix-Lopez *et al.*, 2006).

Os acometidos apresentam um severo retardo mental. Em adição, alguns indivíduos acometidos possuem características físicas como orelhas, face e mandíbula longas. Patologias psiquiátricas como autismo e hiperatividade, também foi relatado em acometidos (Thyagarajan *et al.*, 2008).

1.3.2.5. Retardo Mental ligado ao Cromossomo X

O Retardo Mental ligado ao Cromossomo X (FRAXE) está associado a um sítio frágil localizado em Xq28. Essa doença acomete 1 em cada 50000 meninos nascidos vivos. Esta doença é causada devido ao silenciamento do gene *AFF2* (ou *FMR2*), como consequência de uma expansão do trinucleotídeo CCG localizado na região 5' deste gene. Na população normal o número de repetições CCG varia entre 6 a 35 cópias (Bensaid *et al.*, 2009).

Indivíduos com número de repetições superior a 200, têm suas ilhas CpG hipermetiladas a montante do gene *AFF2*, causando um silenciamento em sua transcrição e perda de função da proteína FMR2 codificada por este gene. A deleção do gene *AFF2* resulta em fenótipos similares àquele causado pela expansão de CGG em *FMR1* (Gu *et al.*, 1996; Gecz, 2000). Esta proteína é expressa nos tecidos periféricos e no cérebro, onde é mais abundante no hipocampo e na amígdala (Miller, W. J. *et al.*, 2000).

Pacientes com FRAXE possuem uma forma mais suave de retardo mental quando comparado a pacientes com FRAXA. Além disso, possuem dificuldade de aprendizado e retardo no desenvolvimento. Alguns pacientes podem ter anormalidades no comportamento como déficit de atenção e hiperatividade, no entanto esses pacientes não manifestam anormalidades físicas (Orr *et al.*, 2007).

1.4. Diagnóstico molecular de doenças causadas por expansões trinucleotídicas

As principais metodologias utilizadas atualmente para a detecção molecular de expansões de *loci* exônicos e não-exônicos de repetições trinucleotídicas associadas às doenças monogênicas, são a análise por *Southern blotting* e a Reação em Cadeia da Polimerase (PCR). O sequenciamento de DNA, considerado o teste padrão ouro, apresenta também uma limitação de resolução de <150 repetições por reação ('*read*'). A análise por *Southern blotting* é altamente robusta em detectar mutações com grandes expansões, no entanto, possui baixa resolução para determinar pré-mutações, além de não permitir a determinação do número exato de repetições do alelo e ser, sobretudo, um método dispendioso e demorado (Mao *et al.*, 2002; Zhou *et al.*, 2004).

O método que utiliza a PCR quantitativa de fluorescência (QF-PCR) tem se mostrado uma ferramenta eficiente tanto pela rapidez na obtenção do diagnóstico quanto pela grande sensibilidade para detecção das doenças causadas por mutações dinâmicas. A QF-PCR tem sido de grande utilidade para a medicina nos exames de pré-natal, exames em recém-nascidos e crianças com suspeita de doenças genéticas. Outro aspecto importante desse método é que ele informa sobre a origem parental da anomalia (Gonzalez-Gonzalez *et al.*, 2005; Abrams *et al.*, 2012).

No Brasil são escassos os estudos populacionais sobre as pré-mutações e mutações de expansões exônicas e de regiões reguladoras 5' associadas com as doenças causadas por expansão de repetições trinucleotídicas. Considerando a questão da miscigenação da população brasileira é importante determinar parâmetros populacionais (frequências e origens parentais) como subsídio a estudos de antropologia genética. O gargalo para esses estudos consiste na dificuldade de se determinar com precisão o número de repetições nos alelos.

O aconselhamento genético sobre o diagnóstico e o padrão de herança das doenças neurológicas causadas por expansões trinucleotídicas é crucial para o compreensivo cuidado dos indivíduos afetados e de suas famílias. Na maioria dos serviços a análise direta de mutações pelo sequenciamento de DNA é indisponível ou economicamente inviável, devido à carência de infraestrutura, somada à natureza heterogênea das mutações patogênicas, tamanho e complexidade das expansões patogênicas envolvidas. O rápido acesso aos testes moleculares específicos é um

elemento de inclusão social, eliminador de desigualdades regionais na área da ciência e tecnologia.

Além do diagnóstico e aconselhamento genético, os benefícios de estudos destas doenças têm sido observados em ambos os campos de neurobiologia e biologia celular. Exemplos incluem o entendimento sobre o papel do controle da tradução na plasticidade sináptica, o papel do RNA no processamento na integridade do funcionamento neuronal e muscular, a importância das enzimas envolvidas no metabolismo energético da célula, e os efeitos expressivos sobre alteração da conformação das proteínas responsáveis pela função e sobrevivência neuronal (Orr *et al.*, 2007).

As doenças neurodegenerativas monogênicas causadas por expansão de repetições trinucleotídicas não são facilmente detectadas e diferenciadas por diagnóstico clínico por apresentarem expressão variada e fenótipos neurológicos semelhantes. Por isso, um diagnóstico acurado depende de testes moleculares que possam detectar a presença de uma mutação no *locus* do gene causador específico da doença (Schelhaas *et al.*, 2000).

Para orientar pacientes que procuram um diagnóstico definitivo para uma destas doenças, é sugerido que haja previamente um aconselhamento genético antes da realização do teste.

Pesquisadores da Federação Mundial de Neurologia (*World Federation of Neurology*) criaram um livro de recomendações e normas, que estabelece todos os princípios éticos para o desenvolvimento do teste baseados no conhecimento e nas técnicas utilizadas em genética molecular, dessa forma, estes só devem ser desenvolvidos em conformidade com todas as condições adequadas, disponibilizando a todos os indivíduos, que desejam fazer o teste, informações relevantes e atualizadas, para então tomar uma decisão de forma esclarecida e voluntária.

O resultado positivo do teste genético impacta profundamente, e de maneira imprevisível, os pacientes e suas famílias. Qualquer profissional que vá aplicar este tipo de teste genético deve procurar obter orientações para os testes a partir do centro de testagem ou de uma organização dedicada aos interesses dos pacientes portadores e suas famílias. Dentre outras normas de procedimento, as diretrizes recomendam que os centros de testes devam adotar as seguintes práticas:

- Deve-se prestar o aconselhamento genético, antes e após o teste, e ainda antes dos resultados serem obtidos;

- Os resultados dos testes devem ser estritamente confidenciais, sendo divulgados apenas à pessoa, e somente para o indivíduo que está sendo testado, independentemente do resultado;

- O participante do teste deve ser encorajado a escolher uma pessoa para acompanhá-lo em todas as etapas do processo de teste: a etapa pré-teste, a realização do teste, a comunicação do resultado e a etapa pós-teste;

- Para proteger os interesses dos menores, nomeadamente a confidencialidade, os testes não devem ser conduzidos para menores de 18 anos sem uma razão médica, como o aparecimento de sintomas de um das doenças na criança (Alonso *et al.*, 2009).

Devido à carência de serviços especializados tanto no diagnóstico quanto no acompanhamento clínico de pacientes acometidos pelas doenças monogênicas causadas por expansão de repetições trinucleotídicas nas regiões Norte e Noroeste Fluminense, o presente estudo teve como meta o desenvolvimento e aplicação de ensaios moleculares para o rastreio de polimorfismos nos genes associados à essas expansões. Os dados gerados neste estudo contribuirão para um maior conhecimento da incidência e prognóstico da doença em pacientes brasileiros.

2. OBJETIVOS

2.1. Objetivo Geral

Desenvolver testes moleculares para a investigação pré- e pós-sintomática de doenças monogênicas causadas por expansão de repetições trinucleotídicas instáveis.

2.2. Objetivos Específicos

- Análise de expansões trinucleotídicas associadas a doenças neurológicas monogênicas por meio de varredura *in silico* “*genome-wide*” para identificação de *loci* de microssatélites trinucleotídicos e sua relação com doenças monogênicas causadas pela expansão desses elementos.
- Desenvolver iniciadores específicos para as quinze doenças monogênicas associadas à expansão de repetições trinucleotídicas, visando identificar os perfis alélicos por ensaio QF-PCR para o rastreamento de polimorfismos localizados nos respectivos genes.
- Desenvolver *in silico* uma versão tetraprimer da técnica da PCR quantitativa por fluorescência com iniciador de repetições trinucleotídicas (“*triplet repeat primed polymerase chain reaction*” - TP-PCR) como ferramenta para a investigação pré- e pós-sintomática de doenças neurodegenerativas monogênicas causadas por instabilidade das repetições.
- Determinar parâmetros populacionais relevantes em amostra da Região Norte Fluminense.
- Desenvolver testes moleculares para a investigação pré- e pós-sintomática em núcleos familiares com acometidos e em indivíduos sintomáticos para subsidiar o aconselhamento genético robusto para estas doenças.

3. METODOLOGIA

3.1. Material biológico

Sangue venoso: amostras de 0,5-2mL de sangue venoso foram coletadas em tubo contendo anticoagulante EDTA e mantidas sob refrigeração até o momento da extração de DNA dos leucócitos.

Células do epitélio bucal: Células bucais foram coletadas por esfregaço bucal utilizando haste flexível, estéril, com ponta de algodão hidrofílico. Este tipo de material foi coletado preferencialmente em crianças menores de cinco anos.

3.2. Sujeitos da pesquisa

Para estimar as frequências alélicas de cada um dos marcadores foram genotipadas amostras de DNA de indivíduos adultos da população geral, não acometidos, não relacionados geneticamente, pertencentes ao banco de DNA do Núcleo de Diagnóstico e Investigação Molecular - NUDIM, gerado a partir de testes de parentesco.

3.3. Aspectos éticos da pesquisa

Este projeto faz parte da grande linha de pesquisa intitulada “*Desenvolvimento de testes moleculares rápidos, altamente eficientes e econômicos para o diagnóstico de doenças genéticas humanas: Fortalecendo a interiorização da Genética no SUS*”, cujo projeto geral foi aprovado pela Comissão Regional de Ética em Seres Humanos da Faculdade de Medicina de Campos, sob folha de rosto **FR278769**. A grande linha de pesquisa é financiada pela FAPERJ e pelo CNPq. Todas as amostras foram coletadas sob assinatura de “Termo de Consentimento Livre e Esclarecido pelos participantes e/ou responsáveis legais” (Anexo).

3.4. Extração de DNA

O material biológico (sangue e de células do epitélio bucal) foi submetido à extração de DNA através de um kit de extração comercial *GE Healthcare®*. Segundo o manual do kit *Ilustra™ Blood genomicPrep Mini Spin (GE Healthcare)*. Esse material foi coletado de indivíduos voluntários.

3.5. Quantificação do DNA genômico

O DNA extraído foi submetido à quantificação direta utilizando o aparelho de espectrofotometria *NanoDrop 2000c™* - *Thermo Scientific*.

3.6. Métodos

3.6.1. Análise exômica e de regiões reguladoras 5' para busca de repetições trinucleotídicas

Foi realizada uma varredura no genoma humano de referência GRCh37/hg19 de Fevereiro 2009, utilizando o programa *Tandem Repeats Database (TRDB)* (Gelfand *et al.*, 2007) para mineração de *locus* trinucleotídicos que estejam associados a doenças causadas por expansão dessas repetições. Este programa utiliza um algoritmo capaz de encontrar os elementos repetitivos pelo percentual de identidade e pela frequência de *indels* (inserções e/ou deleções) comparado a uma sequência referência com repetições perfeitas em série. Quanto maior for o número de repetições e o tipo de repetição, maior será a pontuação final. Por exemplo: 10 repetições perfeitas de tetranucleotídeos apresentam pontuação igual a 80 enquanto 10 de dinucleotídeos apresentam pontuação 40. Vinte repetições dinucleotídicas fazem 80 pontos. As sequências imperfeitas possuem pontuação menor do que as perfeitas com o mesmo número de repetições, exemplo: a sequência ACACACATACACACAC possui pontuação menor do que a sequência ACACACACACACAC. Os descontos na pontuação final dependem da proporção dos erros de pareamento com uma sequência perfeita.

Para seleção exômica e de regiões reguladoras 5' desses *locus* trinucleotídicos foram avaliados aqueles *locus* trinucleotídicos com pontuação mínima de 50 e identidade de 90%.

3.6.2. Desenho dos iniciadores

Foram desenhados iniciadores para duas técnicas diferentes, que serão descritas abaixo:

Para a técnica da Reação em Cadeia da Polimerase Quantitativa por Fluorescência (QF-PCR) foi utilizado o programa *Websat – Find microsatellite and design primers* (Martins *et al.*, 2009) para desenho dos iniciadores, e para verificação de formação de dímeros ou auto-complementaridade dos iniciadores, o programa *OligoCalc - Oligonucleotide Properties Calculator* (Kibbe, 2007). A

sequência de cada *locus* foi analisada utilizando os programas BLAST - *Basic local alignment search tool* (Altschul *et al.*, 1990) e o PCR *in silico* (*iPCR*) do browser da UCSC (Kent *et al.*, 2002) para verificação de homologia com outras regiões do genoma humano, visando estabelecer se os iniciadores eram únicos.

A técnica de PCR com iniciador de repetições trinucleotídicas (TP-PCR), se baseia em uma PCR *triprimer* (que utiliza três iniciadores), sendo apenas um deles, o iniciador 1 marcado com fluorescência, é específico para a região estudada e flanqueia a região 5', e os outros dois iniciadores 2 e 3, são os reversos, e possuem a sequência da região 5' comum (Warner *et al.*, 1996).

A sequência do iniciador 1 foi desenhada utilizando o programa *Websat*, como desenhado para a técnica de QF-PCR.

A sequência 5' comum entre os iniciadores 2 e 3 foi selecionada de um produto artificial com 21pb produzido de maneira randômica *in silico*, e os critérios de seleção foram: (1) não possuir complementaridade entre si; (2) não ter complementaridade com a repetição analisada; (3) não possuir homologia com nenhuma outra sequência do genoma conhecida (Warner *et al.*, 1996; Cagnoli *et al.*, 2004; Ciotti *et al.*, 2004; Cagnoli *et al.*, 2006; Falk *et al.*, 2006; Winnepeninckx *et al.*, 2007; Catalli *et al.*, 2010; Lyon *et al.*, 2010).

A sequência do iniciador 2 é também comum ao iniciador 3, e a sequência do iniciador 3 é composta de uma cauda na região 5' que é sequência comum ao iniciador 2, e na região 3' uma sequência de cinco repetições trinucleotídicas repetidas em tandem do *locus* que está sendo analisado (Warner *et al.*, 1996; Ciotti *et al.*, 2004).

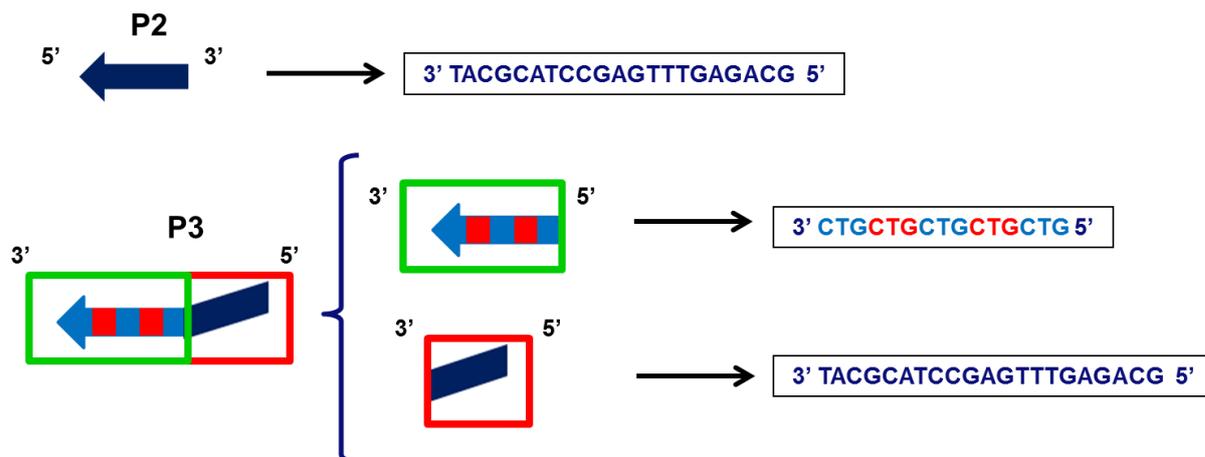


Figura 7. Esquema das seqüências dos iniciadores 2 (P2) e 3 (P3) respectivamente. A seqüência 3' do P3 corresponde a mesma seqüência do P2. E a região 5' de P3 é a seqüência complementar à seqüência alvo do *locus* que está sendo analisado. Neste exemplo o P3 foi desenhado para amplificar uma repetição de trinucleotídeos CAG.

3.6.3. Padronização da QF-PCR convencional

Foram avaliados para os iniciadores desenhados para cada doença separadamente, os efeitos das seguintes variáveis: concentração do DNA alvo, concentração dos iniciadores, e cada passo dos ciclos térmicos utilizados.

Para padronização foram realizados ensaios de gradiente com as seguintes variáveis: (1) temperatura de anelamento, calculada a partir da temperatura ótima de anelamento realizada pelo programa *OligoCalc - Oligonucleotide Properties Calculator*; (2) concentração de iniciadores; (3) concentração de $MgCl_2$; e (4) concentração de DNA. O tampão de reação utilizado teve a seguinte formulação: 0,2mM de dNTP, 1-2mM de $MgCl_2$, 1,25U da enzima *AmpliTaq Gold® DNA Polymerase* (Alves da Silva, 2007).

Na reação de amplificação foi adicionado betaína na concentração final de 2M na reação. Esse agente facilita a separação das fitas de DNA além de aumentar o rendimento e a especificidade dos produtos da PCR por contribuir na amplificação de alvos ricos em CG por aumentar a estabilidade das fitas (Frackman *et al.*, 1998).

3.6.4. PCR com iniciador de repetições trinucleotídicas (TP-PCR)

Regiões com repetições trinucleotídicas podem ser identificadas pela reação em cadeia da polimerase quantitativa de fluorescência quando o comprimento da região repetitiva não é muito longo. Porém, em indivíduos com um grande número de repetições, a PCR pode não ser eficiente na amplificação do alelo longo. Uma

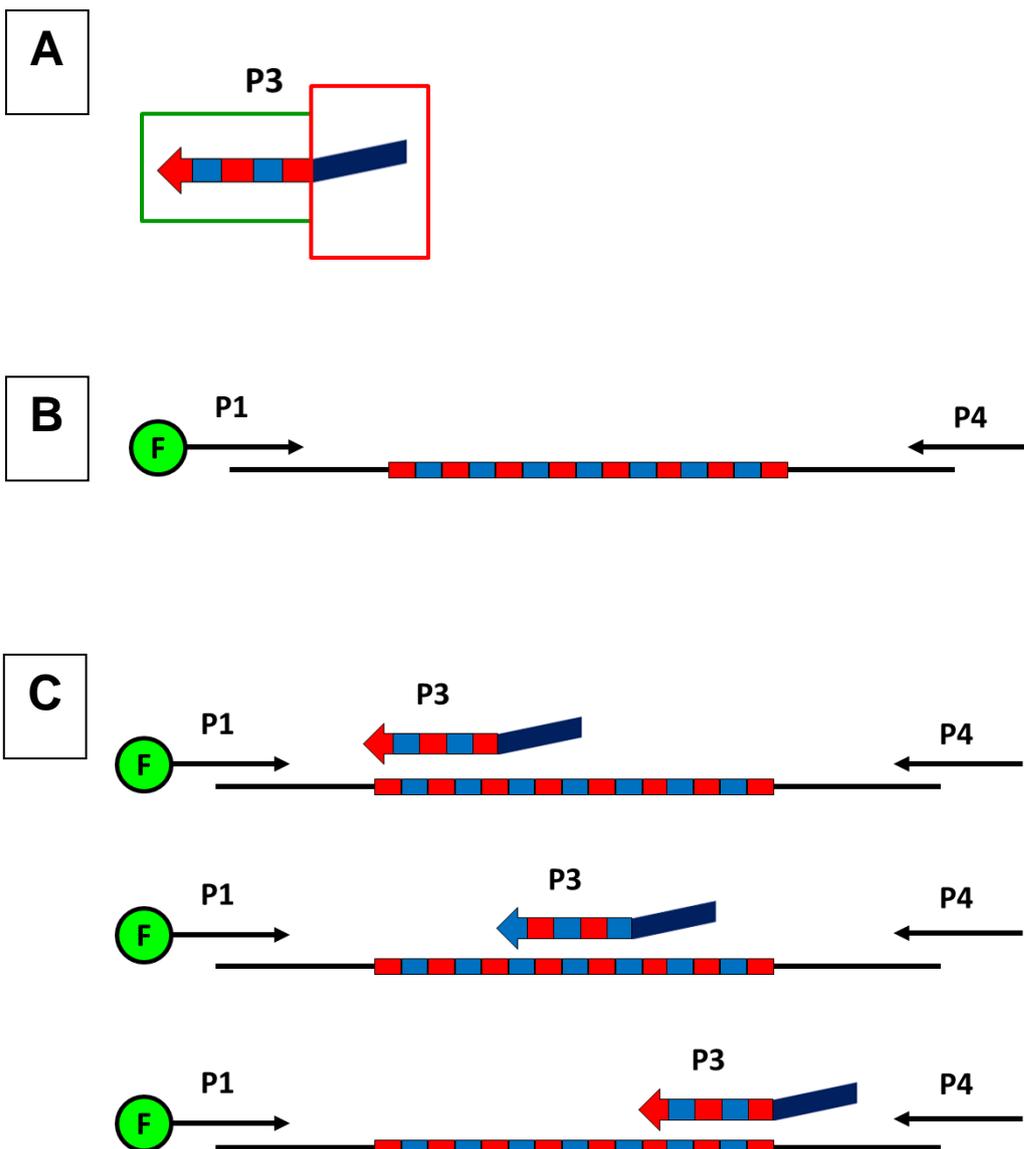
variação da técnica de PCR, denominada *Triplet Repeat Primed PCR* (TP-PCR), foi desenvolvida para aplicação em casos de ocorrência de alelos muito longos (Warner *et al.*, 1996; Ciotti *et al.*, 2004; Falk *et al.*, 2006; Chen *et al.*, 2010; Kakourou *et al.*, 2010; Lyon *et al.*, 2010; Xuncla *et al.*, 2010; Radvansky, Ficek e Kadasi, 2011; Radvansky, Ficek, Minarik *et al.*, 2011; Jama *et al.*, 2013).

