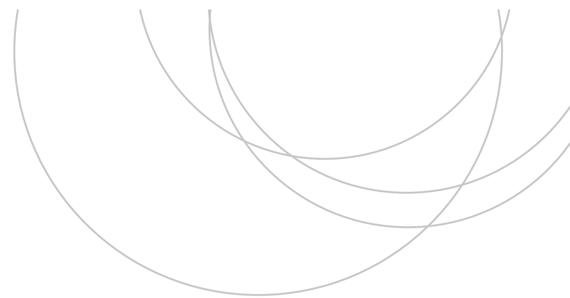
eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

ZIENTZIA ETA TEKNOLOGIA FAKULTATEA
FACULTAD DE CIENCIA Y TECNOLOGÍA



Gradu Amaierako Lana / Trabajo Fin de Grado
Biologia Gradua / Grado en Biología

A new method of molecular diagnosis for the Huntington's disease via PCR.

Egilea/Autor/a:
Jon Poza Villa
Zuzendaria/Director/a:
Ainhoa Iglesias Ara

© 2017, Jon Poza Villa

Table of contents

Abstract 3

Resumen 3

1. Introduction 4
- **1.1 Huntington disease** 5
- **1.2 Genetics of Huntington's disease** 5
- **1.3 Methods for HD molecular diagnosis** 7

2. Hypothesis and objectives 7

3. Materials and methods 8
- **3.1 Individuals** 8
- **3.2 Primers** 9
- **3.3 PCR reagents concentrations, thermal cycler conditions and migration** 10
- **3.4 Regression** 11

4. Results and Discussion 11
- **4.1 Optimization of the PCR with control DNA of non-affected individuals** 11
- **4.1.1 Quality control of control DNA** 11
- **4.1.1 Detection of the normal Htt alleles** 13
- **4.1.1.1 Optimization of the annealing temperature** 13
- **4.1.1.2 Effects of the treatment with DMSO** 15
- **4.1.1.3 The effects of the treatment with betaine** 16
- **4.2 Analysis of the control DNA of affected individuals and test DNA** 18
- **4.2.1 Quality control and minimum concentration of DNA for PCR** 18
- **4.2.2 Detection of the affected Htt alleles with the optimized methodology** 20
- **4.3 Molecular diagnosis of individual 11** 22
- **4.4 Relevant outcomes** 23

5. Conclusion 23

Acknowledgements 24

6. References 24

Abstract

Single gene disorders are the most studied genetic diseases due to its simple inheritance pattern (recessive or dominant). Huntington’s disease (HD) is a single gene disorder which appears between the third and fifth decade of life. It is a progressive neurodegenerative disorder, which leads to 15-20 years of life expectancy after the first symptoms appear. Due to the late onset of the disease, an early diagnosis can be crucial, that is why an easy and cheap method of diagnosis is important. The disease is caused by the number of (CAG)n repeats in the *huntingtin* (Htt) gene; people with over 36 CAG repeats can develop the disorder. Previous molecular diagnosis had been developed, but were expensive or unsuitable for routine use. In this study, we developed a new, cheap and easy method to detect Huntington´s disease via PCR (polymerase chain reaction). To develop the new method, we tested specific primers in positive and negative control samples to verify the effectiveness. We tested the best concentrations of the reagents and conditions for the PCR. Due to the high concentration of the bases C and G, secondary structures were formed in the DNA chain in the amplicon, we used betaine for the disruption of this secondary structures. To measure the exact number of CAG repeats we made a semilogarithmic regression and used its equation. Results showed that the developed method was effective for the detection of the Huntington’s disease in positive and negative controls. With the developed method, we were able to perform the molecular diagnosis of an unknown individual with chances of developing the disease. Taken together, we have developed a new method of molecular diagnosis for the Huntington´s disease via PCR.

Resumen

Los trastornos monogénicos son las enfermedades genéticas más estudiadas debido a su simple modelo de herencia (dominante o recesivo). La enfermedad de Huntington (HD) es un trastorno monogénico que aparece entre la tercera y quinta década de vida. Es una enfermedad progresiva y neurodegenerativa, que lleva a 15-20 años de esperanza de vida después de que los primeros síntomas aparezcan. Debido a la tardía expresión de la enfermedad, un diagnóstico temprano es crucial, es por ello que un diagnóstico barato y sencillo es tan

importante. La enfermedad depende de las repeticiones de (CAG)n en el gen de la *huntingtina* (Htt); las personas con más de 36 copias de CAG pueden desarrollar la enfermedad. Anteriormente se han desarrollado diagnósticos moleculares pero eran caros o inadecuados para el uso rutinario. En este estudio hemos desarrollado nuevo, barato y fácil método para detectar la enfermedad de Huntington vía PCR (reacción en cadena de la polimerasa). Para desarrollar este nuevo método, hemos testado unos primers específicos en controles positivos y negativos para verificar la efectividad. Hemos testado las mejores concentraciones de los reactivos y las condiciones de la PCR. Debido a la gran concentración de bases C y G, se formaron estructuras secundarias en la cadena de ADN del amplicón, utilizamos betaína para deshacer esas estructuras secundarias. Para medir el número exacto de repeticiones de CAG, hicimos una regresión semilogarítmica y usamos la ecuación de esta regresión. Los resultados mostraron que el método desarrollado para detectar la enfermedad de Huntington era efectivo, tanto en controles positivos como negativos. Con el método desarrollado, fuimos capaces de hacer el diagnóstico molecular de un individuo con probabilidades de desarrollar la enfermedad. Resumiendo, hemos creado un método nuevo de diagnóstico molecular para detectar la enfermedad de Huntington vía PCR.

1. Introduction

Genetic diseases are caused by an abnormality or mutation in the genome. These diseases are not contagious, but can be transmitted through the generations. Among these kinds of diseases, the most studied genetic diseases are single gen disorders. This is due to its simple inheritance pattern (recessive or dominant) and the relative simple genetic etiology. The advances in understanding the mechanisms behind this diseases eases new diagnostic test, treatments or inventions to prevent or minimize the disease. It is really important to find new cheap and easy ways of diagnosis, because, an early diagnosis and detection of the disorder, can improve the treatment of the diseased (Genetic alliance, 2010).

1.1 Huntington disease

Huntington's disease (HD), also known as Huntington's chorea, was discovered in 1972 by George Huntington. It is a progressive neurodegenerative disorder, which is hereditary and autosomal dominant. It affects the nervous system, mainly the regions of the brain in charge of the movement control. People with this condition usually have motor disturbance and abnormal body postures, changes in behaviour, emotion judgment and cognition. People with HD may also develop trouble speaking, feeding and swallowing (Roos, 2010).

HD is a rare neurological and psychiatric condition; its frequency can vary among populations. It appears in all the ethnic groups and it affects 5 to 10 people in 100,000 in the Caucasian population. Different phenocopies have been described with lower prevalence than Huntington's disease (Roos, 2010).

The age of manifestation of Huntington's disease can vary among the different individuals with the condition, but it usually appears during the third and fifth decade of life on men and women (Rosales-Reynoso & Barros-Nuñez, 2007). An earlier onset can also be seen, usually in people that have parents with this condition. The progressive increase of the disease leads to 15 to 20 years of life expectancy after the first symptoms appear, being pneumonia the first cause of death. (Vazquez-Cerdas et al., 2011).

1.2 Genetics of Huntington's disease

The gene affected in HD is located in the short arm of the 4th chromosome (4p16.3) (Roos, 2010). This gene is responsible of coding the sequence for the Huntingtin protein (Htt) and has a CAG (cytosine, adenine, guanine) trinucleotide repeat; CAG is the codon responsible of the polyglutamine (or polyQ) site of the protein. The disorder is caused by a polymorphism on the number of repeats of the trinucleotide (CAG)n, and is classified as a polyQ disease, triplet expansion or protein conformation disease. The alteration in the number of the (CAG)n repeats causes a disorder in the configuration of the Huntingtin protein, leading to a

dysfunction of the protein and the already mentioned neurological and psychiatric disorders (Tasset et al.,2009).