O P1 e P4 são os iniciadores gene-específicos que flanqueiam a região de interesse, e apenas um desses iniciadores é marcado com fluorescência. Nesta técnica podem ser utilizados um ou ambos os iniciadores gene-específicos. Além dos iniciadores gene-específicos, dois iniciadores repetição-específicos. O primeiro iniciador é o menor, e corresponde a sequência comum entre ambos iniciadores que é a seguinte: 5' TACGCATCCGAGTTTGAGACG 3' (Warner *et al.*, 1996). Esta é a sequência mais utilizada na literatura para o iniciador 2. O segundo iniciador repetição-específico, o maior dos dois, é desenhado com cinco repetições da sequência complementar à repetição do *locus* que está sendo investigado somado à sequência comum entre eles mostrada anteriormente. O tamanho do *amplicon* nesta técnica varia com o número de repetições trinucleotídicas que a região polimórfica do *locus* específico possui.

Nos ciclos iniciais da PCR, P1 e P4 amplificam a região específica. Nos ciclos posteriores as repetições específicas da região 3' do P3 se ligam nos múltiplos sítios do alelo que contém a repetição do *locus* alvo. Uma concentração 10 vezes menos do P3 em relação ao P2 irá garantir sua exaustão ainda nos primeiros ciclos de amplificação. O P2 irá amplificar a partir da região final dos produtos de amplificação prévia dos ciclos iniciais da PCR, gerados pelo P3. P2 e P3 são iniciadores reversos ao iniciador com fluorescência, e são iniciadores repetição-específicos. No esquema da metodologia mostrado na figura 8 é representada a técnica na versão *tetraprimer*, ou seja, com quatro iniciadores. Para esta versão, o molde utilizado será o produto de amplificação diluído da QF-PCR.

Inicialmente foi utilizado o protocolo de amplificação descrito em (Warner *et al.*, 1996; Ciotti *et al.*, 2004), que utiliza para um volume final 25µl, 200ng de DNA genômico, 0,8M do iniciador 1, 0,8M do iniciador 2 e 0,08M do iniciador 3. O tampão utilizado foi o descrito em Alves Da Silva, 2007 (Alves da Silva, 2007): 0,2 mM dNTP, 1-2 mM MgCl₂, 1,25U da enzima *AmpliTaq Gold® DNA Polymerase*. Inicialmente a reação foi submetida a 30 ciclos, sendo que cada ciclo consiste em: 30 segundos a 94°C para desnaturação, 30 segundos a 60°C para anelamento e 30 segundos a

72°C para extensão, seguido de 10 minutos de extensão final a 72°C. Os produtos da PCR foram então mantidos a 95°C por 2 minutos. Após amplificação, esses produtos da TP-PCR foram também submetidos a eletroforese capilar (Warner *et al.*, 1996; Ciotti *et al.*, 2004). Foi adicionado à reação, outro iniciador gene-específico (P4), para aumentar a especificidade (Chen *et al.*, 2010). Os iniciadores gene-específicos utilizados para essa técnica foram os mesmos utilizados para a técnica de QF-PCR.



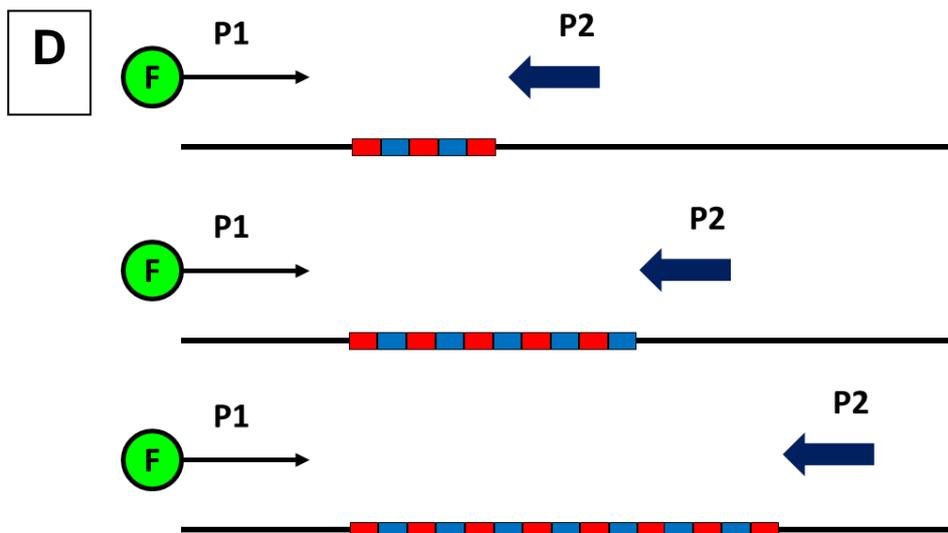


Figura 8. Detalhes do funcionamento da técnica de triplet repeat primed PCR (TP-PCR). (A) Representação do iniciador 3 (P3). A caixa em vermelho mostra a cauda da região 5' que possui sequência idêntica ao iniciador 2 (P2) e não possui complementaridade com nenhuma outra região do genoma alvo, e a caixa em verde mostra a sequência da região 3' de P3 que possui uma sequência de cinco repetições trinucleotídicas em tandem, complementar à repetição do locus alvo de análise. Portanto, P3 pode se anelar em diferentes sítios da repetição. (B) O P1 e P4 são os iniciadores genespecíficos que flanqueiam a região de interesse. A especificidade da reação é devido a esses iniciadores. (C) Nos ciclos iniciais as repetições específicas da região 3' do P3 irão se ligar nos múltiplos sítios do alelo que contém a repetição dando origem a uma variedade de produtos. A especificidade da reação é devido ao iniciador 1 (P1). Uma concentração 10 vezes menos do P3 em relação ao P2 irá garantir sua exaustão ainda nos primeiros ciclos de amplificação. (D) O P2 irá amplificar a partir da região final dos produtos de amplificação prévia dos ciclos iniciais da PCR, gerados pelo P3. Adaptado de (Warner *et al.*, 1996; Ciotti *et al.*, 2004).

3.6.5. Eletroforese capilar

Aos produtos de amplificação foi adicionado formamida (*Hi-Di* Formamida, *Applied Biosystems*) e o padrão de peso molecular *GeneScan LIZ 500*, marcado com o fluoróforo LIZTM (fluorescência laranja) (*Applied Biosystems*). Os produtos de amplificação foram separados por eletroforese capilar com o polímero *POP 4* (contendo 4% de acrilamida), utilizando a plataforma *ABI PrismTM 310 Genetic Analyzer*. Os perfis eletroforéticos foram analisados utilizando os programas *GeneScanTM* e *GenotyperTM*. Os alelos foram nomeados com o valor em pares de nucleotídeos, dos produtos da PCR. Na figura 9, está esquematizado o método padrão de interpretação de um eletroferograma.

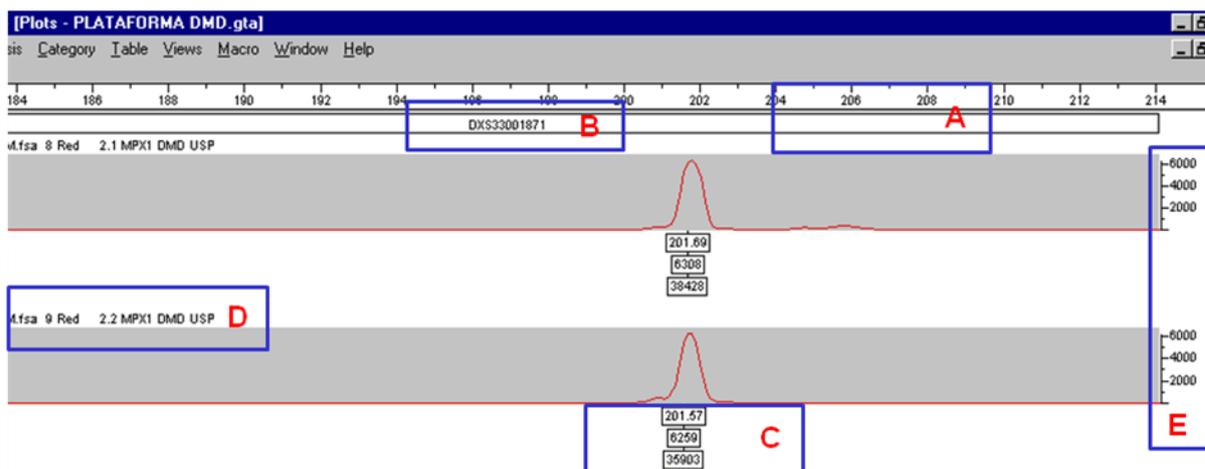


Figura 9. Informações para a interpretação de um eletroferograma. No eixo X, representado em A, escala em número de pares de nucleotídeos (nt). Em B, identificação do marcador utilizado. Em C, informações alélicas representadas, caracterização dos alelos em pares de nucleotídeos, dosagem da proporção alélica em área e intensidade de fluorescência do alelo. Em D, identificação da amostra analisada. No eixo Y, representado em E, escala da atividade de fluorescência em unidades arbitrárias. A intensidade de fluorescência está diretamente relacionada com a quantidade de DNA amplificado. Figura reproduzida com permissão do autor (Alves da Silva, 2007).

3.6.6. Identificação e caracterização dos perfis alélicos

Para a interpretação dos dados foi identificado em todo alelo amplificado apenas o número de pares de bases (pb) que correspondessem às sequências polimórficas analisadas para ser determinados o número de repetições trinucleotídicas de cada *locus* analisado.

Para calcular o número de repetições trinucleotídicas basta subtrair da sequência total amplificada pelos iniciadores gene-específicos o total de nucleotídeos da região não polimórfica (fragmento mínimo a ser amplificado) e dividir o resultado por três, por serem *locus* com polimorfismo de três nucleotídeos. A figura 10 apresenta de forma esquematizada o método genérico adotado como padrão para estimar o número de cópias das repetições específicas do *locus* do gene analisado.

Os resultados que apresentaram alelos com número de repetições superior ao limiar de normalidade foram considerados alelos com pré-mutação ou alelos com mutação, dependendo do número de repetições encontrado e do *locus* que foi analisado.

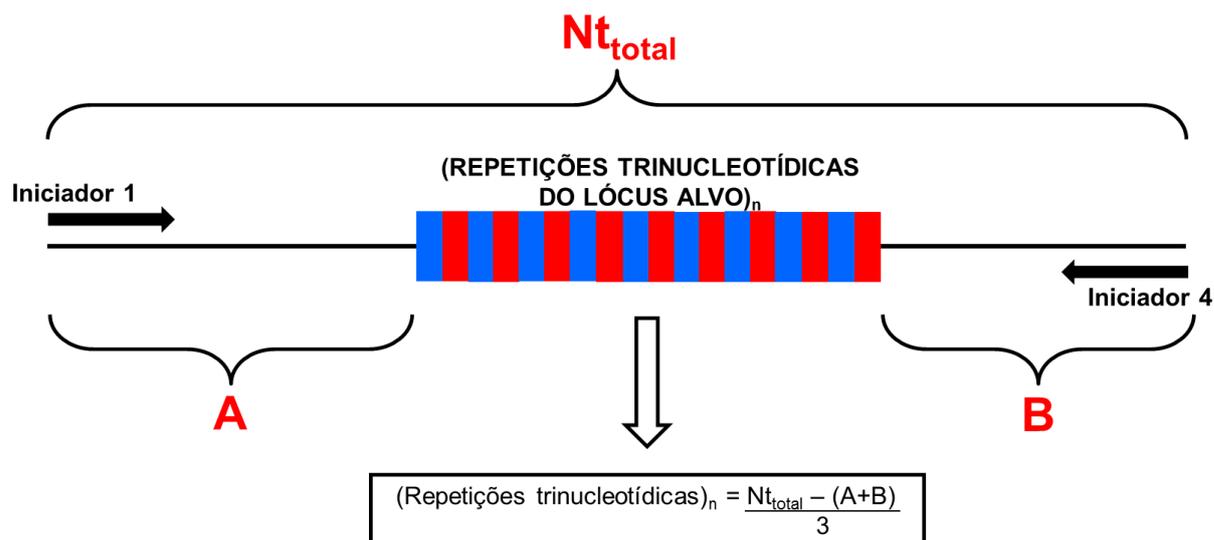


Figura 10. Esquema genérico para estimar o número de repetições trinucleotídicas no gene de interesse. Em azul e vermelho está destacada a sequência polimórfica do gene [Repetições trinucleotídicas]_n. O número de repetições trinucleotídicas do gene é calculado pela determinação tamanho do *amplicon* gerado pela amplificação dos iniciadores 1 e 4 gene-específicos subtraindo o tamanho da região não polimórfica que foi amplificada (região A e B), e dividindo o resultado por três, por serem repetições trinucleotídicas. Nt: nucleotídeos.

3.6.7. Frequências alélicas e dados populacionais

Foram genotipadas amostras de DNA de indivíduos não aparentados geneticamente. Os dados populacionais foram determinados utilizando o software *online PowerStatsV12* distribuído pela *Promega Corp (Promega, Madison, WI)*.

4. RESULTADOS

4.1. Varredura exômica e de regiões reguladoras 5' de TNRs

Para embasamento do estudo foi realizada uma varredura do genoma humano de referência GRCh37/hg19, de fevereiro de 2009. Foram utilizados como critérios, regiões localizadas em éxons e em regiões reguladoras 5' que estivessem associadas a alguma doença, e que apresentassem identidade acima de 90%.

Nesta varredura foram encontradas 74 *loci* de repetições trinucleotídicas envolvidas com doenças. Para todas as regiões encontradas na varredura foi executada uma busca de informações sobre sua localização física em cada cromossomo, o número de cópias da repetição, pontuação, sequência consenso, nome do gene relacionado à região, proteína codificada pelo determinado gene, a doença envolvida com a proteína relacionada, os *links* do OMIM (*Online Mendelian Inheritance in Man*) sobre informações da proteína e da doença respectivamente e a localização citogenética para todos os *locus* de repetições trinucleotídicas encontrado. Os dados relativos aos resultados desta varredura foram compilados e são mostrados na tabela 1.

Após mineração dos dados foi observado que apenas quinze *loci* estão diretamente associados com doenças monogênicas causadas por mutações dinâmicas, todas neurodegenerativas. Na tabela 1, em amarelo são destacadas as quinze doenças monogênicas causadas por expansão de repetições trinucleotídicas.

Após uma revisão sistemática da literatura foram reunidos para estas doenças, os dados mais relevantes sobre cada uma. As informações coletadas encontram-se descritas nas tabelas 2 e 3 para as dez doenças poliglutamínicas e para as cinco não-poliglutamínicas, respectivamente.

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas.

Localização física (a)	Número de cópias (b)	Identidade (%) (c)	Pontuação (d)	Sequência consenso (e)	Nome do gene (Link NCBI) (£)	Proteína (g)	Link OMIM (Proteína) (h)	Doença envolvida (i)	Link OMIM (Doença) (j)	Locus no Mapa Genético (k)
CROMOSSOMO 1										
6529183--6529235	17,666666	94	79	TCC	PLEKHG5	PLECKSTRIN HOMOLOGY DOMAIN-CONTAINING PROTEIN, FAMILY G, MEMBER 5	*611101	SPINAL MUSCULAR ATROPHY, DISTAL, AUTOSOMAL RECESSIVE, 4; DSMA4	#611067	1p36
53793512--53793546	11,666667	100	70	GCA	LRP8	LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 8	*602600	MYOCARDIAL INFARCTION, SUSCEPTIBILITY TO, 1	#608446	1p34, 1q25, 6q25.1, 6p12
53793679--53793717	13,000000	92	51	GCC	DISC2	DISRUPTED IN SCHIZOPHRENIA 2	*606271	SCHIZOPHRENIA 9; SCZD9	#604906	1q42.1
231950684--231950726	14,333333	100	86	AAT	DISC1	DISRUPTED IN SCHIZOPHRENIA 1	*605210	SCHIZOPHRENIA	#181500	1p36.2, 15q15, 14q32.3, 13q34, 13q32, 13q14-q21, 12q24, 11q14-q21, 11p14-p13, 10q22.3, 8p21, 6q13-q26, 1p36.3, 6p22.3, 6p23, 5q23-q35, 3q13.3, 3p25, 1q42.1, 1q42.1, 22q12.3, 22q12.3, 22q11.2, 22q11, 1q32.1, 11p36.2, 15q15, 14q32.3, 13q34, 13q32, 13q14-q21, 12q24, 11q14-q21, 11p14-p13, 10q22.3, 8p21, 6q13-q26, 1p36.3, 6p22.3, 6p23, 5q23-q35, 3q13.3, 3p25, 1q42.1, 1q42.1, 22q12.3, 22q12.3, 22q11.2, 22q11, 1q32.1, 1q42.1
					TSNAX-DISC1	TSNAX-DISC1 readthrough				1q42.1
CROMOSSOMO 2										
50574009--50574039	10,333333	100	62	CGC	NRXN1	NEUREXIN 1	+600565	PITT-HOPKINS-LIKE SYNDROME 2	+600565	2p16.3
51259278--51259328	17,333334	92	68	GGA	BCL11A		*606557	B-CELL CLL/LYMPHOMA 11A	*606557	2p16.1
60780422--60780450	9,666667	100	58	GGC	FBXO41	F-BOX ONLY PROTEIN 41	*609108	ALSTROM SYNDROME	#203800	2p13.2
73496580--73496609	10,000000	96	51	GCG	EIF2AK3	EUKARYOTIC TRANSLATION INITIATION FACTOR 2-ALPHA KINASE 3	*604032	EPIPHYSEAL DYSPLASIA, MULTIPLE, WITH EARLY-ONSET DIABETES MELLITUS	#226980	2p12

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas (Continuação).

148779021--148779093	24,333334	94	110	TGC	MBD5	METHYL-CpG-BINDING DOMAIN PROTEIN 5; MBD5	*611472	MENTAL RETARDATION, AUTOSOMAL DOMINANT 1	#156200	2q23.1
175202175--175202222	16,000000	95	78	GGC	SP9	Sp9 transcription factor homolog (mouse)		BRACHYDACTYLY, TYPE D; BDD	#113200	2q31.1
176957786--176957827	14,000000	95	66	GGC	HOXD13	HOMEODOMAIN PROTEIN 13; HOXD13	*142989	BRACHYDACTYLY, TYPE E1; BDE1 SYNPOLYDACTYLY 1; SPD1 SYNDACTYLY, TYPE V VATER ASSOCIATION BRACHYDACTYLY-SYNDACTYLY SYNDROME	#113300 #186000 #186300 #192350 #610713	2q31-q32 2q31-q32 2q31-q32 2q31-q32 2q31-q32
203241235--203241270	12,000000	100	72	GGC	BMP2	BONE MORPHOGENETIC PROTEIN RECEPTOR, TYPE II; BMP2	*600799	PULMONARY VENOOCCLUSIVE DISEASE; PVOD PULMONARY HYPERTENSION, PRIMARY, 1; PPH1	#265450 #178600	2q33 13q12-q14, 2q33
230456533--230456558	8,666667	100	52	GCT	DNER	DELTA- AND NOTCH-LIKE EPIDERMAL GROWTH FACTOR-RELATED RECEPTOR	*607299	MULTIPLE SCLEROSIS, SUSCEPTIBILITY TO; MS	#126200	6q21, 6p21.3, 6p21.3, 2q37.3
242792876--242792906	10,333333	100	62	GCA	PDCD1	PROGRAMMED CELL DEATH 1; PDCD1	*600244	SYSTEMIC LUPUS ERYTHEMATOSUS; SLE SYSTEMIC LUPUS ERYTHEMATOSUS, SUSCEPTIBILITY TO, 2; SLEB2	#152700 #605218	1q23, 16p13.3, 1q23, 1q22, 11q14, 1p13, 6p21.3, 4q22-q24, 4p16-p15.2, 3p21.3-p21.2 2q37.3
CROMOSSOMO 3										
46751074--46751100	9,000000	100	54	AAG	TMIE			DEAFNESS, AUTOSOMAL RECESSIVE 6; DFNB6	#600971	3p21
63898361--63898392	10,666667	100	64	GCA	ATXN7	ATAXIN 7	*607640	SPINOCEREBELLAR ATAXIA 7	#164500	3p21.1-p12
CROMOSSOMO 4										
3076604--3076667	21,333334	98	119	CAG	HTT	HUNTINGTIN; HTT	*613004	HUNTINGTON DISEASE; HD	#143100	4p16.3
3076666--3076695	10,000000	96	51	GCC	HTT	HUNTINGTIN; HTT	*613004	HUNTINGTON DISEASE; HD	#143100	4p16.3
41747993--41748039	15,666667	91	58	GCC	PHOX2B	PAIRED-LIKE HOMEODOMAIN 2B	*603851	CENTRAL HYPOVENTILATION SYNDROME, CONGENITAL; CCHS HIRSCHSPRUNG DISEASE, SUSCEPTIBILITY TO, 1; HSCR1 NEUROBLASTOMA, SUSCEPTIBILITY TO, 2; NBLST2	#209880 #142623 #613013	12q22-q23, 11p13, 10q11.2, 5p13.1-p12, 4p12, 20q13.2-q13.3 10q11.2, 4p12, Xq28 4p12
46392008--46392045	12,666667	94	58	GGA	GABRA2	GAMMA-AMINOBUTYRIC ACID RECEPTOR, ALPHA-2	*137140	ALCOHOL DEPENDENCE	#103780	13q14-q21, 13q14, 7q31, 4q22, 4q22, 4p13-p12

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas (Continuação).

CROMOSSOMO 5											
									PHEOCHROMOCYTOMA	#171300	5p13.1-p12 12q22-q23, 11p13, 10q11.2, 5p13.1-p12, 4p12, 20q13.2- q13.3
37815590--37815650	20,333334	93	86	TCC	GDNF	GLIAL CELL LINE- DERIVED NEUROTROPHIC FACTOR	*600837	CENTRAL HYPOVENTILATION SYNDROME, CONGENITAL; CCHS	#209880		
88179120--88179177	19,000000	93	80	GAG		MADS BOX TRANSCRIPTION ENHANCER FACTOR 2, POLYPEPTIDE C	*600662	HIRSCHSPRUNG DISEASE, SUSCEPTIBILITY TO, 3; HSCR3	#613711	5p13.1-p12	
107716707--107716749	14,333333	93	59	CCG	MEF2C			MENTAL RETARDATION, STEREOTYPIC MOVEMENTS, EPILEPSY, AND/OR CEREBRAL MALFORMATIONS	#613443	5q14	
131593351--131593380	10,000000	96	51	GGC	PDLIM4	PDZ AND LIM DOMAIN PROTEIN 4	*603422	OSTEOPOROSIS	#166710	12q12-q14, 11q13.4, 11p12, 7q22.1, 7q21.3, 5q31.1, 20p12.3, 17q21.31-q22	
140953557--140953597	13,666667	97	73	AGG	DIAPH1	DIAPHANOUS, DROSOPHILA, HOMOLOG OF, 1	*602121	DEAFNESS, AUTOSOMAL DOMINANT 1	#124900	5q31	
146258291--146258322	10,666667	100	64	GCT	PPP2R2B	PROTEIN PHOSPHATASE 2, REGULATORY SUBUNIT B, BETA	*604325	SPINOCEREBELLAR ATAXIA 12; SCA12	#604326	5q31-q33	
178772260--178772285	8,666667	100	52	GCA	ADAMTS2	A DISINTEGRIN-LIKE AND METALLOPROTEINASE WITH THROMBOSPONDIN TYPE 1 MOTIF, 2	*604539	EHLERS-DANLOS SYNDROME, TYPE VII, AUTOSOMAL RECESSIVE	#225410	5q23	
CROMOSSOMO 6											
16327865--16327955	30,333334	97	164	TGC	ATXN1	ATAXIN 1	*601556	SPINOCEREBELLAR ATAXIA 1; SCA1	#164400	6p23	
16761716--16761775	20,400000	90	72	CTC							
24173218--24173249	10,666667	100	64	AAT	DCDC2	DOUBLECORTIN DOMAIN-CONTAINING PROTEIN 2	*605755	DYSLEXIA, SUSCEPTIBILITY TO, 2; DYX2	#600202	6p22.2	
35478743--35478788	15,333333	91	56	CCT	TULP1	TUBBY-LIKE PROTEIN 1	*602280	RETINITIS PIGMENTOSA 14; RP14	#600132	6p21.3	
38139657--38139686	10,000000	100	60	AAT	BTBD9	BTB/POZ DOMAIN- CONTAINING PROTEIN 9	*611237	RESTLESS LEGS SYNDROME, SUSCEPTIBILITY TO, 6	#611185	6p21	
45390419--45390484	22,000000	95	105	CAG		RUNT-RELATED TRANSCRIPTION FACTOR 2	*600211	CLEIDOCRANIAL DYSPLASIA; CCD	#119600	6p21	
45390487--45390531	15,000000	95	72	GGC	RUNX2						
102517171--102517214	14,666667	100	88	AAT	GRIK2	GLUTAMATE RECEPTOR, IONOTROPIC, KAINATE 2	*138244	MENTAL RETARDATION, AUTOSOMAL RECESSIVE 6; MRT6	#611092	6q21	
170870995--170871105	37,000000	95	177	GCA	TBP	TATA BOX-BINDING PROTEIN	*600075	SPINOCEREBELLAR ATAXIA 17; SCA17 (HUNTINGTON DISEASE- LIKE 4; HDL4) PARKINSON DISEASE, LATE- ONSET; PD	#607136 #168600	6q27 17q21.1, 6q27, 4q22	
CROMOSSOMO 7											
39379288--39379318	10,333333	100	62	CAG	POU6F2	POU DOMAIN, CLASS 6, TRANSCRIPTION FACTOR 2	*609062	WILMS TUMOR 5; WT5	#601583	7p14-p13	
156802644--156802678	11,666667	97	61	GCG	MNX1	MOTOR NEURON AND PANCREAS HOMEBOX 1	*142994	CURRARINO SYNDROME	#176450	7q36	

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas (Continuação).

CROMOSSOMO 8										
61591324--61591371	16,000000	97	87	GCG	CHD7	CHROMODOMAIN HELICASE DNA- BINDING PROTEIN 7	*608892	KALLMANN SYNDROME 5; KAL5	#612370	8q12.1
								SCOLIOSIS, IDIOPATHIC, SUSCEPTIBILITY TO, 3; IS3	#608765	8q12.1
								CHARGE SYNDROME (CHARGE ASSOCIATION--COLOBOMA, HEART ANOMALY, CHOANAL ATRESIA, RETARDATION, GENITAL AND EAR ANOMALIES HALL-HITTNER SYNDROME; HHS)	#214800	8q12.1, 7q21.11
								HYPOGONADOTROPIC HYPOGONADISM	#146110	12q13-q21, 9q34.3, 8q12.1, 8p11.2-p11.1, 4q25, 3p21.1, 19p13.3
CROMOSSOMO 9										
79999548--79999580	11,000000	100	66	TGA	VPS13A	VACUOLAR PROTEIN SORTING 13, YEAST, HOMOLOG OF, A	*605978	CHOREOACANTHOCYTOSIS; CHAC	#200150	9q21
95237025--95237068	14,666667	100	88	TCA	ASPN	ASPORIN	*608135	OSTEOARTHRITIS SUSCEPTIBILITY 3; OS3	#607850	9q21.3-q22
								INTERVERTEBRAL DISC DISEASE; IDD	#603932	15q22, 9q21.3- q22, 1p21, 6q27, 1p33-p32.2, 20q13.3
100616695--100616733	13,000000	97	69	GCC	FOXE1	FORKHEAD BOX E1	*602617	HYPOTHYROIDISM, ATHYROIDAL, WITH SPIKY HAIR AND CLEFT PALATE (BAMFORTH-LAZARUS SYNDROME)	#241850	9q22
102590605--102590650	15,333333	93	65	ACC	NR4A3	NUCLEAR RECEPTOR SUBFAMILY 4, GROUP A, MEMBER 3	*600542	CHONDROSARCOMA, EXTRASKELETAL MYXOID	#612237	17q11.1-q11.2, 9q22, 3q11-q12
140087005--140087052	16,000000	93	69	CCT	C9orf75	TAPERIN	*613354	DEAFNESS, AUTOSOMAL RECESSIVE 79; DFN7	#613307	9q34.3
36216638--36216683	15,000000	97	83	ATT	GNE	UDP-N- ACETYLGLUCOSAMINE 2-EPIMERASE/N- ACETYLMANNOSAMINE KINASE	*603824	SIALURIA	#269921	9p13.3
								INCLUSION BODY MYOPATHY 2, AUTOSOMAL RECESSIVE; IBM2	#600737	9p13.3
								NONAKA MYOPATHY; NM	#605820	9p13.3
CROMOSSOMO 10										
79397345--79397388	14,666667	95	70	CCG	KCNMA1	POTASSIUM CHANNEL, CALCIUM-ACTIVATED, LARGE CONDUCTANCE, SUBFAMILY M, ALPHA MEMBER 1	*600150	GENERALIZED EPILEPSY AND PAROXYSMAL DYSKINESIA; GEPD	#609446	10q22.3
CROMOSSOMO 11										
65325326--65325365	13,500000	95	64	CAG	LTBP3	LATENT TRANSFORMING GROWTH FACTOR-BETA- BINDING PROTEIN 3	*602090	TOOTH AGENESIS, SELECTIVE, 6; STHAG6	#613097	11q12
119077000--119077033	11,333333	100	68	CGG	CBL	CAS-BR-M MURINE ECOTROPIC RETROVIRAL TRANSFORMING SEQUENCE HOMOLOG	*165360	NOONAN SYNDROME-LIKE DISORDER	#613563	11q23.3

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas (Continuação).

CROMOSSOMO 12										
49427251--49427286	12,000000	97	63	TGC	MLL2	MYELOID/LYMPHOID OR MIXED LINEAGE LEUKEMIA 2	*602113	KABUKI SYNDROME	#147920	12q12-q14
112036754--112036823	23,333334	98	131	GCT	ATXN2	ATAXIN 2	*601517	SPINOCEREBELLAR ATAXIA 2 (WADIA-SWAMI SYNDROME)	#183090	12q24
112944751--112944797	15,666667	100	94	TGA	PTPN11	PROTEIN-TYROSINE PHOSPHATASE, NONRECEPTOR-TYPE, 11	*176876	NOONAN SYNDROME 1; NS1 (FEMALE PSEUDO-TURNER SYNDROME)	#163950	12q24.1
						MOVED TO 163950 AND 610733	^163955	METACHONDROMATOSIS	#156250	12q24.1
								LEOPARD SYNDROME 1	#151100	12q24.1
								JUVENILE MYELOMONOCYTIC LEUKEMIA; JMML	#607785	17q11.2, 12q24.1, 5q31
7045880--7045938	19,666666	96	100	CAG	ATN1	ATROPHIN 1	*607462	DENTATORUBRAL-PALLIDOLUSIAN ATROPHY; DRPLA (ATAXIA, CHOREA, SEIZURES, AND DEMENTIA)	#125370	12p13.31
103352152--103352208	20,333334	91	77	GCA	ASCL1	ACHAETE-SCUTE COMPLEX, DROSOPHILA, HOMOLOG OF, 1	*100790	CENTRAL HYPOVENTILATION SYNDROME, CONGENITAL; CCHS	#209880	12q22-q23, 11p13, 10q11.2, 5p13.1-p12, 4p12, 20q13.2- q13.3
CROMOSSOMO 13										
100637703--100637748	15,333333	93	65	GCG	ZIC2	ZINC FINGER PROTEIN OF CEREBELLUM 2	*603073	HOLOPROSENCYPHALY 5; HPE5	#609637	13q32
70713484--70713517	11,333333	100	68	TAC		ATAXIN 8 OPPOSITE STRAND	*603680	SPINOCEREBELLAR ATAXIA 8; SCA8	#608768	13q21
70713516--70713561	15,333333	100	92	CTG	ATXN8OS					
CROMOSSOMO 14										
29236611--29236657	15,666667	93	67	CCA	FOXG1	FORKHEAD BOX G1 (ONCOGENE QIN)	*164874	RETT SYNDROME, CONGENITAL VARIANT	#613454	14q13
CROMOSSOMO 15										
								PROGRESSIVE EXTERNAL OPHTHALMOPLEGIA WITH MITOCHONDRIAL DNA DELETIONS, AUTOSOMAL DOMINANT, 1; PEOA1	#157640	15q25
								MITOCHONDRIAL DNA DEPLETION SYNDROME 4A (ALPERS TYPE); MTDPS4A (ALPERS DIFFUSE DEGENERATION OF CEREBRAL GRAY MATTER WITH HEPATIC CIRRHOSIS)	#203700	15q25
89876820--89876860	13,666667	97	73	GCT	POLG	POLYMERASE, DNA, GAMMA	*174763	PROGRESSIVE EXTERNAL OPHTHALMOPLEGIA WITH MITOCHONDRIAL DNA DELETIONS, AUTOSOMAL RECESSIVE; PEOB	#258450	15q25
								SENSORY ATAXIC NEUROPATHY, DYSARTHRIA, AND OPHTHALMOPARESIS; SANDO	#607459	15q25
								MITOCHONDRIAL DNA DEPLETION SYNDROME 4B (MNGIE TYPE); MTDPS4B	#613662	15q25
43910867--43910899	11,000000	100	66	CAG	STRC	STEREOCILIN	*606440	DEAFNESS, AUTOSOMAL RECESSIVE 16; DFNB16	#603720	15q15

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas (Continuação).

100252710--100252743	11,333333	100	68	CAG	MEF2A	MADS BOX TRANSCRIPTION ENHANCER FACTOR 2, POLYPEPTIDE A	*600660	CORONARY ARTERY DISEASE, AUTOSOMAL DOMINANT, 1; ADCAD1	#608320	15q26
CROMOSSOMO 16										
3778401--3778454	18,000000	92	72	TGC	CREBBP	CREB-BINDING PROTEIN	*600140	RUBINSTEIN-TAYBI SYNDROME 1; RSTS1	#180849	16p13.3
72821594--72821657	21,333334	93	92	GCC						16q22.3-q23.1,1q25,17p11,16q22.3-q23.1,16q22.1,13q12.3,11p11.2,10q25,10p15,8p22,7q11.23,7p11-q21,7p22,3p26,Xq11-q12,1q42.2-q43,22q12.3,22q12.1,20q13,19q
72831358--72831408	17,000000	92	66	TTG						
72992585--72992664	26,666666	90	88	TCC	ZFHX3	ZINC FINGER HOMEBOX 3	*104155	PROSTATE CANCER	#176807	
79633073--79633124	17,666666	92	70	CCG	MAF	V-MAF AVIAN MUSCULOAPONEUROTIC FIBROSARCOMA ONCOGENE HOMOLOG (PROTOONCOGENE MAF)	*177075	CATARACT, PULVERULENT, JUVENILE-ONSET	#610202	16q22-q23
51175656--51175685	10,000000	100	60	GCT	SALL1	SAL-LIKE 1	*602218	TOWNES-BROCKS SYNDROME; TBS (ANUS, IMPERFORATE, WITH HAND, FOOT, AND EAR ANOMALIES)	#107480	16q12.1
57983264--57983309	15,333333	93	65	CCT	CNGB1	CYCLIC NUCLEOTIDE-GATED CHANNEL, BETA-1	*600724	RETINITIS PIGMENTOSA; RP	#268000	16q13,14q31.3,2q14.1,1q41,19q13.3
CROMOSSOMO 17										
38978747--38978789	14,333333	95	68	CTC	KRT10	KERATIN 10	*148080	ICHTHYOSIS, CYCLIC, WITH EPIDERMOLYTIC HYPERKERATOSIS ERYTHRODERMA, ICHTHYOSIFORM, CONGENITAL RETICULAR; CRIE (AARAU DISEASE, INCLUDED) EPIDERMOLYTIC HYPERKERATOSIS; EHK	#607602	17q21-q22
								SMITH-MAGENIS SYNDROME; SMS	#609165	17q21-q22
									#113800	17q21-q22,12q13
17697094--17697134	13,666667	100	82	CAG	RAI1	RETINOIC ACID-INDUCED GENE 1	*607642		#182290	17p11.2
CROMOSSOMO 18										
CROMOSSOMO 19										
11558341--11558405	21,666666	90	76	GAG	PRKCSH	PROTEIN KINASE C SUBSTRATE, 80-KD, HEAVY CHAIN	*177060	POLYCYSTIC LIVER DISEASE; PCLD	#174050	19p13.2-p13.1,6q21,19p13.2-p13.1
13318673--13318712	13,333333	100	80	CTG	CACNA1A	CALCIUM CHANNEL, VOLTAGE-DEPENDENT, P/Q TYPE, ALPHA-1A SUBUNIT	*601011	SPINOCEREBELLAR ATAXIA 6; SCA6	#183086	19p13
								MIGRAINE, FAMILIAL HEMIPLEGIC, 1; FHM1	#141500	19p13
								EPISODIC ATAXIA, TYPE 2; EA2	#108500	19p13
46273463--46273524	20,666666	100	124	CAG	DMPK	DYSTROPHIA MYOTONICA PROTEIN KINASE	*605377	DYSTROPHIA MYOTONICA 1	#160900	19q13.2-q13.3

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas (Continuação).

CROMOSSOMO 20										
CROMOSSOMO 21										
CROMOSSOMO 22										
28194881--28194962	27,333334	95	128	GCT						22q12.3- qter,10q23.31, 22q12.3-qter, 22q12.2
28195603--28195653	17,000000	92	66	TGC	MNI	MENINGIOMA 1 GENE	*156100	MENINGIOMA, FAMILIAL	#607174	
CROMOSSOMO X										
25031771--25031814	14,666667	95	70	GCC	ARX	ARISTALESS-RELATED HOMEBOX, X-LINKED	*300382	PARTINGTON X-LINKED MENTAL RETARDATION SYNDROME; PRTS (MENTAL RETARDATION, X- LINKED, WITH DYSTONIC MOVEMENTS, ATAXIA, AND SEIZURE) EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE, 1; EIEE1 MENTAL RETARDATION, X- LINKED, WITH OR WITHOUT SEIZURES, ARX-RELATED; MRXARX LISSENCEPHALY, X-LINKED, 2; LISX2 (LISSENCEPHALY, X-LINKED, WITH AMBIGUOUS GENITALIA; XLAG) CORPUS CALLOSUM, AGENESIS OF, WITH ABNORMAL GENITALIA	#309510	Xp22.13
									#308350	Xp22.13
									#300419	Xp22.13
									#300215	Xp22.13
									#300004	Xp22.13
41377018--41377075	19,333334	98	107	AGC	CASK	CALCIUM/CALMODULIN- DEPENDENT SERINE PROTEIN KINASE	*300172	FG SYNDROME 4; FGS4 MENTAL RETARDATION AND MICROCEPHALY WITH PONTINE AND CEREBELLAR HYPOPLASIA	#300422	Xp11.4
									#300749	Xp11.4
47030561--47030600	13,333333	97	71	GGA	RBM10	RNA-BINDING MOTIF PROTEIN 10	*300080	TARP SYNDROME; TARPS	#311900	Xp11.23,Xp11.23- q13.3
									#313200	Xq11-q12
									#312300	Xq11-q12
									#300633	Xq11-q12
									#300068	Xq11-q12
66765159--66765261	33,333332	93	145	GCA	AR	ANDROGEN RECEPTOR	*313700	ANDROGEN INSENSITIVITY, PARTIAL HYPOSPADIAS 1, X-LINKED; HYSPI ANDROGEN INSENSITIVITY SYNDROME; AIS		1q25, 17p11, 16q22.3-q23.1, 16q22.1, 13q12.3, 11p11.2, 10q25, 10p15, 8p22, 7q11.23, 7p11- q21, 7p22, 3p26, Xq11-q12, 1q42.2-q43, 22q12.3, 22q12.1, 20q13, 19q
								PROSTATE CANCER	#176807	

Tabela 1. Relação das regiões trinucleotídicas exônicas e regiões reguladoras 5' no genoma humano. As doenças destacadas em amarelo são as que possuem associação com doenças associadas à expansão de repetições trinucleotídicas (Continuação).