As mentioned before, the disorder is caused by the number of (CAG)n repeats. Normal alleles have between 9 and 36 repeats of the trinucleotide; people with these alleles will not develop the disorder and the number of trinucleotides will segregate without polymorphism in the 99% of the cases. People with 29 to 34 repeats of the polyQ sequence will not develop the disorder during their lifetime, but the genotype can be unstable in the germinal cells, increasing the number of repeats in the next generation and increasing the chance of developing the disorder (Rosales-Reynoso & Barros-Nuñez, 2007). Higher instability has been seen in the spermatogenesis, rather than in the ovogenesis, this is the reason why is more common to develop the HD when the father has the mutated allele (Tasset et al., 2009). An incomplete penetrance is observed in people with 35 to 39 repeats of the CAG trinucleotide; this people will be meiotically unstable and can develop Huntington’s disease. People with more than 40 glutamines in the sequence will have complete penetrance and will always develop the disorder (Rosales-Reynoso & Barros-Nuñez, 2007).

The diagnosis of Huntington’s disease is very unusual due to its late development, the diagnosis is made after the symptoms appear, familiar records and genetic analysis are the most commons ways of identification. Knowing the different ranges of the disease, polymerase chain reaction (PCR) has been used as a method for the diagnosis and identification of Huntington’s disease (www.ninds.nih.gov). Because people with mutated alleles don’t show different phenotypes, the molecular diagnosis is the only way of early detection of the disorder. Also, its late development, makes even more important the early detection of the condition, this is the reason why developing an easy and cheap method of molecular diagnosis to detect HD is so important.

1.3 Methods for HD molecular diagnosis

Waner et al. (1993) developed an efective PCR assay to detect the disorder; previous PCR assays had been designed, but were difficult to perform and unsuitable for routine use. The main trouble for the amplification of the sequence is the high concentration of the bases guanine and cytosine (G and C) in the DNA. This G and C high concentration causes secondary structures in the DNA due to the binding of this two bases. To enhance the conditions of the PCR and disrupt the secondary structures, commercially available kits can be used to amplify CG rich sequences. The problem with these kits is that they are expensive; other additives like DMSO, betaine, formamide or glycerol, can be used as PCR enhancer, making lower costs and the final products being able to be visualized by ethidium bromide stained agarose gel electrophoresis (Bhagya et al., 2013). Warner et al. (1993) used DMSO (dimethylsulphoxide) for the disruption of these secondary structures. DMSO changes the melting characteristics of the DNA and disrupts base paring; this can enhance the breakdown of secondary structures, but can also create nonspecific bands that can alter the interpretation of the results. Whereas, betaine is an isostabilicing agent that equalizes the contribution of CG and AT base paring to stabilize the DNA (Frackman et al., 1998). When DMSO is used, usually is mixed with betaine too, because DMSO denatures de DNA and betaine stabilizes the denatured DNA (Bhagya eta al., 2013). Furthermore, their method gave products from 80 to 143 base pair (bp) for normal alleles (11 to 32 CAG repeats); and 173 or more base pairs for mutated alleles (42 CAG repeats), these PCR products are small and require a 6% denaturing polyacrylamide gel.

2. Hypothesis and objectives

Since Warner et al.'s (1993) method for detecting HD was old and with the actual information about HD, my **hypothesis** was that it would be possible to improve the method that they created, making it easier to detected diseased people with this condition. The **main objective** of this work was to make a new effective method to detect Huntington's disease and the number of bp of each allele via PCR. Four mutated individuals were used to develop a method of detecting the disease, to check the effectiveness of the method, an individual with

chances of having HD (due to familiar record) was diagnosed with the created method. Different **aims** are going to be tested during the work, first, using betaine instead of DMSO to disrupt the secondary structures created by the high concentration of the bases G and C. It is expected better results in the disruption of the DNA using betaine, rather than DMSO, due to the isostabilizing effect of the betaine. It is also expected to appear nonspecific bands when no DMSO or betaine is used, because, as explained, the high concentration of CG can create secondary structures. Because of the length of the products of the PCR assay performed by Warner et al. (1993) were too short and required a denaturing polyacrylamide gel to determinate the absolute size, new primers had been designed for longer PCR products. Another aim was to prove that doing a semilogarithmic regression was posible to calculate the absolute size of the PCR products using only a 2-4% agarose gel, avoiding the use of DMSO and the denaturing polyacrylamide gel. The last aim was to find the best concentrations of reagents and conditions for the PCR, that's the reason why different reagent's concentrations and temperatures in the PCR program were changed during the work.