66766357--66766408	17,333334	100	104	GGC	AR	ANDROGEN RECEPTOR	*313700	SPINAL AND BULBAR MUSCULAR ATROPHY, X-LINKED 1; SMAX1	#313200	Xq11-q12
								ANDROGEN INSENSITIVITY, PARTIAL	#312300	Xq11-q12
70360589--70360698	36,750000	90	123	AGC	MED12	MEDIATOR COMPLEX SUBUNIT 12 (TRINUCLEOTIDE REPEAT-CONTAINING GENE 11; TNRC11)	*300188	HYPOSPADIAS 1, X-LINKED; HYSPI	#300633	Xq11-q12
								ANDROGEN INSENSITIVITY SYNDROME; AIS	#300068	Xq11-q12
76907770--76907810	13,666667	95	64	TCC	ATRX	ATR-X GENE	*300032	1q25, 17p11, 16q22.3-q23.1, 16q22.1, 13q12.3, 11p11.2, 10q25, 10p15, 8p22, 7q11.23, 7p11-q21, 7p22, 3p26, Xq11-q12, 1q42.2-q43, 22q12.3, 22q12.1, 20q13, 19q	#176807	PROSTATE CANCER
								OPITZ-KAVEGGIA SYNDROME; OKS	#305450	Xq13
77303052--77303084	11,000000	100	66	TTG	ATP7A	ATPase, Cu(2+)-TRANSPORTING, ALPHA POLYPEPTIDE	*300011	LUJAN-FRYNS SYNDROME	#309520	Xq13
								ALPHA-THALASSEMIA MYELODYSPLASIA SYNDROME; ATMDS	#300448	Xq13
99664708--99664738	10,333333	100	62	CCG	PCDH19	PROTODADHERIN 19	*300460	ALPHA-THALASSEMIA/MENTAL RETARDATION SYNDROME, NONDELETION TYPE, X-LINKED; ATRX	#301040	Xq13
								MENTAL RETARDATION-HYPOTONIC FACIES SYNDROME, X-LINKED, 1; MRXHF1	#309580	Xq13
146993555--146993629	25,000000	94	116	GGC	ASPMR1	ANTISENSE FMR1 GENE	*300805	MENKES DISEASE	#309400	Xq12-q13
								OCIPITAL HORN SYNDROME; OHS	#304150	Xq12-q13
146993555--146993629	25,000000	94	116	GGC	FMR1	FMR1 GENE	*300805	SPINAL MUSCULAR ATROPHY, DISTAL, X-LINKED 3; SMAX3	#300489	Xq13.1-q21, Xq12-q13
								EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE, 9; EIEE9	#300088	Xq22
								FRAGILE X TREMOR/ATAXIA SYNDROME; FXTAS	#300623	Xq27.3
								FRAGILE X MENTAL RETARDATION SYNDROME	#300624	Xq27.3
								PREMATURE OVARIAN FAILURE 1; POF	#311360	Xq26-q28

CROMOSSOMO Y

- (a) Distância em pares de base a partir do centrômero (Xp-tel) como determinado utilizando o banco de dados do genoma de referência. O valor mostrado indica a distância de cada elemento microssatélite.
- (b) Número de cópias da repetição.
- (c) Porcentagem de identidade entre as cópias da repetição.
- (d) O valor da pontuação mínima de alinhamento para o microssatélite é dependente do tamanho da unidade de repetição.
- (e) Padrão de sequência consenso do elemento de repetição.
- (f) Nome do gene localizado no *locus* da repetição trinucleotídica encontrada, segundo o banco de genes do NCBI.
- (g) Nome da proteína relacionada ao *loci* trinucleotídico
- (h) *Link* sobre informações da determinada proteína relacionada ao *locus* analisado segundo banco de dados *OMIM* (*Online Mendelian Inheritance in Man*).
- (i) Doença relacionada ao *locus* determinado
- (j) *Link* sobre informações da determinada doença relacionada ao *locus* analisado segundo banco de dados *OMIM*.
- (k) *Locus* cromossômico relacionado a doença de acordo com o banco de dados *OMIM*.

Tabela 2. Informações referentes às doenças poliglutamínicas.

Doenças Poliglutamínicas (causadas por expansão de repetições CAG)										
Doenças	Link OMIM	Localização	Gene	Produto	Alelos			Sequência	Origem parental principal da mutação	Principais características clínicas
					Normais	Pré-mutação	Mutação			
Ataxia espinocerebelar 1 (SCA1)	#164400	6p22.3	<i>ATXN1</i> (éxon 8)	Ataxina-1	6–39	40	41–83	CAG	Paterna	Ataxia, fala arrastada, espasticidade e dificuldade cognitiva
Ataxia espinocerebelar 2 (SCA2)	#183090	12q24.12	<i>ATXN2</i> (éxon 1)	Ataxina-2	1-31	32	33–200	CAG	Paterna	Ataxia, polineuropatia, diminuição dos reflexos e variante infantil com retinopatia
Ataxia espinocerebelar 3 (ou Machado-Joseph) (SCA3 ou MJD)	#109150	14q32.12	<i>ATXN3</i> (éxon 8)	Ataxina-3	12–40	41–85	>86	CAG	Paterna	Ataxia, parkinsonismo e espasticidade
Ataxia espinocerebelar 6 (SCA6)	#183086	19p13.2	<i>CACNA1A</i> (éxon 47)	Subunidade alfa 1A do canal de cálcio neuronal	<18	19	20–33	CAG	Paterna	Ataxia, disartria, nistagmo e tremores
Ataxia espinocerebelar 7 (SCA7)	#164500	3p14.1	<i>ATXN7</i> (éxon 3)	Ataxina-7	4–17	18–33	34-460	CAG	Paterna	Ataxia, cegueira e insuficiência cardíaca em crianças
Ataxia espinocerebelar 12 (SCA12)	#604326	5q32	<i>PPP2R2B</i> (5' UTR)	Subunidade da proteína reguladora PP2A	7–28	29–66	67–78	CAG	Paterna e Materna (mais instável quando de origem paterna)	Ataxia e convulsões
Ataxia espinocerebelar 17 (SCA17)	#607136	6q27	<i>TBP</i> (éxon 3)	TATA box-ligante de proteína	25–42	43–48	49–66	CAG	Paterna > Materna	Ataxia, declínio cognitivo, convulsões e problemas psiquiátricos
Atrofia dentatorrubropalidoluisiana (DRPLA)	#125370	12p13.31	<i>ATN1</i> (éxon 5)	Atrofina-1	6–35	36–48	49–88	CAG	Paterna	Ataxia, convulsões, demência e coreoatetose

Tabela 2. Informações referentes às doenças poliglutamínicas.

Doenças Poliglutamínicas (causadas por expansão de repetições CAG)										
Doenças	Link OMIM	Localização	Gene	Produto	Alelos			Sequência	Origem parental principal da mutação	Principais características clínicas
					Normais	Pré-mutação	Mutação			
Atrofia muscular bulboespinal (Kennedy) (SBMA)	#313200	Xq12	<i>AR</i> (ou <i>HUMARA</i>) (éxon 1)	Receptor de andrógenos	13–31	32–39	>40	CAG	Paterna	Fraqueza motora, dificuldade de deglutição, ginecomastia, diminuição da fertilidade
Doença de Huntington (HD)	#143100	4p16.3	<i>HTT</i> (éxon 1)	Huntingtina	6–29	30–37	38–180	CAG	Paterna	Coréia, distonia, déficits cognitivos, problemas psiquiátricos

Tabela 3. Informações referentes às doenças não-poliglutamínicas.

Doenças Não-poliglutamínicas (causadas por expansão de repetições não-CAG)										
Doenças	Link OMIM	Localização	Gene	Produto	Alelos			Sequência	Origem parental principal da mutação	Principais características clínicas
					Normais	Pré-mutação	Mutação			
Ataxia de Friedreich (FRDA)	#229300	9q21.11	<i>FXN</i> (intron 1)	Frataxina	5-30	31-100	101-1000	GAA	Recessiva	Ataxia sensorial, cardiomiopatia e diabetes
Ataxia espinocerebelar 8 (SCA8)	#608768	13q21	<i>ATXN8OS</i> (3' UTR)	Transcrito /não traduzido	15-34	35-89	90-250	CTG	Materna	Ataxia, fala arrastada e nistagmo
Distrofia Miotônica 1 (DM1)	#160900	19q13.32	<i>DMPK</i> (3' UTR)	Distrofia miotônica-proteína quinase	5-37	38-50	>50	CTG	Materna	Miotonia, fraqueza, defeitos de condução cardíaca, resistência à insulina, catarata, atrofia testicular e retardo mental em forma congênita
Síndrome do X-frágil (FRAXA)	#300624	Xq27.3	<i>FMR1</i> (5' UTR)	Proteína ligante de RNA - FMRP	4-50	51-200	>200	CGG	Materna	Retardo mental, macroorquidismo, defeitos do tecido conjuntivo e anormalidades comportamentais
Retardo Mental ligado ao Cromossomo X (FRAXE)	#300948	Xq28	<i>AFF2</i> (5' UTR)	Proteína ligante de RNA - FMR2	4-39	40-200	>200	GCC	Materna	Retardo mental, dificuldade de aprendizado e retardo no desenvolvimento

4.2. Desenho de marcadores para as quinze doenças causadas por expansão de repetições trinucleotídicas

A tabela 4 mostra os iniciadores gene e repetições-específicos desenhados *in silico* para o *locus* relacionado a cada doença, e o tamanho dos *amplicons* dos respectivos marcadores. Os marcadores foram nomeados com a inicial de três letras D(número do cromossomo)S (Cópia única dentro do determinado cromossomo) acompanhado pelo número que indica a distância em pares de base a partir do telômero do braço curto do cromossomo em questão. Por exemplo, D12S7045859 é um marcador microssatélite com cópia única dentro do cromossomo 12 com localização em 7.045.859pb no genoma de referência deste cromossomo.

Os iniciadores repetição-específicos foram desenhados para a técnica de TP-PCR, de acordo com os critérios descritos na literatura. O tamanho do *amplicon* nesta técnica varia com o número de repetições trinucleotídicas que a região polimórfica do *locus* específico possui. Os iniciadores gene-específicos utilizados para essa técnica foram os mesmos utilizados para a técnica de QF-PCR.

Tabela 4. Relação dos iniciadores gene e repetição-específicos desenhados e tamanho dos *amplicons* dos marcadores para as quinze doenças causadas por expansão de repetições trinucleotídicas. Em amarelo está sendo destacada a região comum entre o iniciador 2 e 3 (P2 e P3).

Marcador	Doença	Iniciadores gene-específicos		<i>Amplicon</i> (pb)	Iniciadores Repetição-específicos	
		Iniciador sentido (P1)	Iniciador antissentido (P4)		Iniciador de repetições trinucleotídicas (P3)	P2
DOENÇAS POLIGLUTAMÍNICAS						
D6S16327787	Ataxia espinocerebelar 1 (SCA1)	CAGACGCCGGGACACAAG	TGTGGACGTACTGGTTCTGCTG	184	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG
D12S112036658	Ataxia espinocerebelar 2 (SCA2)	CCTCACCATGTCGCTGAAG	AGGAGACCGAGGACGAGG	188	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG
D14S92537354	Ataxia espinocerebelar 3 ou Doença de Machado-Joseph (SCA3 ou MJD)	GTTTCAGACAGCAGCAAAAGC	GCCTTACCTAGATCACTCCAA	133	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG
D19S13318390	Ataxia espinocerebelar 6 (SCA6)	GTGTCCTATTCCCCTGTGATCC	GACCCGCCTCTCCATCCT	225	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG
D3S63898258	Ataxia espinocerebelar 7 (SCA7)	GAAAGAATGTCGGAGCGG	TTCAGGACTGGGCAGAGG	270	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG
D5S146258259	Ataxia espinocerebelar 12 (SCA12)	CACTGCTGCTGGGAAAGAGT	ACTCACCTCACCCACAC	150	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG
D6S170870832	Ataxia espinocerebelar 17 (SCA17)	TGGAAGAGCAACAAAGGCA	GTTGACTGCTGAACGGCTG	163	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG
D12S7045859	Atrofia dentatorrubropalidolusiana (DRPLA)	ATCACCACCAGCAACAGCA	GAGCTACCGCCTCCAGT	130	TACGCATCCCAGTTTGAGACG	TACGCATCCCAGTTTGAGACG

Tabela 4. Relação dos iniciadores gene e repetição-específicos desenhados e tamanho dos *amplicons* dos marcadores para as quinze doenças causadas por expansão de repetições trinucleotídicas. Em amarelo está sendo destacada a região comum entre o iniciador 2 e 3 (P2 e P3) (Continuação).

Marcador	Doença	Iniciadores gene-específicos			Iniciadores Repetição-específicos	
		Iniciador sentido (P1)	Iniciador antissentido (P4)	<i>Amplicon</i> (pb)	Iniciador de repetições trinucleotídicas (P3)	P2
DXS66765056	Atrofia muscular bulboespinal ou Doença de Kennedy (SBMA)	GTGCGCGAAGTGATCCAGAA	CCAGGACCAGGTAGCCTGTG	244	TACGCATCCCAGTTTGAGACG GCTGCTGCTGCTGCTG	TACGCATCCCAGTTTGAGACG
D4S3076583	Doença de Huntington (HD)	TGGCGACCCTGGAAAAGCTGAT	GCGGTGGCGGCTGTTGCTGCT	167	TACGCATCCCAGTTTGAGACG GCTGCTGCTGCTGCTG	TACGCATCCCAGTTTGAGACG
DOENÇAS NÃO-POLIGLUTAMÍNICAS						
D9S71652012	Ataxia de Friedreich (FRDA)	GACTAACCTGGCCAACATGG	GCCCGGCTAACTTTTCTTTA	103	TACGCATCCCAGTTTGAGACG TTCTTCTTCTTCTTC	TACGCATCCCAGTTTGAGACG
D13S70713460	Ataxia espinocerebelar 8 (SCA8)	CCTTCATGTTAGAAAACCTGGC	GCAGTATGAGGAAGTGTGAAA	181	TACGCATCCCAGTTTGAGACG CAGCAGCAGCAGCAGC	TACGCATCCCAGTTTGAGACG
D19S46273433	Distrofia Miotônica 1 (DM1)	GGTCCTGTAGCCGGGAA	GGCCGAAAGAAAGAAATGGT	112	TACGCATCCCAGTTTGAGACG CAGCAGCAGCAGCAGC	TACGCATCCCAGTTTGAGACG
DXS146993454	Síndrome do X-frágil (FRAXA)	CTCCGTTTCGGTTTCACTTC	AGCTCCTCCATCTTCTTTCAG	255	TACGCATCCCAGTTTGAGACG CGCCGCCGCCGCCGCCG	TACGCATCCCAGTTTGAGACG
DXS147582106	Retardo Mental ligado ao Cromossomo X (FRAXE)	TGTGTAAGTGTGTGATGCTGCCG	CACTCGGGTCCGCTCGGT	279	TACGCATCCCAGTTTGAGACG CGGCGGCGGCGGCGGCGG	TACGCATCCCAGTTTGAGACG

específico. Os próximos alelos serão apenas a soma de mais uma repetição, pois há durante a amplificação o anelamento contínuo dos iniciadores, uma a uma repetição. Assim, o alelo 90 corresponde a 6 repetições CAG, o alelo 93, a 7 repetições. E assim por diante até o alelo maior, de 129pb que corresponde a 19 repetições. Ou seja, nesta sequência mostrada, o indivíduo possui 19 repetições CAG no *locus* específico para o gene *HTT* associado à Doença de Huntington.

4.3. Padronização dos ensaios para três dos marcadores desenhados

Para este trabalho foram padronizadas as condições dos ensaios para os marcadores apenas da Doença de Machado-Joseph, Doença de Kennedy, Síndrome do X-frágil, Ataxia espinocerebelar 2 e Distrofia miotônica para QF-PCR e para as três primeiras doenças para técnica de TP-PCR. Para a padronização foram analisados os seguintes parâmetros: efeito da concentração de DNA alvo, concentração de $MgCl_2$, temperatura e tempo de anelamento e de extensão dos iniciadores na amplificação da sequência de interesse, e da extensão final, além do número de ciclos utilizados para amplificação. Cada marcador testado *in vitro* foi padronizado em reação uniplex.

4.3.1. Padronização de ensaio para QF-PCR

Para QF-PCR, primeiramente foram realizadas curvas de concentração de DNA genômico com 10, 20, 30, 40 e 50ng por reação, gradiente com variações na concentração de $MgCl_2$ de 1,0mM; 1,5mM; e 2,0mM e gradiente de temperatura de anelamento entre 57 e 62°C, com uma variação de 1°C entre as temperaturas. Para realização dos ensaios foi utilizada uma variação na concentração de 0,32 a 0,8 μ M de cada iniciador gene-específico, e 2M de betaína em um volume final de 12,5 μ L para as reações de QF-PCR e TP-PCR. O ciclo inicial de amplificação utilizado foi o seguinte: Desnaturação inicial de 95°C por 11min, seguido de 28 ciclos de 94°C/1 min para desnaturação, gradiente de 57 a 62°C para o anelamento dos iniciadores por 1min e 72°C/1min para extensão da fita, e após, mantido a 60°C por 1hora para extensão final para então ser armazenado a 4°C.

No entanto, após testes foi observado que a quantidade de ciclos, tempo da fase de anelamento e de extensão do ciclo e extensão final também deveriam ser modificados. Para cada um dos ensaios foi encontrado, após diversos testes, a combinação das suas condições ideais de amplificação, como mostrado resumidamente na tabela 5.

Todos os ensaios apresentaram melhor amplificação com 2,0mM de $MgCl_2$, concentração de 0,8 μ M de cada iniciador, e tempo de desnaturação de um minuto. Para cada passo da amplificação foram realizados testes separadamente, visando à otimização da reação.

Tabela 5. Condições otimizadas para a QF-PCR dos cinco marcadores testados *in vitro*.

Marcador	Temperatura de anelamento (°C)	Número de ciclos	Tempo de desnaturação (min)	Tempo de anelamento (min)	Tempo de extensão (min)	Temperatura de extensão final (°C)	Tempo de extensão final (min)	Concentração de DNA (ng/μL)
D14S92537354	58	28	1	1	1	68	60	20
DXS66765056	59	28	1	1	1	68	60	50
DXS146993454	65	35	1	1	2	72	7	50
D12S112036658	61	30	1	1,5	1,5	68	60	30
D19S46273433	58	35	1	1	1,5	72	10	15

Para o marcador desenhado para o diagnóstico da Doença de Machado-Joseph (D14S92537354), o ensaio que apresentou uma melhor resolução consistiu na utilização de uma reação de amplificação com 28 ciclos, mantendo as temperaturas e duração de cada fase do ciclo inicial. A temperatura ideal de anelamento foi de 58°C e concentração de DNA genômico de 20ng/μL (figura 13A).

Já para o ensaio com os marcadores DXS66765056 e DXS146993454 desenhados no *locus* dos genes causadores da Doença de Kennedy (SBMA) e Síndrome do X-frágil (FRAXA), respectivamente foram utilizados 35 ciclos para amplificação na reação de QF-PCR, e concentração de 50ng/μL de DNA genômico. No entanto a temperatura de anelamento em SBMA foi de 59°C e para FRAXA, 65°C. Os eletroferogramas representativos após otimização dos ensaios são mostrados nas figuras 13B e 13C.

Para obter uma melhor resolução no eletroferograma, a reação de QF-PCR do marcador para Ataxia Espinocerebelar 2 (D12S112036658) foi otimizada, utilizando 30 ciclos para amplificação, temperatura de anelamento de 61°C e 30ng/μL de DNA genômico (eletroferograma na figura 13D). E para o marcador da Distrofia miotônica (D19S46273433) foram necessários 35 ciclos com temperatura de anelamento de 58°C e utilizando 15ng/μL para uma melhor resolução do eletroferograma (figura 13E).

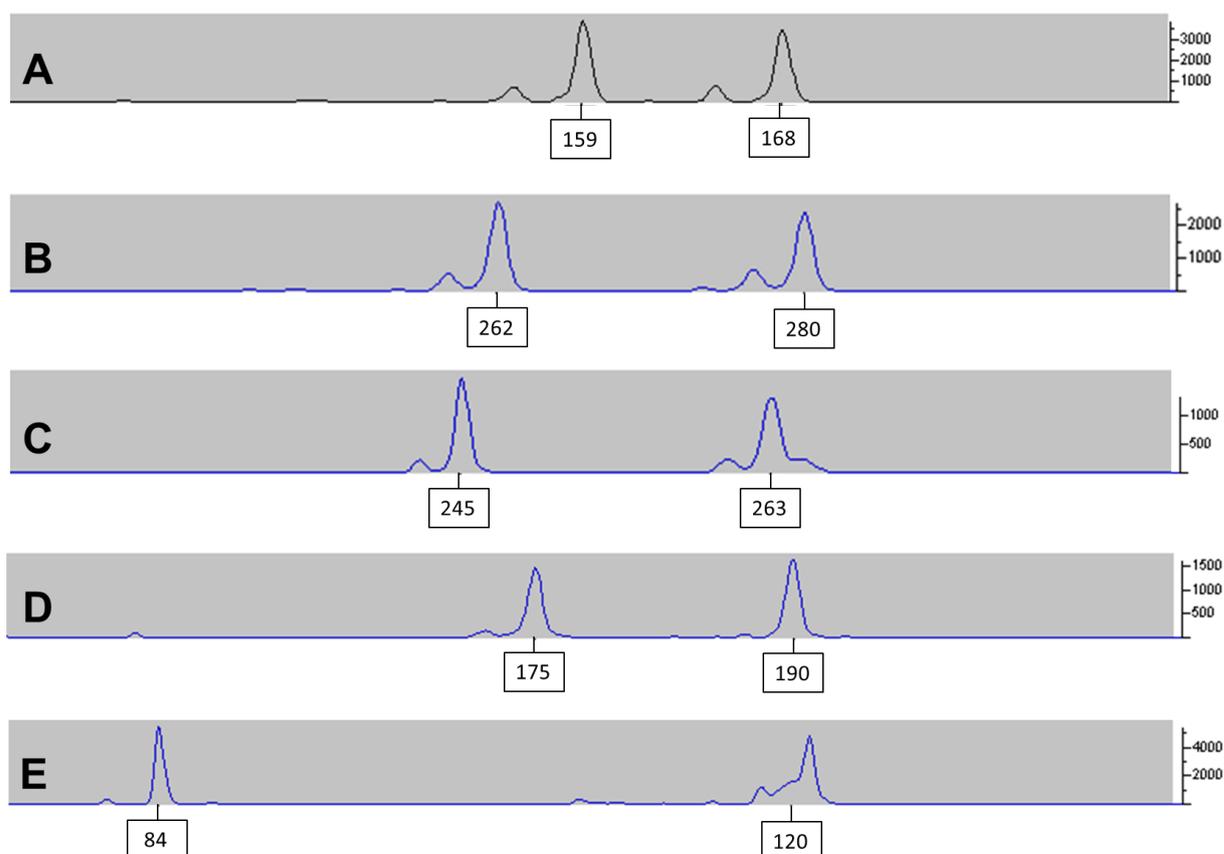


Figura 13. Eletroferogramas representativos dos perfis alélicos de amplificação sob condições otimizadas. (A) Perfil alélico de amplificação do ensaio para o marcador D14S92537354 utilizado para diagnóstico da Doença de Machado-Joseph; (B) Perfil alélico de amplificação do ensaio para o marcador DXS66765056 utilizado para diagnóstico da Doença de Kennedy; (C) Perfil alélico de amplificação do ensaio para o marcador DXS146993454 utilizado para diagnóstico da Síndrome do X-frágil; (D) Perfil alélico de amplificação do ensaio para o marcador D12S112036658 utilizado para diagnóstico da Ataxia Espinocerebelar 2; (E) Perfil alélico de amplificação do ensaio para o marcador D19S46273433 utilizado para diagnóstico da Distrofia Miotônica.

O número de repetições trinucleotídicas aproximado dos *loci* analisados foram obtidos pela técnica de QF-PCR, onde foi inferido, excluindo o número de nucleotídeos não variáveis do total de nucleotídeos e dividido o resultado por três. Na tabela 6 é discriminado o número de nucleotídeos não variáveis de cada um dos quinze marcadores. A figura 14 ilustra o esquema para detecção do número de repetições do *locus* do gene *AR*, utilizando a técnica de QF-PCR com iniciadores gene-específicos.

4.3.2. Padronização de ensaio para TP-PCR

Para a técnica de TP-PCR foi realizada apenas a padronização para a Doença de Machado-Joseph, Doença de Kennedy e para a Síndrome do X-frágil, não tendo sido realizados estudos de frequência alélica utilizando esta técnica.

Para a Doença de Machado-Joseph e Doença de Kennedy, foi utilizada a versão *tetraprimer* do ensaio de TP-PCR, sendo dois iniciadores repetição-específicos e os outros dois iniciadores, gene-específicos. O exemplo dessa versão da técnica é ilustrado pelo esquema da figura 15 para o marcador D14S92537354. Já para a Síndrome do X-frágil foi desenhada a versão *triprimer* do ensaio, sendo dois repetição-específicos e um iniciador gene-específico, marcado com fluorescência. Para este ensaio o iniciador marcado foi o antissentido. O esquema para essa versão da técnica é mostrado na figura 16.

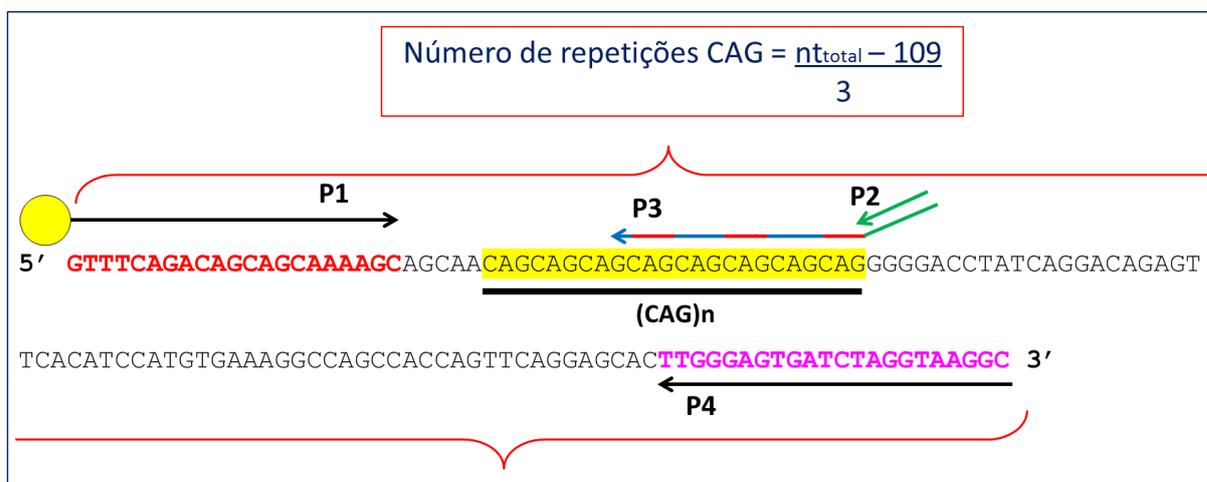


Figura 15. Esquema para detecção do número de repetições utilizando a versão *tetraprimer* do ensaio de TP-PCR. Expansões (CAG)>52 no gene *ATXN3*, deste exemplo causam Ataxia espinocerebelar 3 (síndrome de Machado-Joseph), com penetrância incompleta. Em amarelo é destacada a sequência polimórfica do gene *ATXN3* [(CAG)*n*]. Os *amplicons* gerados pela associação do iniciador P4 com os iniciadores P2 e P3 irá identificar o número de repetições CGG de cada alelo do indivíduo. E o *amplicon* obtido da amplificação do iniciador P1 com o P4 irá determinar seus alelos.



Figura 16. Esquema para detecção do número de repetições utilizando a versão *triprimer* do ensaio de TP-PCR. No exemplo é mostrado o desenho para o gene *FMR1*, relacionado com a Síndrome do X-frágil. Em amarelo está destacada a sequência polimórfica [CGG] n do gene *FMR1*. O *amplicon* gerado pela associação do iniciador P4 com os iniciadores P2 e P3 irá determinar o número de repetições CGG de cada alelo no indivíduo.

Para padronização da técnica de TP-PCR, as mesmas variáveis da QF-PCR foram testadas. Foi observado que o tempo de extensão do ciclo e final nessa técnica deveria ser aumentado, para permitir uma completa extensão dos longos produtos na reação PCR (Warner *et al.*, 1996). Foram realizados diversos testes de amplificação por essa técnica para cada uma dos três marcadores testados *in vitro*.

Inicialmente os marcadores foram testados utilizando os ciclos padronizados para QF-PCR. Posteriormente foi realizado um gradiente de tempo de extensão do ciclo. A extensão do ciclo foi testada com 1; 1,5; 2 e 2,5min. O número de ciclos também foi aumentado, assim como o tempo e temperatura de anelamento. À medida que uma dessas variáveis era modificada, as outras se mantiveram para verificar apenas o efeito da variável testada. Já o tempo de extensão final foi aumentado somente na Síndrome do X-frágil.

Para a versão *tetraprimer*, foi realizada a técnica de TP-PCR associada à de QF-PCR, que é feita primeiro para aumentar o sinal de amplificação, pois o molde para a TP-PCR serão os produtos amplificados por QF-PCR. Assim, à reação de TP-PCR propriamente dita, são adicionados somente os iniciadores repetição-específicos e o iniciador sentido gene-específico, para amplificação nesta segunda PCR. E no eletroferograma são mostrados os produtos de amplificação das duas técnicas de PCR, ou seja, o produto de amplificação dos quatro iniciadores.

Para a QF-PCR associado à TP-PCR, foram realizadas diluições do produto de amplificação de 1/10 (1 μ L do produto para 10 μ L de água miliQ) até 1/100, com uma variação de 10 μ L entre cada diluição. Essa diluição é utilizada como molde para a reação de TP-PCR.

Na Doença de Machado-Joseph, o tempo de extensão do ciclo teve que ser aumentado para 2min. A temperatura de anelamento e de extensão e o tempo de extensão final foram mantidos. A melhor diluição foi a de 1 μ L do produto de amplificação para 30 μ L de água miliQ. E o número de ciclos foi aumentado para 30. Essas foram as condições que apresentaram uma melhor resolução no eletroferograma para o ensaio do o marcador D14S92537354, como mostrado na figura 17.

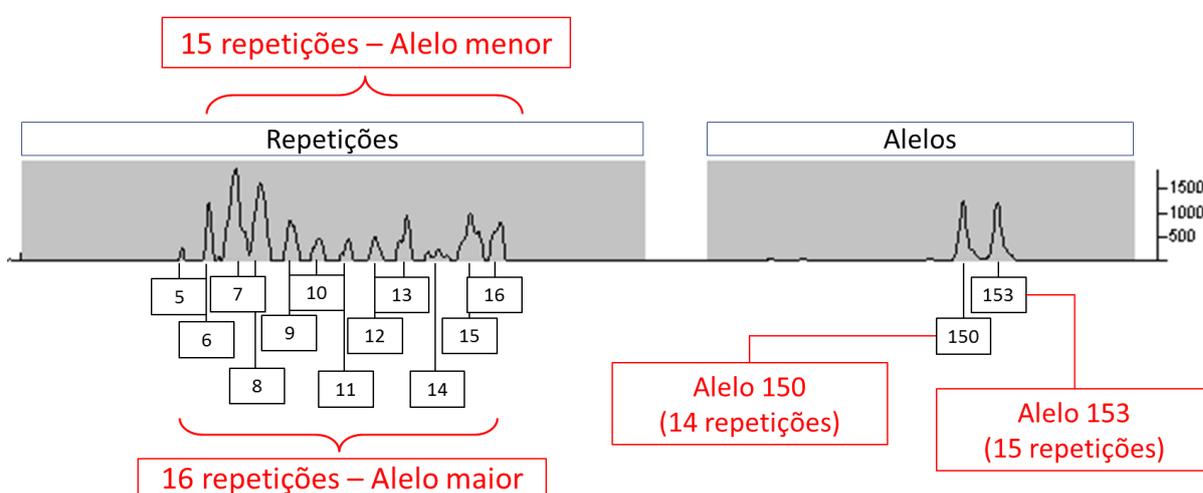


Figura 17. Eletroferograma representativo dos perfis dos produtos de amplificação da região CAG do gene *ATXN3*. Estão sendo determinados o número de repetições (CAG) n em cada alelo identificado. Os alelos 150 e 153, com 14 e 15 repetições, respectivamente, correspondem à amplificação com iniciadores gene-específicos. E com a técnica de TP-PCR foram encontradas 15 e 16 repetições, respectivamente. Para amplificação do marcador D14S92537354 foi utilizada a técnica de QF-PCR associada à de TP-PCR.

Para a Doença de Kennedy, o tempo de extensão passou de 1min para 2min e o tempo de anelamento, para 1,5min. A temperatura de anelamento foi diminuída para 58°C e o número de ciclos permaneceu o mesmo. A diluição para esta reação foi de 1 μ L do produto de amplificação para 10 μ L de água miliQ. Na figura 18 é exibido o eletroferograma representativo da padronização deste ensaio para o marcador DXS66765056.

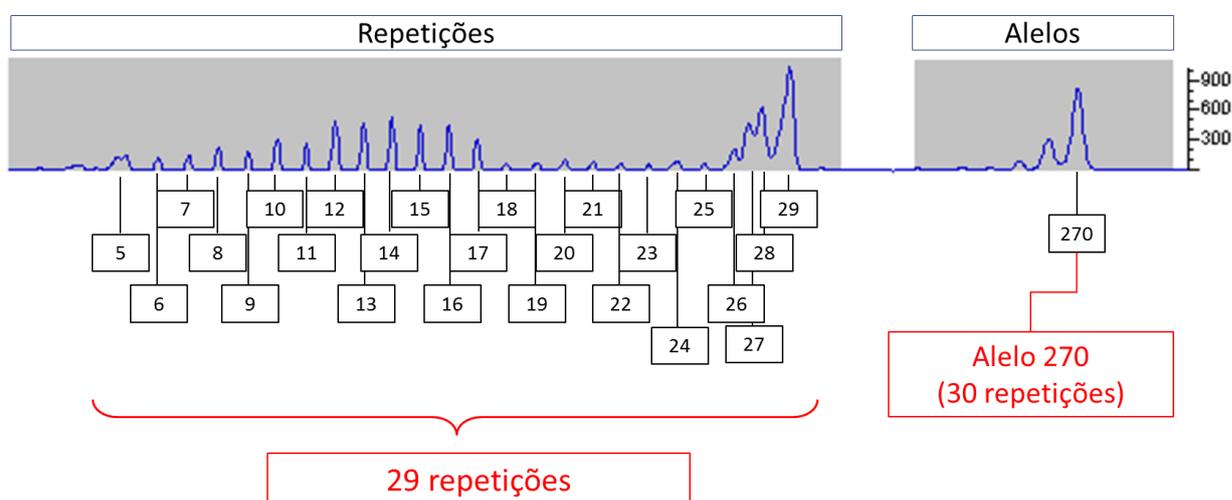


Figura 18. Eletroferograma representativo dos perfis dos produtos de amplificação da região CAG do gene AR. Estão sendo determinados os números de repetições $(CAG)_n$ em cada alelo identificado. O alelo 270 corresponde à amplificação com iniciadores gene-específicos e observa-se que há aproximadamente 30 repetições CAG. Utilizando os iniciadores repetição-específicos, foram encontradas 29 repetições CAG. Para amplificação do marcador DXS66765056 foi utilizada a técnica de QF-PCR associada à de TP-PCR.

Para o marcador DXS146993454 foi utilizada a versão *triprimer* da técnica de TP-PCR. O ensaio que apresentou uma melhor amplificação para este marcador utilizou 30 ciclos para amplificação e temperatura de anelamento de 64°C por 2 minutos. O tempo de extensão do ciclo foi mantido com 2min e o tempo de extensão final foi aumentado para 20min. Para esta reação foi adicionado 50ng de DNA genômico. O eletroferograma do ensaio padronizado para o gene *FMR1* é ilustrado na figura 19.

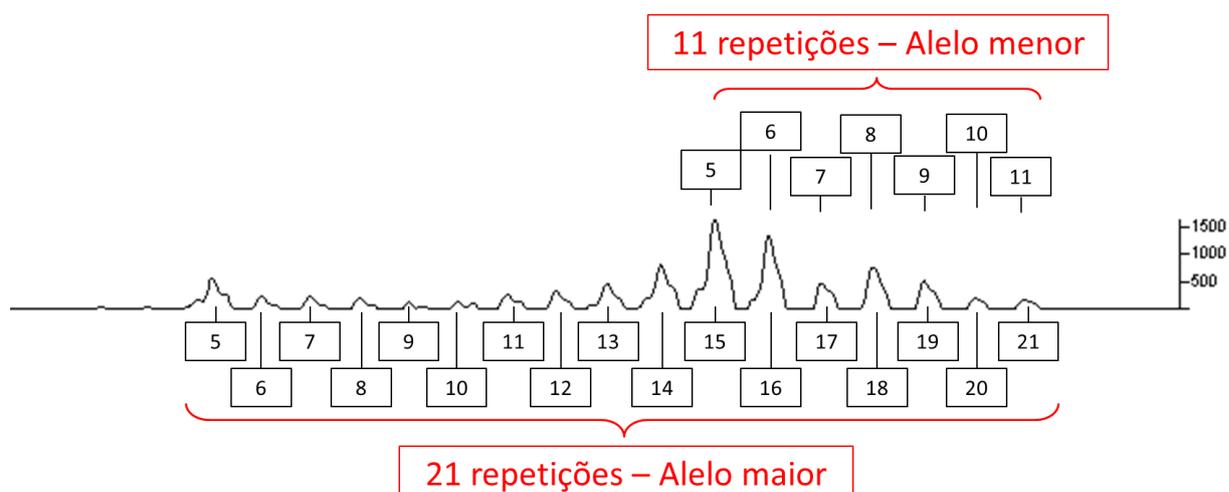


Figura 19. Eletoferograma representativo dos perfis dos produtos de amplificação da região CGG do gene *FMR1*. O indivíduo possui 11 repetições CGG no alelo menor e 21 repetições no maior.

4.4. Dados populacionais relativos aos marcadores

Para dados populacionais foram tipadas amostras apenas para os marcadores D14S92537354, DXS146993454, e DXS66765056, relacionados ao diagnóstico da doença de Machado-Joseph, Síndrome do X-frágil, e Doença de Kennedy, respectivamente. Após padronização destes ensaios para QF-PCR, estes foram validados através da utilização de amostras controle, referentes a indivíduos normais não aparentados. Os ensaios foram satisfatórios e confirmaram o perfil genético esperado para as amostras.

Foram determinados parâmetros populacionais pela tipagem de mulheres não relacionadas geneticamente para os marcadores D14S92537354, DXS146993454, e DXS66765056, selecionadas aleatoriamente do banco de DNA do laboratório. As variações em comprimento para um determinado alelo foram sempre múltiplas de três, por serem marcadores de repetições trinucleotídicas. Os alelos foram nomeados conforme o tamanho (em pares de bases) do produto de amplificação observado. As frequências de cada alelo foram determinadas e estão apresentadas na tabela 7.

Em um total de 200 alelos, os marcadores para o polimorfismo $(CAG)_n$ do gene *ATXN3* e *AR* apresentaram taxa de heterozigose de 0,85 e 0,84,

respectivamente. E o marcador para o polimorfismo $(CGG)_n$ do gene *FMR1*, em um total de 170 alelos, apresentou taxa de heterozigose de 0,29.

Os padrões de combinações entre os alelos encontrados na população normal e a distribuição da heterozigose e homozigose na população amostral para estas doenças foram obtidos e estão apresentados nas tabelas 8, 9 e 10, respectivamente.

Tabela 7. Frequências alélicas para os marcadores D14S92537354, DXS66765056 e DXS146993454 para os genes *ATXN3*, *AR*, e *FMR1* respectivamente (n=indivíduos).