3. Materials and methods

3.1 Individuals

Seven anonymous individuals were used for the experiment. Four of the individuals had the mutated allele for Huntington's disease (individuals 6, 7, 9 and 15); other two individuals had normal alleles and were used as control (individuals 3 and 4). The developed method was used to genotype an unknown individual (individual 11), with potential chances of developing the condition (due to its family record) and make a diagnosis. All mutated individual's samples and individual 11 were bought from the Corriel Institute of Medical Research (www.corriel.org). The samples used as control were obtained from the laboratory's DNA collection. Because individual 3 and 4 showed the same amplification length during the control, only individual 4 was used in in the PCRs with mutated alleles as a normal allele sample. The numbers of (CAG) repeats were known for the mutated alleles (individuals with Huntington's disease), the data of the normal alleles (control samples) was unknown (Table 1). The DNA concentrations of the other samples were, 6: 293.3 ng/µl; 7: 319.8 ng/µl; 9: 306.8 ng/µl; 11: 303.2 ng/µl and 15: 295.5 ng/µl.

Table 1. Identification number of the anonymous individuals, the (CAG)n expansion of the mutated and control alleles of each individual (alleles with more than 40 (CAG) repeats are considered mutated alleles) and the expected length of the PCR amplification products for each allele. Individual 11 is going to be diagnosed to identify if it has the HD.

Individual	*Number of (CAG) repeats*	*Expected length (in base pair)*
6	48	229
7	55 & 15	250 & 130
9	70 & 15	295 & 130
15	66 & 16	283 & 133
11 (diseased?)	**No data**	**No data**
3	No data	No data
4	No data	No data

3.2 Primers

Two control primers were used to determinate the viability of the DNA before using the specific Htt primers. The primer's sequence used for the determination of the quality and viability of the DNA were, forward: 5'-GCA GCA GCT GGA CAA TGT CA-3'; reverse: 5'-GCC CCA GCA GAG GAA GAA AA-3'. The PCR product was expected to have 429 base pair.The genomic sequences of the specific primers were, Htt-F (Foward): 5'-CGA CCC TGG AAA AGC TGA TG-3'; Htt-R (Reverse): 5'-CTG AGG AAG CTG AGG AGG C-3'. The PCR product sequence had 85 base pair + (CAG)n trinucleotide repeats (Fig 1). Products between 112 to 193 bp range were given for normal alleles (people with 9 to 36 repeats) and 205 bp or bigger products for mutated alleles (>40 repeats, complete penetrance). People with incomplete penetrance (35 to 39 repeats) gave products in the 190 bp and 202 bp range.

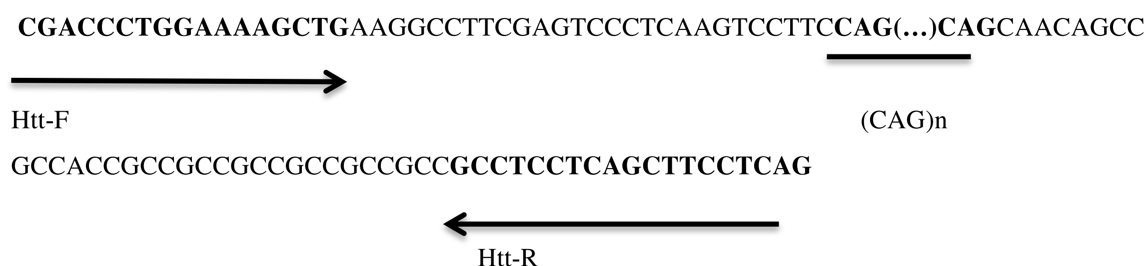


Fig. 1. Sequence of amplification of the PCR analysis. Arrows indicate the binding side of the primers (Htt-F and Htt-R). (CAG)n repeats underlined, the number of repeats vary on each individual, giving different lengths on the PCR products.

3.3 PCR reagents concentrations, thermal cycler conditions and migration

The different PCRs that were performed had equal concentrations of primers (specific or non-specific), dNTPs, Tris-HCl (pH=8.3), KCl and Taq polymerase. All the PCRs were performed in a 50 µl volume, using 200 µM of each dNTPs (800 µM of total dNTPs), 200 µM of each primer, 10 mM of Tris-HCl (pH=8.3), 50 mM of KCl and 1.25 units of Taq polymerase. $MgCl_2$, DNA, dimethylsulfoxide (DMSO) and betaine (Trimethylglycine) concentration vary depending the PCR assay.

The thermo cycler conditions were equal in the first and third stage; the second step of the second stage varies in the different PCR assays (the first and third steps of the second stage were equal in all PCR assays). In the first stage, the thermal cycler was heated to 94°C for 3 min; them, the second stage had 35 cycles, the first step was at 94°C for 1 min, the second step lasted 90 s in all PCR (but the temperature changed depending on the analysis) and the third step was at 72°C for 150 s, after all the cycles were complete the final stage lasted 7 min at 72ºC.

All the polymerase chain reactions (PCR) were performed in an Applied Biosystems Veriti Thermal Cycler. To determinate the length of the products GeneRuler™ 1kb Plus DNA Ladder was used (Fig 2). The products of the PCR were resolved in a 2% agarose gel and stained with bromide ethidium. All the migrations were performed with 10 µl of each sample and 2,5 µl of the DNA ladder.

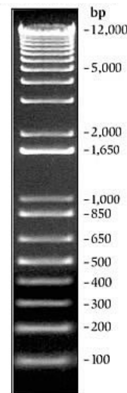


Fig. 2. Length (in base pair, bp) of the GeneRuler™ 1kb plus DNA ladder in a 1% agarose gel. This DNA ladder was used in all the PCR assays of the work.

3.4 Regression

With the distance of migration and the $\log_{10}$ of the length (base pair), of the DNA ladder bands, two semilogarithmic regressions were made. The equation of the regressions were used to calculate and identify the length of the bands of the different individuals. One of the regressions was used to estimate the bp number of bands <200 bp and the other regression was used to calculate bands >200 bp.

4. Results and Discussion

4.1 Optimization of the PCR with control DNA of non-affected individuals

4.1.1 Quality control of control DNA

Before using the specific primers designed to detect the Huntington's disease, control primers were used to test the viability of the control DNA. The main reason was to make sure that the control DNA had enough quality to be used as template for PCR reaction. If it wasn't able to amplify, a different sample of control DNA should have been used instead. The other reason was to check that the designed primers to detect Huntington's disease were able to amplify

DNA. After testing the viability of the DNA with the control primers, when using the specific primers if no DNA was amplified, it would be a mistake on the primers and not a problem on the DNA.

The PCR was performed with the concentrations mentioned in the Materials and methods, with 3 mM of $MgCl_2$ and without DMSO or betaine. The thermo cycler was heated to 94°C for 3 min, them 35 cycles at 94°C for 1 min, 56°C for 90 s and 72°C for 150 s, after all the cycles were complete the final stage lasted 7 min at 72ºC. A 2% agarose gel was used to resolve the products of the PCR (Fig 3).

Two different control DNA samples were used, individuals 3 and 4 (Fig 3). The concentration of DNA of the samples was unknown. Different volumes (0.5 and 1 µl) of DNA were used for the PCR assay.

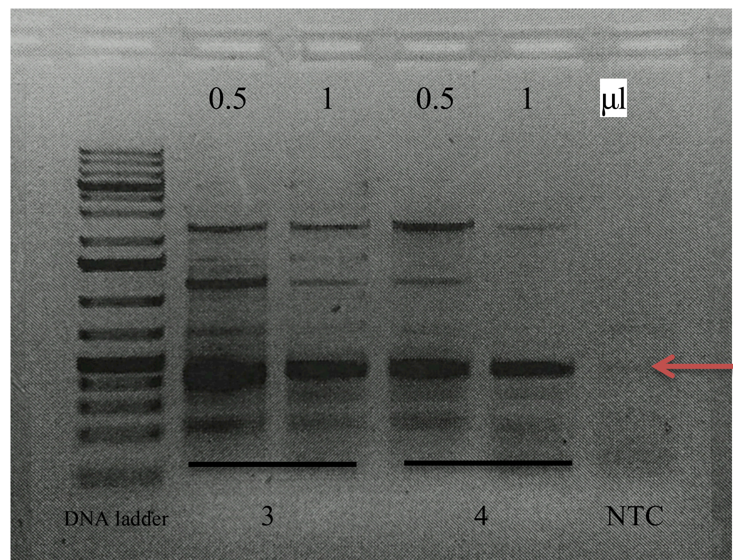


Fig. 3. PCR analysis of the control DNA samples (individuals 3 and 4) with non-specific primers in a 2% agarose gel stained with bromide ethidium. Individual 3 (0.5 and 1 ul DNA; lanes 1 and 2), individual 4 (0.5 and 1ul DNA; lanes 3 and 4), the DNA ladder and the non-template control (NTC) . Arrow indicates the expected 429 bp band.