Alelo	D14S92537354 (n=100)	DXS66765056 (n=100)	DXS146993454 (n=85)
126	0,02		
129	0,24		
141	0,02		
144	0,01		
147	0,03		
150	0,07		
153	0,1		
156	0,215		
159	0,04		
162	0,01		
168	0,11		
171	0,015		
174	0,035		
177	0,015		
180	0,005		
183	0,01		
186	0,025		
189	0,005		
192	0,01		
195	0,01		
204	0,005		
206			0,012
221			0,014
233			0,007
236			0,041
242			0,061
245			0,054
248			0,02
250		0,025	
251			0,02
253		0,02	
256		0,01	
257			0,061
259		0,064	
260			0,324
262		0,079	
263			0,338
265		0,089	

Tabela 7. Frequências alélicas para os marcadores D14S92537354, DXS66765056 e DXS146993454 para os genes ATXN3, AR, e FMR1 respectivamente (n=indivíduos).
(Continuação).

Alelo	D14S92537354 (n=100)	DXS66765056 (n=100)	DXS146993454 (n=85)
266			0,041
268		0,129	
269			0,014
271		0,208	
274		0,084	
277		0,074	
278			0,007
280		0,104	
283		0,05	
286		0,035	
289		0,02	
295		0,005	
298		0,005	

Tabela 10. Padrões de combinações entre os alelos do gene *FMR1* encontrados na população normal. Distribuição da heterozigose e homozigose na população amostral. Em sombreado é mostrado o número de alelos com genótipo homozigoto para o determinado alelo e seu respectivo número de repetições CGG correspondente.

Alelos		206	221	233	236	242	245	248	251	257	260	263	266	269	278
Repetições		4	9	13	14	16	17	18	19	21	22	23	24	25	28
206	4	1													
221	9		1												
233	13														
236	14				2										
242	16					4									
245	17						2								
248	18														
251	19								1						
257	21							2							
260	22			1	1		1		1		22				
263	23		1		1		2	1			5	23			
266	24					1	1				2	2			
269	25									1	1				
278	28												1		

Após genotipagem dos 100 indivíduos para a Doença de Machado-Joseph, foram calculados os números de repetições $(CAG)_n$, que variou entre 6 e 32 com distribuição bimodal, sendo os alelos com 7 e 16 repetições CAG os mais frequentes. Esses dados são apresentados no gráfico 1.

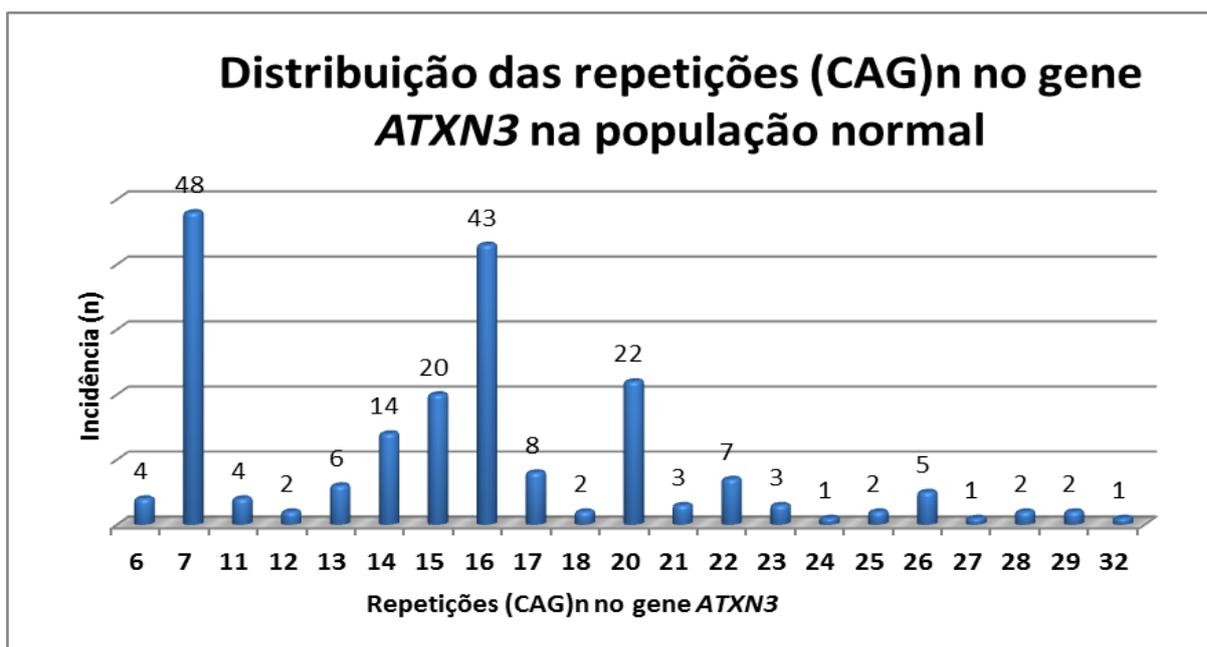


Gráfico 1. Frequência de distribuição das repetições $(CAG)_n$ no gene *ATXN3*, relacionado à Doença de Machado-Joseph. 100 amostras do banco de DNA do NUDIM foram aleatoriamente selecionadas garantindo o anonimato dos doadores.

Para a Síndrome do X-frágil foi realizada a genotipagem de 85 indivíduos. 14 alelos foram encontrados e calculados os números de repetições $(CGG)_n$, que variou entre 4 e 28 com distribuição bimodal de 22 e 23 repetições CGG. Nesta varredura não foram encontrados alelos patogênicos. Esses dados são apresentados no gráfico 2.

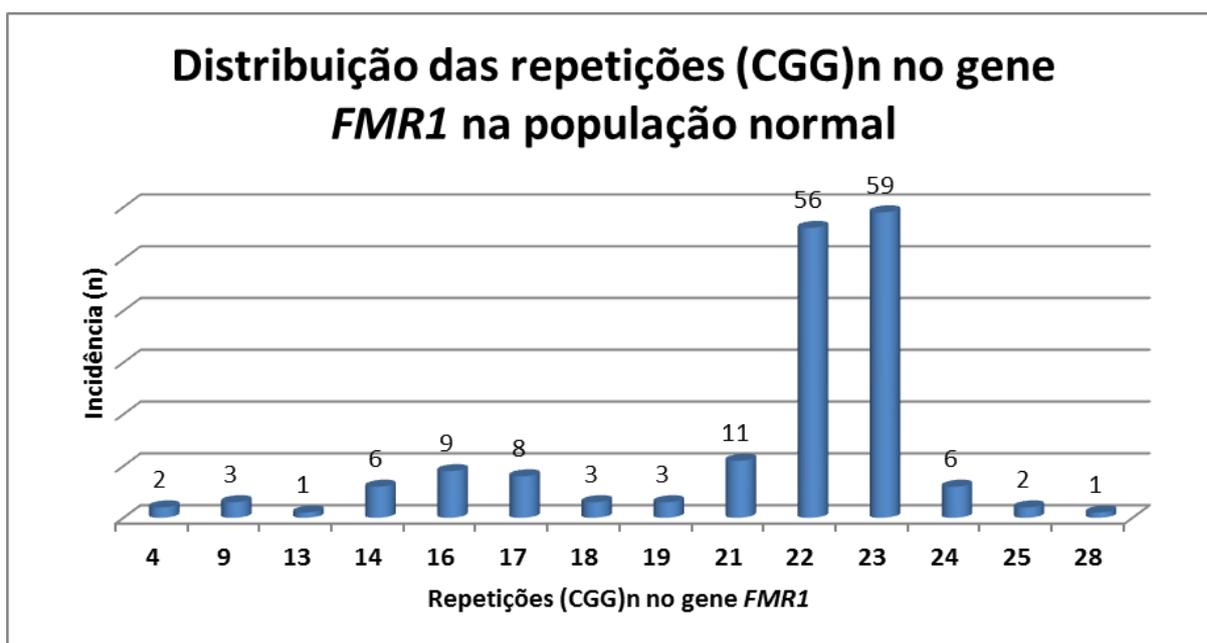


Gráfico 2. Frequência de distribuição das repetições (CGG)_n no gene *FMR1*, relacionado à Síndrome do X-frágil. 85 amostras do banco de DNA do NUDIM foram aleatoriamente selecionadas garantindo o anonimato dos doadores.

Para a Doença de Kennedy, nesta subpopulação, o número de repetições variou de 24 e 39, com distribuição unimodal, tendo como moda, 31 repetições (gráfico 3).

Como mostrado anteriormente, na tabela 2, alelos com número de repetições na faixa de 32 até 39 repetições CAG no gene *AR* caracterizam alelos com pré-mutação. No gráfico 3 esse grupo está sendo destacado em amarelo, onde se observa que 75 dos 200 alelos analisados possuem alelos com número de repetições na faixa de pré-mutação. Os indivíduos que apresentaram pelo menos um alelo com pré-mutação, tiveram seus núcleos familiares tipados com o marcador DXS66765056 para investigação de possíveis expansões na geração posterior, por se tratar de alelos instáveis.

No tocante à possibilidade de maternidade, os alelos expostos na tabela 11 deixam evidente que em todos os marcadores genéticos polimórficos analisados, o perfil alélico de 226F apresenta alelos em comum com o perfil alélico de 226M, estabelecendo forte evidência a favor da condição de maternidade em relação à 226M (Índice Cumulativo de Maternidade, CMLR = 651834; Probabilidade de Maternidade, MP = 99,9998%). Ou seja, o alelo expandido transmitido não é um alelo de exclusão de parentesco, conforme verificado pelo teste de maternidade. Assim, conclui-se que houve uma expansão de duas repetições CAG, como mostrado pelo heredograma na figura 21.

Tabela 11. Evidência genética a favor da maternidade. Perfis alélicos nucleares (genótipos) identificados pelas análises de marcadores genéticos nas amostras biológicas doadas por 226M (Mãe) e 226F (Filho).

MARCADOR	LR ^(*)	FÓRMULA	FREQUÊNCIAS ^(**)	226M	226F	AMO ^(***)
D8S1179	1,96046	1 / 4t	t=0,128	11/15	14/15	14/15
D21S11	0,98552	1 / 4r	r=0,254	30/32,2	28/30	28/30
D7S820	1,93765	(p+q) / 4pq	p=0,26 q=0,256	10/11	10/11	10/11
CSF1PO	1,53075	1 / 2r	r=0,327	12	10/12	10/12
D3S1358	3,96259	1 / p	p=0,252	16	16	16
TH01	3,22179	(p+r) / 4pr	p=0,206 r=0,124	6/8	6/8	6/8
D13S317	1,67628	1 / 2p	p=0,298	11/12	11	11
D16S539	1,56606	1 / 4p	p=0,16	9/11	9/10	9/10
D2S1338	2,03045	1 / 4s	s=0,123	16/19	19/24	19/24
D19S433	53,511	(q+u) / 4qu	q=0,237 u=0,00477	13/17	13/17	13/17
VWA	0,91084	1 / 4s	s=0,274	14/17	16/17	16/17
TPOX	3,41871	1 / s	s=0,293	11	11	11
D18S51	6,55844	1 / p	p=0,152	16	16	16
D5S818	0,67624	1 / 4r	r=0,37	10/12	12/13	12/13
FGA	2,26183	1 / 4u	u=0,111	20/25	23/25	23/25
Amelogenina				X	XY	
CMLR^(****)	651834					
MP%^(*****)	99,9998					

(¹) **LR**: Índice¹ de Maternidade ou Parentesco Genético; razão de verossimilhança² ou probabilidade relativa da evidência genética estabelecida; valor numérico relativo aos cenários (hipóteses) de parentesco genético principal e alternativo, calculado para cada sistema alélico tipado de acordo com o algoritmo algébrico descrito por Brenner (Brenner *et al.*, 1994).

(^{**}) Frequências alélicas estimadas para a população brasileira, conforme descrito para o sistema *AmpFiSTR*[™] (Grattapaglia *et al.*, 2001).

(^{***}) **A**lelos **M**aternos **O**brigatórios (**AMO**), consistentes com os perfis alélicos determinados, que foram identificados utilizando o Algoritmo Algébrico Simbólico de Parentesco Genético.

(^{****}) **CMLR**: Índice Cumulativo de Maternidade ou Parentesco Genético; valor cumulativo dos índices de probabilidade relativa (**LR**) da evidência genética, calculado para o conjunto de sistemas alélicos estabelecidos para cada indivíduo tipado utilizando a equação $CLR = LR1 \times LR2 \times \dots \times LRn$, para os sistemas alélicos 1, 2...n. Quanto maior o índice cumulativo, mais forte é a evidência genética estabelecida a favor da hipótese principal. **CLR** é verificado por meio da razão entre a probabilidade dos alelos pertencentes ao perfil genético do(a) filho(a) e que estejam em coincidência com o perfil alélico do suposto pai e/ou de supostos ascendentes paternos serem oriundos desse suposto pai e/ou da linhagem ascendente masculina, e a probabilidade deles serem oriundos de qualquer indivíduo (não genotipado) da população.

(^{*****}) **MP%**: Probabilidade Relativa de Parentesco Genético ou Probabilidade Relativa de Maternidade, calculada de acordo com o Teorema de Bayes a partir do Índice de Maternidade utilizando a equação $MP\% = [CLR/(CLR + 1)] \times 100$; representa a probabilidade de a suposta mãe ser a mãe biológica, progenitora ou ascendente feminino.

1 Símbolo numérico que se associa a outro para caracterizar um novo símbolo.

2 Verossimilhança: Número positivo e menor que a unidade, que se associa a um evento aleatório, e que se mede pela frequência relativa da sua ocorrência numa longa sucessão de eventos.

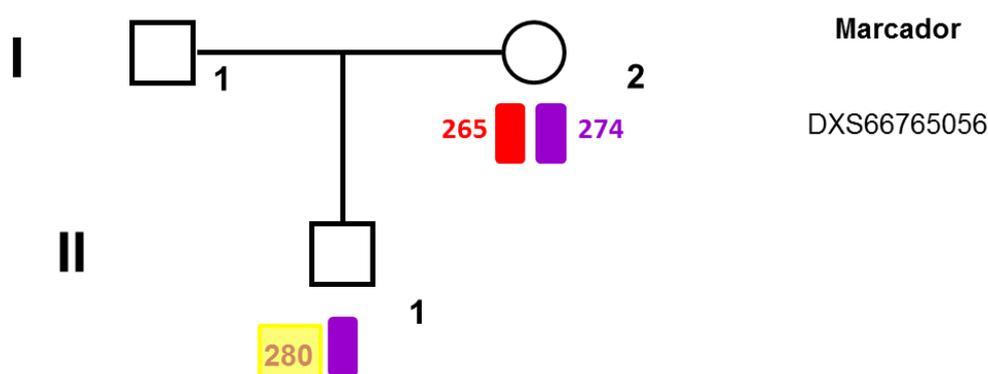


Figura 21. Heredograma do núcleo familiar 226 com uma expansão no alelo 274/280 no gene AR. Os indivíduos foram tipados com o marcador DXS66765056. Os quadrados representam os homens e os círculos, as mulheres. Nomes dos alelos são o comprimento do produto em pares de bases. A mãe (indivíduo I:2) possui 32 repetições CAG no *locus* do gene *AR*, e o filho (indivíduo II:1), 34 repetições CAG, ou seja, apresentou uma expansão de duas repetições CAG, mas que ainda não é um número de repetições que exceda o limiar de normalidade causador da doença de Kennedy. Em amarelo está sendo destacado o alelo que sofreu expansão.

5. DISCUSSÃO

Os critérios utilizados na varredura do exoma humano e em regiões reguladoras 5' permitiram identificar 74 *loci* candidatos. Destes, apenas quinze *loci* estão relacionados a doenças monogênicas causadas por expansão de repetições trinucleotídicas. Para os 59 *loci* restantes não foram encontrados estudos que os associem a doenças causadas pela expansão de repetições trinucleotídicas ou são *loci* em genes vinculados a doenças multifatoriais.

Neste estudo foram desenvolvidos ensaios moleculares da QF-PCR e TP-PCR para avaliação rápida das repetições trinucleotídicas associadas com as doenças de Machado Joseph e Kennedy, e a síndrome do X-frágil.

Para a técnica de QF-PCR foram utilizados apenas os iniciadores gene-específicos, sendo um deles marcado com fluóforo. Este ensaio é específico para o *locus* analisado, no entanto a técnica não permite a enumeração do número de repetições, sendo apenas uma estimativa do comprimento da repetição (Tassone *et al.*, 2008; Addis *et al.*, 2012). A QF-PCR foi recentemente validada como um método eficiente para estimação do número de cópias das repetições no gene *HTT*, relacionado à Doença de Huntington, detectando pré-mutações e mutações (Agostinho *et al.*, 2012).

Os ensaios da TP-PCR, que incluíram os iniciadores repetição-específicos em associação a um ou ambos os iniciadores gene-específicos, foram eficientes para detecção de alelos com grandes expansões (Warner *et al.*, 1996; Ciotti *et al.*, 2004; Jama *et al.*, 2013). Na versão *triprimer*, o ensaio TP-PCR só determina expansões a partir do iniciador 5' que é marcado com um fluóforo, para a detecção dos produtos de amplificação (Kakourou *et al.*, 2010; Xuncla *et al.*, 2010). Já a versão *tetraprimer* (Kakourou *et al.*, 2010; Lyon *et al.*, 2010; Addis *et al.*, 2012; Teo *et al.*, 2012; Jama *et al.*, 2013) além de detectar grandes alelos, permite a enumeração das cópias trinucleotídicas em uma única reação, pois refere as expansões a partir de ambos os iniciadores externos à repetição (Teo *et al.*, 2012).

Pelos testes realizados neste trabalho, comparando ambas as técnicas de QF e TP-PCR para as doenças de Machado-Joseph e Kennedy pode-se deduzir que, utilizando a técnica de QF-PCR que apenas infere o número de repetições do *locus*, o erro de cálculo na enumeração das repetições é provável que seja de apenas uma a duas repetições para mais ou para menos que quando utilizando a técnica de TP-

PCR, que mostra o número exato de repetições. No entanto para validação de ambas as técnicas, é necessário realizar o sequenciamento dos *loci* para determinar o número exato de repetições.

Destacamos o estudo do microssatélite (CAG)_n no gene *ATXN3*, relacionado com a Doença de Machado-Joseph, por ser a ataxia mais frequente no mundo e no Brasil, podendo chegar a 30% de todos os casos de ataxia no mundo (Schols *et al.*, 2004; Emmel, 2007; Freund *et al.*, 2009) e até a 92% no Brasil (Bettencourt *et al.*, 2011). A taxa de heterozigose observada para o marcador D14S92537354 relacionado a esta doença foi de 0,85, similar à encontrada na literatura (Freund *et al.*, 2009). Foram encontrados 21 alelos diferentes, variando entre 6 a 32 repetições. Todas as repetições encontradas caracterizam alelos normais. O ensaio para esta doença mostrou-se eficiente em determinar o número de cópias do *locus*.

A síndrome do X-frágil é a maior causa não sindrômica de retardo mental ligado ao cromossomo X (Felix-Lopez *et al.*, 2006), por isso a importância de ter na região um estudo sobre essa doença. Na genotipagem do marcador DXS146993454, desenhado para o gene *FMR1*, foram encontrados 14 alelos diferentes, variando de 4 a 28 repetições, sendo os alelos com 22 e 23 repetições CGG os mais frequentes.

A taxa de heterozigose para este marcador foi 0,29, considerada baixa quando comparada aos dados da literatura, que apresentam heterozigose entre 0,45 e 0,76 (Peixoto *et al.*, 1998; Arrieta *et al.*, 2003; Felix-Lopez *et al.*, 2006). Essa discrepância pode ser explicada pela possível presença de alelos com números de pares de base superior a 500pb nestes indivíduos, ou seja, *loci* com número de repetições superior a 102, que é o limite de análise utilizando o marcador de peso molecular LIZ500™, utilizado neste estudo. Ou seja, não se pode concluir que na população testada há um grande número de indivíduos homozigotos para o *locus* de *FMR1* ou que os indivíduos homozigotos, possuem outro alelo que não pôde ser detectado pela metodologia utilizada, e na realidade seriam heterozigotos.

Pela tipagem do marcador DXS66765056 relacionado ao gene *AR*, foram encontrados 15 alelos, e uma taxa de heterozigose de 0,84, similar à frequência observada na literatura, que é de 0,89 (Desmarais *et al.*, 1998). Dos 200 alelos tipados, 75 apresentaram entre de 32 a 39 repetições CAG. Esses 75 alelos representam 64 indivíduos com pelo menos um alelo apresentando pré-mutação.