The results showed that the control DNA was able to amplify with non-specific primers; also that 0.5 µl amplified better than 1µl. The expected 429 bp band appeared in all 4 lanes of the electrophoresis (Fig 3).

4.1.1 Detection of the normal Htt alleles

4.1.1.1 Optimization of the annealing temperature

After testing that the control DNA was able to amplify, the specific primers for detecting Huntington's disease were used with the control DNA. Only the control DNA was used (individuals 3 and 4), to make sure that the designed primers were able to amplify.

The PCR was performed in a 50 µl volume, using the same concentrations as the ones used with the non-specific primers (but with the designed primers instead). After heating the thermocycler to 94°C for 3 min, 35 cycles of 1min at 94°C, 90 s at 55°C or at 53°C (two different temperatures were used in this stage) and 150 s at 72°C; followed by 7 min at 72°C. The PCR products were resolved in a 2% agarose (Fig 4).

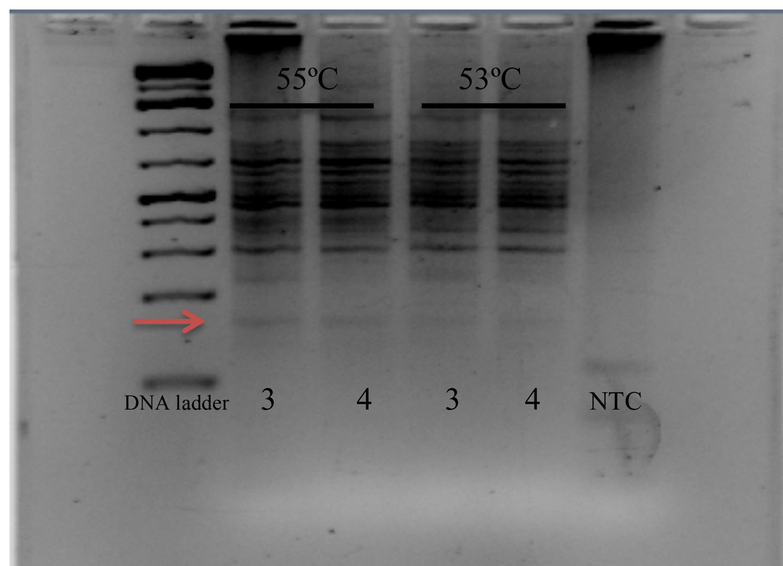


Fig. 4. PCR analysis of the control DNA samples (individuals 3 and 4) with the specific Htt primers, performed in a 2% agarose gel stained with bromide ethidium. Different temperatures were used, individual 3 at 55°C (lane 1), individual 4 at 55°C (lane 2), individual 3 at 53°C (lane 3), individual 4 at 53°C (lane 4), DNA ladder and Non template control (NTC). Arrow shows the expected product size for a normal allele (>196 bp).

The bands that were expected to appear, were supposed to have between 112 and 196 base pairs. The results mainly showed bands between 500 and 700 base pairs, but a weak band of the expected size appeared in the samples that in the second stage of the 35 cycles were heated at 55°C (Fig 4).

Because the highest temperature used in the PCR (55°C) showed a weak line, higher temperatures were used to determinate which temperature was better for the amplification of the DNA. Six different temperatures were used during the second stage of the 35 cycles. The temperatures used went from 55ºC to 60°C for each individual of the control DNA. The PCR products were resolved in a 2% agarose gel (Fig 5).