Além disso, o alelo mais frequente encontrado foi o alelo com 31 repetições, que é um número de repetições normal, mas que representa um alelo longo. A tipagem por *iPCR* da sequência do gene *AR* apresenta um alelo com pré-mutação, com 33 repetições, que condiz com a grande ocorrência de indivíduos genotipados apresentando pelo menos um alelo com pré-mutação.

Este grande número de alelos com pré-mutação na população normal pode ser um indício que, por serem alelos instáveis, evolutivamente nas próximas gerações a frequência da Doença de Kennedy na população poderá aumentar, decorrente do processo de expansão de CAG no *locus* do gene *AR* (Mirkin, 2006).

Os 64 indivíduos que apresentaram pré-mutações, tiveram seus núcleos familiares investigados em busca de possíveis expansões transgeracionais. Desta busca, apenas uma família apresentou expansão no *locus*. Esta família apresentou a expansão de duas repetições, mas não patogênica, passando de 32 para 34 repetições CAG.

Mesmo que nos outros 63 núcleos familiares nenhuma expansão transgeracional tenha sido encontrada não significa que o alelo é estável, podendo apresentar indivíduos acometidos nas próximas gerações. No entanto, esses dados mostram que apesar da frequência de pré-mutação ser alta, não foram encontrados alelos expandidos, o que poderia explicar a baixa prevalência da doença de Kennedy na população, de apenas 1 a cada 50000 homens nascidos vivos (Kaimen-Maciel *et al.*, 1998).

O gene *AR* é o mais utilizado na literatura como marcador do padrão de inativação do cromossomo X, por apresentar elevada taxa de heterozigose (Allen *et al.*, 1992; Iitsuka *et al.*, 2001; Indsto *et al.*, 2001; Sato *et al.*, 2003; Thouin *et al.*, 2003; Gentilini *et al.*, 2012; Jobanputra *et al.*, 2012; Manoukian *et al.*, 2013; Mossner *et al.*, 2013). No entanto, a alta frequência de indivíduos com pré-mutação explica porque o marcador para o gene *AR* não deve ser utilizado para aplicações forenses, pois ele não é um marcador neutro (Szibor *et al.*, 2005). Além disso, esse marcador é utilizado também em amplos estudos de associação, pois o número de repetições CAG neste *locus* está relacionado a aumento de risco de diversas patologias humanas. Dentre várias outras doenças, o polimorfismo neste gene está relacionado ao aparecimento de câncer de mama e à síndrome do ovário policístico (Lin *et al.*, 2012; Gottlieb *et al.*, 2013).

A caracterização dos diferentes marcadores moleculares relacionados às doenças monogênicas causadas por expansão de repetições trinucleotídicas conduziu ao desenvolvimento de testes sensíveis e específicos para a análise direta das doenças utilizando como ferramenta a PCR. Como os sintomas dessas doenças se sobrepõem e há uma intensa variabilidade fenotípica, esta ferramenta molecular se tornou útil para o diagnóstico diferencial e confirmatório, bem como para a testagem preditiva. A identificação correta da patologia torna possível a aplicação de terapias apropriadas e de um aconselhamento genético adequado (Emmel, 2007). Este estudo também tem uma grande importância no Brasil, pois no país não existem estudos publicados em periódicos internacionais sobre a técnica de TP-PCR.

Este trabalho permitiu o desenvolvimento de marcadores para o diagnóstico molecular pré- e pós-sintomático das doenças de Machado-Joseph e Kennedy e para a síndrome do X-frágil de maneira rápida e acurada. Além disso, esse estudo trouxe uma importante contribuição para a melhoria do diagnóstico destas doenças na região, além de facilitar e possibilitar o aconselhamento genético a núcleos familiares com acometidos.

6. CONCLUSÕES

- A abordagem *in silico* mostrou que 59 *loci* de repetições trinucleotídicas possuem perfis semelhantes aos quinze *loci* relacionados às doenças causadas por expansão de trinucleotídeos. Contudo, para esses 59 *loci* não se conhece o envolvimento de expansões instáveis na etiologia das doenças genéticas.
- A associação de ambas as técnicas de QF-PCR e TP-PCR aumenta o sinal de amplificação dos perfis alélicos gerados além de aumentar o poder de resolução dos resultados.
- O desenvolvimento de marcadores para as doenças de Machado-Joseph e Kennedy e para a síndrome do X-frágil possibilitou a criação de testes que permitem caracterizar de maneira rápida, essas doenças.
- As taxas de heterozigose observadas para as doenças de Machado-Joseph e Kennedy foram semelhantes às encontradas em outras populações. No entanto, para a síndrome do X-frágil a taxa encontrada difere da literatura possivelmente devido a diferenças metodológicas utilizadas no presente estudo.
- Um alto índice de alelos com pré-mutação foi encontrado pela tipagem do marcador para o gene *AR*, relacionado à doença de Kennedy.

7. REFERÊNCIAS BIBLIOGRÁFICAS

Abrams, L., Cronister, A., Brown, W. T., Tassone, F., Sherman, S. L., Finucane, B., McConkie-Rosell, A., Hagerman, R., Kaufmann, W. E., Picker, J., Coffey, S., Skinner, D., Johnson, V., Miller, R. e Berry-Kravis, E. Newborn, carrier, and early childhood screening recommendations for fragile X. *Pediatrics*, v.130, n.6, Dec, p.1126-35. 2012.

Addis, M., Serrenti, M., Meloni, C., Cau, M. e Melis, M. A. Triplet-primed PCR is more sensitive than southern blotting-long PCR for the diagnosis of myotonic dystrophy type1. *Genet Test Mol Biomarkers*, v.16, n.12, Dec, p.1428-31. 2012.

Agostinho, L. A., Rocha, C. F., Medina-Acosta, E., Barboza, H. N., da Silva, A. F., Pereira, S. P., da Silva Idos, S., Paradela, E. R., Figueiredo, A. L., Nogueira Ede, M., Alvarenga, R. M., Hernan Cabello, P., dos Santos, S. R. e Paiva, C. L. Haplotype analysis of the CAG and CCG repeats in 21 Brazilian families with Huntington's disease. *J Hum Genet*, v.57, n.12, Dec, p.796-803. 2012.

Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M. e Belmont, J. W. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet*, v.51, n.6, Dec, p.1229-39. 1992.

Alonso, M. E., Ochoa, A., Sosa, A. L., Rodriguez, Y., Chavez, M., Boll, C., Yescas, P., Macias, R. e Rasmussen, A. Presymptomatic diagnosis in Huntington's disease: the Mexican experience. *Genet Test Mol Biomarkers*, v.13, n.6, Dec, p.717-20. 2009.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. e Lipman, D. J. Basic local alignment search tool. *J Mol Biol*, v.215, n.3, Oct 5, p.403-10. 1990.

Alves da Silva, A. F. *Recombinação em 21q dificulta a determinação da origem meiótica da não disjunção em portadores da Síndrome de Down*. (Monografia de Conclusão de Curso (Ciências Biológicas com ênfase em Biotecnologia)). Laboratório de Biotecnologia, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, 2007. 63 p.

Arrieta, I., Penagarikano, O., Telez, M., Ortega, B., Flores, P., Criado, B., Veiga, I., Peixoto, A. L. e Lostao, C. M. The FMR1 CGG repeat and linked microsatellite markers in two Basque valleys. *Heredity (Edinb)*, v.90, n.3, Mar, p.206-11. 2003.

Bagni, C., Tassone, F., Neri, G. e Hagerman, R. Fragile X syndrome: causes, diagnosis, mechanisms, and therapeutics. *J Clin Invest*, v.122, n.12, Dec 3, p.4314-22. 2012.

Bahl, S., Viridi, K., Mittal, U., Sachdeva, M. P., Kalla, A. K., Holmes, S. E., O'Hearn, E., Margolis, R. L., Jain, S., Srivastava, A. K. e Mukerji, M. Evidence of a common founder for SCA12 in the Indian population. *Ann Hum Genet*, v.69, n.Pt 5, Sep, p.528-34. 2005.

Bensaid, M., Melko, M., Bechara, E. G., Davidovic, L., Berretta, A., Catania, M. V., Gecz, J., Lalli, E. e Bardoni, B. FRAXE-associated mental retardation protein (FMR2) is an RNA-binding protein with high affinity for G-quartet RNA forming structure. Nucleic Acids Res, v.37, n.4, Mar, p.1269-79. 2009.

Bettencourt, C. e Lima, M. Machado-Joseph Disease: from first descriptions to new perspectives. Orphanet J Rare Dis, v.6, p.35. 2011.

Bhide, P. G., Day, M., Sapp, E., Schwarz, C., Sheth, A., Kim, J., Young, A. B., Penney, J., Golden, J., Aronin, N. e DiFiglia, M. Expression of normal and mutant huntingtin in the developing brain. J Neurosci, v.16, n.17, Sep 1, p.5523-35. 1996.

Brenner, C. H., Rittner, C. e Schneider, P. M. Calculation of paternity probabilities from multilocus DNA profiles. Electrophoresis, v.15, n.2, p.170-174. 1994.

Bruni, A. C., Takahashi-Fujigasaki, J., Maltecca, F., Foncin, J. F., Servadio, A., Casari, G., D'Adamo, P., Maletta, R., Curcio, S. A., De Michele, G., Filla, A., El Hachimi, K. H. e Duyckaerts, C. Behavioral disorder, dementia, ataxia, and rigidity in a large family with TATA box-binding protein mutation. Arch Neurol, v.61, n.8, Aug, p.1314-20. 2004.

Burk, K., Abele, M., Fetter, M., Dichgans, J., Skalej, M., Laccone, F., Didierjean, O., Brice, A. e Klockgether, T. Autosomal dominant cerebellar ataxia type I clinical features and MRI in families with SCA1, SCA2 and SCA3. Brain, v.119 (Pt 5), Oct, p.1497-505. 1996.

Cagnoli, C., Michielotto, C., Matsuura, T., Ashizawa, T., Margolis, R. L., Holmes, S. E., Gellera, C., Migone, N. e Brusco, A. Detection of large pathogenic expansions in FRDA1, SCA10, and SCA12 genes using a simple fluorescent repeat-primed PCR assay. J Mol Diagn, v.6, n.2, May, p.96-100. 2004.

Cagnoli, C., Stevanin, G., Michielotto, C., Gerbino Promis, G., Brussino, A., Pappi, P., Durr, A., Dragone, E., Viemont, M., Gellera, C., Brice, A., Migone, N. e Brusco, A. Large pathogenic expansions in the SCA2 and SCA7 genes can be detected by fluorescent repeat-primed polymerase chain reaction assay. J Mol Diagn, v.8, n.1, Feb, p.128-32. 2006.

Cancel, G., Durr, A., Didierjean, O., Imbert, G., Burk, K., Lezin, A., Belal, S., Benomar, A., Abada-Bendib, M., Vial, C., Guimaraes, J., Chneiweiss, H., Stevanin, G., Yvert, G., Abbas, N., Saudou, F., Lebre, A. S., Yahyaoui, M., Hentati, F., Vernant, J. C., *et al.* Molecular and clinical correlations in spinocerebellar ataxia 2: a study of 32 families. Hum Mol Genet, v.6, n.5, May, p.709-15. 1997.

Cannon, B., Pan, C., Chen, L., Hadd, A. G. e Russell, R. A dual-mode single-molecule fluorescence assay for the detection of expanded CGG repeats in Fragile X syndrome. Mol Biotechnol, v.53, n.1, Jan, p.19-28. 2012.

Catalli, C., Morgante, A., Iraci, R., Rinaldi, F., Botta, A. e Novelli, G. Validation of sensitivity and specificity of tetraplet-primed PCR (TP-PCR) in the molecular diagnosis of myotonic dystrophy type 2 (DM2). J Mol Diagn, v.12, n.5, Sep, p.601-6. 2010.

Cavagnari, B. M. [Gene therapy: a therapeutic option for neoplasias, infections and monogenic diseases]. Arch Argent Pediatr, v.109, n.4, Aug, p.326-32. 2011.

Cellini, E., Forleo, P., Nacmias, B., Tedde, A., Latorraca, S., Piacentini, S., Parnetti, L., Gallai, V. e Sorbi, S. Clinical and genetic analysis of hereditary and sporadic ataxia in central Italy. Brain Res Bull, v.56, n.3-4, Oct-Nov 1, p.363-6. 2001.

Chen, L., Hadd, A., Sah, S., Filipovic-Sadic, S., Krosting, J., Sekinger, E., Pan, R., Hagerman, P. J., Stenzel, T. T., Tassone, F. e Latham, G. J. An information-rich CGG repeat primed PCR that detects the full range of fragile X expanded alleles and minimizes the need for southern blot analysis. J Mol Diagn, v.12, n.5, Sep, p.589-600. 2010.

Chen, P., Ma, M., Shang, H., Su, D., Zhang, S. e Yang, Y. [Copy number variation of trinucleotide repeat in dynamic mutation sites of autosomal dominant cerebellar ataxias related genes]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi, v.26, n.6, Dec, p.626-33. 2009.

Ciotti, P., Di Maria, E., Bellone, E., Ajmar, F. e Mandich, P. Triplet repeat primed PCR (TP PCR) in molecular diagnostic testing for Friedreich ataxia. J Mol Diagn, v.6, n.4, Nov, p.285-9. 2004.

Cirillo, D., Agostini, F., Klus, P., Marchese, D., Rodriguez, S., Bolognesi, B. e Tartaglia, G. G. Neurodegenerative diseases: Quantitative predictions of protein-RNA interactions. Rna, v.19, n.2, Feb, p.129-40. 2012.

Cleary, J. D., Nichol, K., Wang, Y. H. e Pearson, C. E. Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. Nat Genet, v.31, n.1, May, p.37-46. 2002.

Cleary, J. D. e Pearson, C. E. The contribution of cis-elements to disease-associated repeat instability: clinical and experimental evidence. Cytogenet Genome Res, v.100, n.1-4, p.25-55. 2003.

_____. Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. Trends Genet, v.21, n.5, May, p.272-80. 2005.

Cummings, C. J. e Zoghbi, H. Y. Trinucleotide repeats: mechanisms and pathophysiology. Annu Rev Genomics Hum Genet, v.1, p.281-328. 2000.

Di Prospero, N. A. e Fischbeck, K. H. Therapeutics development for triplet repeat expansion diseases. Nat Rev Genet, v.6, n.10, Oct, p.756-65. 2005.

Dorsman, J. C., Smoor, M. A., Maat-Schieman, M. L., Bout, M., Siesling, S., van Duinen, S. G., Verschuuren, J. J., den Dunnen, J. T., Roos, R. A. e van Ommen, G. J. Analysis of the subcellular localization of huntingtin with a set of rabbit polyclonal antibodies in cultured mammalian cells of neuronal origin: comparison with the distribution of huntingtin in Huntington's disease autopsy brain. Philos Trans R Soc Lond B Biol Sci, v.354, n.1386, Jun 29, p.1061-7. 1999.

Duenas, A. M., Goold, R. e Giunti, P. Molecular pathogenesis of spinocerebellar ataxias. Brain, v.129, n.Pt 6, Jun, p.1357-70. 2006.

Emmel, V. E. Análise de repetições CAG nos genes SCA1, SCA2, SCA3 e SCA6 em pacientes com suspeita clínica de ataxia espinocerebelar. (Dissertação de Mestrado). Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007. 88 p.

Falk, M., Vojtiskova, M., Lukas, Z., Kroupova, I. e Froster, U. Simple procedure for automatic detection of unstable alleles in the myotonic dystrophy and Huntington's disease loci. Genet Test, v.10, n.2, Summer, p.85-97. 2006.

Felix-Lopez, X. A., Arguello-Garcia, R., Cerda-Flores, R. M., Penaloza-Espinoza, R. I., Buentello-Malo, L., Estrada-Mena, F. J., Ramos-Kuri, M., Gomez, F. S. e Arenas-Aranda, D. J. FMR1 CGG repeat distribution and linked microsatellite-SNP haplotypes in normal Mexican Mestizo and indigenous populations. Hum Biol, v.78, n.5, Oct, p.579-98. 2006.

Frackman, S., Kobs, G., Simpson, D. e Storts, D. Betaine and DMSO: Enhancing Agents for PCR. Promega Notes. Promega Corporation., v.65, p.27-30. 1998.

Freund, A. A., Scola, R. H., Teive, H. A., Arndt, R. C., da Costa-Ribeiro, M. C., Alle, L. F. e Werneck, L. C. Spinocerebellar ataxias: microsatellite and allele frequency in unaffected and affected individuals. Arq Neuropsiquiatr, v.67, n.4, Dec, p.1124-32. 2009.

Gao, R., Matsuura, T., Coolbaugh, M., Zuhlke, C., Nakamura, K., Rasmussen, A., Siciliano, M. J., Ashizawa, T. e Lin, X. Instability of expanded CAG/CAA repeats in spinocerebellar ataxia type 17. Eur J Hum Genet, v.16, n.2, Feb, p.215-22. 2008.

Gatchel, J. R. e Zoghbi, H. Y. Diseases of unstable repeat expansion: mechanisms and common principles. Nat Rev Genet, v.6, n.10, Oct, p.743-55. 2005.

Gecz, J. The FMR2 gene, FRAXE and non-specific X-linked mental retardation: clinical and molecular aspects. Ann Hum Genet, v.64, n.Pt 2, Mar, p.95-106. 2000.

Gelfand, Y., Rodriguez, A. e Benson, G. TRDB--the Tandem Repeats Database. Nucleic Acids Res, v.35, n.Database issue, Jan, p.D80-7. 2007.

Gentilini, D., Castaldi, D., Mari, D., Monti, D., Franceschi, C., Di Blasio, A. M. e Vitale, G. Age-dependent skewing of X chromosome inactivation appears delayed in centenarians' offspring. Is there a role for allelic imbalance in healthy aging and longevity? Aging Cell, v.11, n.2, Apr, p.277-83. 2012.

Geschwind, D. H., Perlman, S., Figueroa, C. P., Treiman, L. J. e Pulst, S. M. The prevalence and wide clinical spectrum of the spinocerebellar ataxia type 2 trinucleotide repeat in patients with autosomal dominant cerebellar ataxia. Am J Hum Genet, v.60, n.4, Apr, p.842-50. 1997.

Gispert, S., Twells, R., Orozco, G., Brice, A., Weber, J., Heredero, L., Scheufler, K., Riley, B., Allotey, R., Nothers, C., Hillermann, R., Lunkes, A., Khati, C., Stevanin, G.,

Hernandez, A., Magariño, C., Klockgether, T., Durr, A., Chneiweiss, H., Enczmann, J., *et al.* Chromosomal assignment of the second locus for autosomal dominant cerebellar ataxia (SCA2) to chromosome 12q23-24.1. Nature Genetics, v.4, p.295-299. 1993.

Gonzalez-Gonzalez, C., Garcia-Hoyos, M., Trujillo-Tiebas, M. J., Lorda-Sanchez, I., de Alba, M. R., Infantes, F., Gallego, J., Diaz-Recasens, J., Ayuso, C. e Ramos, C. Application of fetal DNA detection in maternal plasma: a prenatal diagnosis unit experience. J Histochem Cytochem, v.53, n.3, Mar, p.307-14. 2005.

Gottlieb, B., Alvarado, C., Wang, C., Gharizadeh, B., Babrzadeh, F., Richards, B., Batist, G., Basik, M., Beitel, L. K. e Trifiro, M. Making Sense of Intratumor Genetic Heterogeneity: Altered Frequency of Androgen Receptor CAG Repeat Length Variants in Breast Cancer Tissues. Hum Mutat, Feb 1. 2013.

Grattapaglia, D., Schmidt, A. B., Costa e Silva, C., Stringher, C., Fernandes, A. P. e Ferreira, M. E. Brazilian population database for the 13 STR loci of the AmpFISTR Profiler Plus and Cofiler multiplex kits. Forensic Sci Int, v.118, n.1, Apr 15, p.91-4. 2001.

Gu, Y., Shen, Y., Gibbs, R. A. e Nelson, D. L. Identification of FMR2, a novel gene associated with the FRAXE CCG repeat and CpG island. Nat Genet, v.13, n.1, May, p.109-13. 1996.

Hannan, A. J. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'. Trends Genet, v.26, n.2, Feb, p.59-65. 2010.

Holmes, S. E., O'Hearn, E. E., McInnis, M. G., Gorelick-Feldman, D. A., Kleiderlein, J. J., Callahan, C., Kwak, N. G., Ingersoll-Ashworth, R. G., Sherr, M., Sumner, A. J., Sharp, A. H., Ananth, U., Seltzer, W. K., Boss, M. A., Vieria-Saecker, A. M., Epplen, J. T., Riess, O., Ross, C. A. e Margolis, R. L. Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. Nat Genet, v.23, n.4, Dec, p.391-2. 1999.

Hori, T., Tsuji, S. e Yamauchi, M. [Dynamic mutation and its molecular cytogenetics]. Nihon Rinsho, v.57, n.4, Apr, p.771-7. 1999.

Hurles, M. How homologous recombination generates a mutable genome. Hum Genomics, v.2, p.179-186. 2005.

Huynh, D. P., Del Bigio, M. R., Ho, D. H. e Pulst, S. M. Expression of ataxin-2 in brains from normal individuals and patients with Alzheimer's disease and spinocerebellar ataxia 2. Ann Neurol, v.45, n.2, Feb, p.232-41. 1999.

Iitsuka, Y., Bock, A., Nguyen, D. D., Samango-Sprouse, C. A., Simpson, J. L. e Bischoff, F. Z. Evidence of skewed X-chromosome inactivation in 47,XXY and 48,XXYY Klinefelter patients. Am J Med Genet, v.98, n.1, Jan 1, p.25-31. 2001.

Indsto, J. O., Cachia, A. R., Kefford, R. F. e Mann, G. J. X inactivation, DNA deletion, and microsatellite instability in common acquired melanocytic nevi. Clin Cancer Res, v.7, n.12, Dec, p.4054-9. 2001.

Jama, M., Millson, A., Miller, C. E. e Lyon, E. Triplet repeat primed PCR simplifies testing for huntington disease. J Mol Diagn, v.15, n.2, Mar, p.255-62. 2013.

Jobanputra, V., Levy, B., Kinney, A., Brown, S., Shirazi, M., Yu, C., Kline, J. e Warburton, D. Copy number changes on the X chromosome in women with and without highly skewed X-chromosome inactivation. Cytogenet Genome Res, v.136, n.4, p.264-9. 2012.

Kaimen-Maciel, D. R., Medeiros, M., Clímaco, V., Kelian, G. R., Silva, L. S. T., Souza, M. M. e Raskin, S. Atrofia Muscular Bulbo Espinhal Recessiva Ligada ao Cromossomo X (Doença de Kennedy). Arq Neuropsiquiatr, v.56, n.3-B, p.639-645. 1998.

Kakourou, G., Dhanjal, S., Mamas, T., Serhal, P., Delhanty, J. D. e SenGupta, S. B. Modification of the triplet repeat primed polymerase chain reaction method for detection of the CTG repeat expansion in myotonic dystrophy type 1: application in preimplantation genetic diagnosis. Fertil Steril, v.94, n.5, Oct, p.1674-9. 2010.

Kashi, Y. e King, D. G. Simple sequence repeats as advantageous mutators in evolution. Trends Genet, v.22, p.253-259. 2006.

Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., Kawakami, H., Nakamura, S., Nishimura, M., Akiguchi, I. e et al. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. Nat Genet, v.8, n.3, Nov, p.221-8. 1994.

Kennedy, L., Evans, E., Chen, C. M., Craven, L., Detloff, P. J., Ennis, M. e Shelbourne, P. F. Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. Hum Mol Genet, v.12, n.24, Dec 15, p.3359-67. 2003.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. e Haussler, D. The human genome browser at UCSC. Genome Res, v.12, n.6, Jun, p.996-1006. 2002.

Kibbe, W. A. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res, v.35, n.Web Server issue, Jul, p.W43-6. 2007.

King, M. W. The Trinucleotide Repeat Disorders: IU School of Medicine / miking at iupui.edu 2011.

Kozlowski, P., Sobczak, K. e Krzyzosiak, W. J. Trinucleotide repeats: triggers for genomic disorders? Genome Med, v.2, n.4, p.29. 2010.

Kunkel, T. A. Nucleotide repeats. Slippery DNA and diseases. Nature, v.365, n.6443, Sep 16, p.207-8. 1993.

La Spada, A. R. Trinucleotide repeat instability: genetic features and molecular mechanisms. Brain Pathol, v.7, n.3, Jul, p.943-63. 1997.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., *et al.* Initial sequencing and analysis of the human genome. Nature, v.409, n.6822, Feb 15, p.860-921. 2001.

Landles, C. e Bates, G. P. Huntingtin and the molecular pathogenesis of Huntington's disease. Fourth in molecular medicine review series. EMBO Rep, v.5, n.10, Oct, p.958-63. 2004.

Lastres-Becker, I., Rub, U. e Auburger, G. Spinocerebellar ataxia 2 (SCA2). Cerebellum, v.7, n.2, p.115-24. 2008.

Levy, S., Sutton, G., Ng, P. C., Feuk, L., Halpern, A. L., Walenz, B. P., Axelrod, N., Huang, J., Kirkness, E. F., Denisov, G., Lin, Y., MacDonald, J. R., Pang, A. W., Shago, M., Stockwell, T. B., Tsiamouri, A., Bafna, V., Bansal, V., Kravitz, S. A., Busam, D. A., *et al.* The diploid genome sequence of an individual human. PLoS Biol, v.5, n.10, Sep 4, p.e254. 2007.

Lin, L. H., Baracat, M. C., Maciel, G. A., Soares, J. M., Jr. e Baracat, E. C. Androgen receptor gene polymorphism and polycystic ovary syndrome. Int J Gynaecol Obstet, v.120, n.2, Feb, p.115-8. 2012.

Lokkegaard, T., Nielsen, J. E., Hasholt, L., Fenger, K., Werdelin, L., Tranebjaerg, L., Lauritzen, M., Colding-Jorgensen, E., Gronbech-Jensen, M., Henriksen, O. A. e Sorensen, S. A. Machado-Joseph disease in three Scandinavian families. J Neurol Sci, v.156, n.2, Apr 1, p.152-7. 1998.

Lutz, R. E. Trinucleotide repeat disorders. Semin Pediatr Neurol, v.14, n.1, Mar, p.26-33. 2007.

Lyon, E., Laver, T., Yu, P., Jama, M., Young, K., Zoccoli, M. e Marlowe, N. A simple, high-throughput assay for Fragile X expanded alleles using triple repeat primed PCR and capillary electrophoresis. J Mol Diagn, v.12, n.4, Jul, p.505-11. 2010.

Manoukian, S., Verderio, P., Tabano, S., Colapietro, P., Pizzamiglio, S., Grati, F. R., Calvello, M., Peissel, B., Burn, J., Pensotti, V., Allemani, C., Sirchia, S. M., Radice, P. e Miozzo, M. X chromosome inactivation pattern in BRCA gene mutation carriers. Eur J Cancer, v.49, n.5, Mar, p.1136-41. 2013.

Mao, R., Aylsworth, A. S., Potter, N., Wilson, W. G., Breningstall, G., Wick, M. J., Babovic-Vuksanovic, D., Nance, M., Patterson, M. C., Gomez, C. M. e Snow, K. Childhood-onset ataxia: testing for large CAG-repeats in SCA2 and SCA7. Am J Med Genet, v.110, n.4, Jul 15, p.338-45. 2002.

Martins, W. S., Lucas, D. C., Neves, K. F. e Bertioli, D. J. WebSat--a web software for microsatellite marker development. Bioinformatics, v.3, n.6, p.282-3. 2009.

McMurray, C. T. Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet, v.11, n.11, Nov, p.786-99. 2010.

Miller, J. W., Urbinati, C. R., Teng-Umnuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A. e Swanson, M. S. Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. Embo J, v.19, n.17, Sep 1, p.4439-48. 2000.

Miller, W. J., Skinner, J. A., Foss, G. S. e Davies, K. E. Localization of the fragile X mental retardation 2 (FMR2) protein in mammalian brain. Eur J Neurosci, v.12, n.1, Jan, p.381-4. 2000.

Mirkin, S. M. DNA structures, repeat expansions and human hereditary disorders. Curr Opin Struct Biol, v.16, n.3, Jun, p.351-8. 2006.

_____. Expandable DNA repeats and human disease. Nature, v.447, n.7147, Jun 21, p.932-40. 2007.

Mor-Cohen, R., Magal, N., Gadoth, N., Shohat, T. e Shohat, M. Correlation between the incidence of myotonic dystrophy in different groups in Israel and the number of CTG trinucleotide repeats in the myotonin gene. Am J Med Genet, v.71, n.2, Aug 8, p.156-9. 1997.

Moreno, P. A., Velez, P. E., Martinez, E., Garreta, L. E., Diaz, N., Amador, S., Tischer, I., Gutierrez, J. M., Naik, A. K., Tobar, F. e Garcia, F. The human genome: a multifractal analysis. BMC Genomics, v.12, n.1, Oct 14, p.506. 2011.

Morfini, G., Pigino, G. e Brady, S. T. Polyglutamine expansion diseases: failing to deliver. Trends Mol Med, v.11, n.2, Feb, p.64-70. 2005.

Mossner, M., Nolte, F., Hutter, G., Reins, J., Klaumunzer, M., Nowak, V., Oblander, J., Ackermann, K., Will, S., Rohl, H., Neumann, U., Neumann, M., Hopfer, O., Baldus, C. D., Hofmann, W. K. e Nowak, D. Skewed X-inactivation patterns in ageing healthy and myelodysplastic haematopoiesis determined by a pyrosequencing based transcriptional clonality assay. J Med Genet, v.50, n.2, Feb, p.108-17. 2013.

Mutsuddi, M. e Rebay, I. Molecular genetics of spinocerebellar ataxia type 8 (SCA8). RNA Biol, v.2, n.2, Apr, p.49-52. 2005.

Nucifora, F. C., Jr., Ellerby, L. M., Wellington, C. L., Wood, J. D., Herring, W. J., Sawa, A., Hayden, M. R., Dawson, V. L., Dawson, T. M. e Ross, C. A. Nuclear localization of a non-caspase truncation product of atrophin-1, with an expanded polyglutamine repeat, increases cellular toxicity. J Biol Chem, v.278, n.15, Apr 11, p.13047-55. 2003.

Orr, H. T., Chung, M. Y., Banfi, S., Kwiatkowski, T. J., Jr., Servadio, A., Beaudet, A. L., McCall, A. E., Duvick, L. A., Ranum, L. P. e Zoghbi, H. Y. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nat Genet, v.4, n.3, Jul, p.221-6. 1993.

Orr, H. T. e Zoghbi, H. Y. Trinucleotide repeat disorders. Annu Rev Neurosci, v.30, p.575-621. 2007.

Pearson, C. E., Nichol Edamura, K. e Cleary, J. D. Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet, v.6, n.10, Oct, p.729-42. 2005.

Peixoto, A., dos Santos, M. R., Seruca, R., Amorim, A. e Castedo, S. Analysis of FMR1 and flanking microsatellite markers in normal and fragile X chromosomes in Portugal: evidence for a "protector" haplotype. Eur J Hum Genet, v.6, n.5, Sep-Oct, p.518-22. 1998.

Peltonen, L., Perola, M., Naukkarinen, J. e Palotie, A. Lessons from studying monogenic disease for common disease. Hum Mol Genet, v.15 Spec No 1, Apr 15, p.R67-74. 2006.

Radvansky, J., Ficek, A. e Kadasi, L. Upgrading molecular diagnostics of myotonic dystrophies: multiplexing for simultaneous characterization of the DMPK and ZNF9 repeat motifs. Mol Cell Probes, v.25, n.4, Aug, p.182-5. 2011.

Radvansky, J., Ficek, A., Minarik, G., Palffy, R. e Kadasi, L. Effect of unexpected sequence interruptions to conventional PCR and repeat primed PCR in myotonic dystrophy type 1 testing. Diagn Mol Pathol, v.20, n.1, Mar, p.48-51. 2011.

Restituto, S., Thompson, R. M., Eliet, J., Raiké, R. S., Riedl, M., Charnet, P. e Gomez, C. M. The polyglutamine expansion in spinocerebellar ataxia type 6 causes a beta subunit-specific enhanced activation of P/Q-type calcium channels in *Xenopus* oocytes. J Neurosci, v.20, n.17, Sep 1, p.6394-403. 2000.

Richards, R. I. Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. Hum Mol Genet, v.10, n.20, Oct 1, p.2187-94. 2001.

Richards, R. I. e Sutherland, G. R. Dynamic mutations: a new class of mutations causing human disease. Cell, v.70, n.5, Sep 4, p.709-12. 1992.

_____. Dynamic mutation: possible mechanisms and significance in human disease. Trends Biochem Sci, v.22, n.11, Nov, p.432-6. 1997.

Rosales-Reynoso, M. A., Ochoa-Hernandez, A. B. e Barros-Nunez, P. [Diseases caused by triplet expansion]. Rev Neurol, v.49, n.2, Jul 16-31, p.79-87. 2009.

Rosenberg, R. N. Machado-Joseph disease: an autosomal dominant motor system degeneration. Mov Disord, v.7, n.3, p.193-203. 1992.

Ross, C. A. e Poirier, M. A. Protein aggregation and neurodegenerative disease. Nat Med, v.10 Suppl, Jul, p.S10-7. 2004.

Sato, K., Hashiyada, M., Uehara, S., Nata, M. e Okamura, K. CpG dinucleotide methylation patterns in the human androgen receptor gene and X-chromosome inactivation in peripheral blood leukocytes of phenotypically normal women. J Hum Genet, v.48, n.7, p.374-9. 2003.

Schelhaas, H. J., Ippel, P. F., Beemer, F. A. e Hageman, G. Similarities and differences in the phenotype, genotype and pathogenesis of different spinocerebellar ataxias. Eur J Neurol, v.7, n.3, May, p.309-14. 2000.

Schmitt, I., Evert, B. O., Khazneh, H., Klockgether, T. e Wuellner, U. The human MJD gene: genomic structure and functional characterization of the promoter region. Gene, v.314, Sep 18, p.81-8. 2003.

Schols, L., Bauer, P., Schmidt, T., Schulte, T. e Riess, O. Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. Lancet Neurol, v.3, n.5, May, p.291-304. 2004.

Servadio, A., Koshy, B., Armstrong, D., Antalffy, B., Orr, H. T. e Zoghbi, H. Y. Expression analysis of the ataxin-1 protein in tissues from normal and spinocerebellar ataxia type 1 individuals. Nat Genet, v.10, n.1, May, p.94-8. 1995.

Stevanin, G., Herman, A., Durr, A., Jodice, C., Frontali, M., Agid, Y. e Brice, A. Are (CTG)_n expansions at the SCA8 locus rare polymorphisms? Nat Genet, v.24, n.3, Mar, p.213; author reply 215. 2000.

Szibor, R., Hering, S. e Edelmann, J. The HumARA genotype is linked to spinal and bulbar muscular dystrophy and some further disease risks and should no longer be used as a DNA marker for forensic purposes. Int J Legal Med, v.119, n.3, May, p.179-80. 2005.

Tassone, F., Pan, R., Amiri, K., Taylor, A. K. e Hagerman, P. J. A rapid polymerase chain reaction-based screening method for identification of all expanded alleles of the fragile X (FMR1) gene in newborn and high-risk populations. J Mol Diagn, v.10, n.1, Jan, p.43-9. 2008.

Teo, C. R., Law, H. Y., Lee, C. G. e Chong, S. S. Screening for CGG repeat expansion in the FMR1 gene by melting curve analysis of combined 5' and 3' direct triplet-primed PCRs. Clin Chem, v.58, n.3, Mar, p.568-79. 2012.

Thouin, M. M., Giron, J. M. e Hoffman, E. P. Detection of nonrandom X chromosome inactivation. Curr Protoc Hum Genet, v.Chapter 9, Feb, p.Unit9 7. 2003.

Thyagarajan, B., Bower, M., Berger, M., Jones, S., Dolan, M. e Wang, X. A novel polymorphism in the FMR1 gene: implications for clinical testing of fragile X syndrome. Arch Pathol Lab Med, v.132, n.1, Jan, p.95-8. 2008.

Todd, P. K. e Paulson, H. L. RNA-mediated neurodegeneration in repeat expansion disorders. Ann Neurol, v.67, n.3, Mar, p.291-300. 2010.

Trott, A., Jardim, L. B., Ludwig, H. T., Saute, J. A., Artigas, O., Kieling, C., Wanderley, H. Y., Rieder, C. R., Monte, T. L., Socal, M., Alonso, I., Ferro, A., Carvalho, T., do Ceu Moreira, M., Mendonca, P., Ferreirinha, F., Silveira, I., Sequeiros, J., Giugliani, R. e Saraiva-Pereira, M. L. Spinocerebellar ataxias in 114 Brazilian families: clinical and molecular findings. Clin Genet, v.70, n.2, Aug, p.173-6. 2006.

Trott, A., Maris, A. F. e Miranda, G. B. Spinocerebellar ataxias caused by expanded polyglutamine: a review. Rev. Neurocienc., v.18, n.4, p.512-522. 2010.

Vonsattel, J. P. e DiFiglia, M. Huntington disease. J Neuropathol Exp Neurol, v.57, n.5, May, p.369-84. 1998.

Warner, J. P., Barron, L. H., Goudie, D., Kelly, K., Dow, D., Fitzpatrick, D. R. e Brock, D. J. A general method for the detection of large CAG repeat expansions by fluorescent PCR. J Med Genet, v.33, n.12, Dec, p.1022-6. 1996.

Wheeler, V. C., Lebel, L. A., Vrbanac, V., Teed, A., te Riele, H. e MacDonald, M. E. Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. Hum Mol Genet, v.12, n.3, Feb 1, p.273-81. 2003.

Winnepenninckx, B., Debacker, K., Ramsay, J., Smeets, D., Smits, A., FitzPatrick, D. R. e Kooy, R. F. CGG-repeat expansion in the DIP2B gene is associated with the fragile site FRA12A on chromosome 12q13.1. Am J Hum Genet, v.80, n.2, Feb, p.221-31. 2007.

Xuncla, M., Rodriguez-Revenga, L., Madrigal, I., Jimenez, D., Mila, M. e Badenas, C. Protocol proposal for Friedreich ataxia molecular diagnosis using fluorescent and triplet repeat primed polymerase chain reaction. Transl Res, v.156, n.5, Nov, p.309-14. 2010.

Yrigollen, C., Durbin-Johnson, B., Gane, L., Nelson, D., Hagerman, R., Hagerman, P. e Tassone, F. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med, v.14, n.8, Aug, p.729-36. 2012.

Zhou, Y., Law, H. Y., Boehm, C. D., Yoon, C. S., Cutting, G. R., Ng, I. S. e Chong, S. S. Robust fragile X (CGG)_n genotype classification using a methylation specific triple PCR assay. J Med Genet, v.41, n.4, Apr, p.e45. 2004.

Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D. W., Amos, C., Dobyns, W. B., Subramony, S. H., Zoghbi, H. Y. e Lee, C. C. Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. Nat Genet, v.15, n.1, Jan, p.62-9. 1997.

Zoghbi, H. Y. Spinocerebellar ataxia type 1. Clin Neurosci, v.3, n.1, p.5-11. 1995.

_____. The expanding world of ataxins. Nat Genet, v.14, n.3, Nov, p.237-8. 1996.

ANEXO**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

Eu _____, tendo sido satisfatoriamente informado(a) sobre o estudo supracitado, e seus objetivos gerais, declaro por livre e espontânea vontade que concordo em participar do mesmo. Autorizo a utilização da amostra biológica de mim coletada apenas para este projeto, podendo o DNA extraído dessa amostra ser armazenado para utilização em novo projeto de pesquisa, para o qual deverei ser contatado(a) a fim de conceder ou não autorização para seu uso. Declaro estar ciente das seguintes medidas de proteção contra qualquer tipo de discriminação e/ou estigmatização individual ou coletiva, que serão utilizadas durante a pesquisa: a minha identidade será mantida sob sigilo absoluto; todas as dúvidas que porventura ocorram durante o processo desta pesquisa serão devidamente esclarecidas; não haverá nenhum tipo de constrangimento ou ação em contra, caso decida por revogar a autorização aqui concedida para utilização das informações obtidas a partir dos estudos da amostra biológica e dos dados, de mim coletados, a qualquer momento. Declaro estar ciente que terei a opção de tomar ou não conhecimento dos resultados da pesquisa na forma de parecer técnico individual quanto ao meu perfil genético determinado para os marcadores supracitados. Declaro ainda não ter nenhum tipo de interesse econômico, em qualquer época, sobre os possíveis resultados alcançados durante as etapas subsequentes do andamento deste estudo, deixando claro ser o meu interesse particular somente o de colaborador (a) anônimo (a).

Campos dos Goytacazes, RJ, ____ de _____ de 20__.

Assinatura do voluntário (ou mãe, pai e/ou responsável legal)

Enrique Medina-Acosta, MSc., PhD.

Coordenador do Núcleo de Diagnóstico e Investigação Molecular – NUDIM

Universidade Estadual do Norte Fluminense Darcy Ribeiro

Matrícula funcional nº. 0000549-6

Responsável pela pesquisa laboratorial