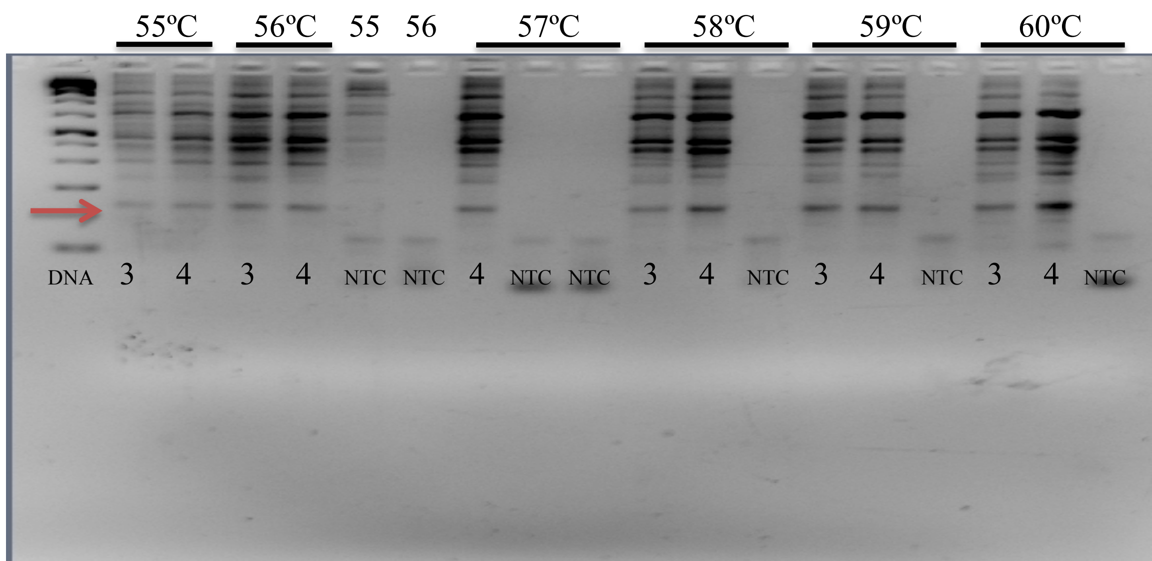


Fig. 5. PCR analysis of the control DNA (individuals 3 and 4) in a 2% agarose gel, stained with bromide ethidium. Different temperatures were used in the reaction, 55ºC (individuals 3 and 4, lanes 1 and 2), 56ºC (individuals 3 and 4, lanes 3 and 4) 57ºC (individual 4, lane 7), 58ºC (individuals 3 and 4, lanes 10 and 11), 59ºC (individuals 3 and 4, lanes 13 and 14), 60ºC (individuals 3 and 4, lanes 16 and 17), the DNA ladder and the respectives non template controls (NTC) for each temperature. Arrow shows the expected product size for a normal allele (>196 bp).

The products of the PCR increased with the temperature, especially in temperatures 57ºC and 58ºC. In temperatures higher than 58ºC the products of the PCR decreased in quantity.

Although the expected band size appeared (band between 112-196 base pair), non-specific bands between 300 and 700 base pair appeared again (Fig 5).

As explained in the Introduction, due to the repetition of (CAG)n , a high concentration of the nucleotides cytosine (C) and (G) appear in the sequence. As expected, this high concentration of cytosine and guanine causes folds on the structure of the DNA chain, possibly due to the absence of DMSO or betaine in the PCR conditions. The secondary structures created by the linkage of C and G couldn't migrate as desired in the agarose gel, creating the non-specific bands that appeared in Fig 5 and Fig 6. It could have been because of the high concentration of $MgCl_2$. So one question to ask was, if due to the high concentration of $MgCl_{2,}$ the Taq polymerase was malfunctioning and hence creating more non-specific bands. Thus, another question to address was if the treatment with DMSO or betaine could improve the result by diminishing the secondary structures formation in the template DNA and, therefore enhance specific amplicon formation.

4.1.1.2 Effects of the treatment with DMSO

Like Warner et al. (1993) did, DMSO was used for the disruption of the secondary structures created by the bases C and G. Also $MgCl_2$ concentration was lowered for the optimization of the polymerase, better results in the polymerase activity were expected when less concentration of $MgCl_2$ was used. Four different amplifications, with different concentrations of the reagents, were performed with the individuals 3 and 4.

The concentrations of the reagents for the PCR were the ones mentioned in the Materials and methods section. The assays were performed with 3 mM of $MgCl_2$, 3 mM of $MgCl_2$ and 10% DMSO, 1.5 mM $MgCl_2$ and with 1.5 mM $MgCl_2$ and 10% DMSO. The temperature used in the second step of the second stage was 58ºC.

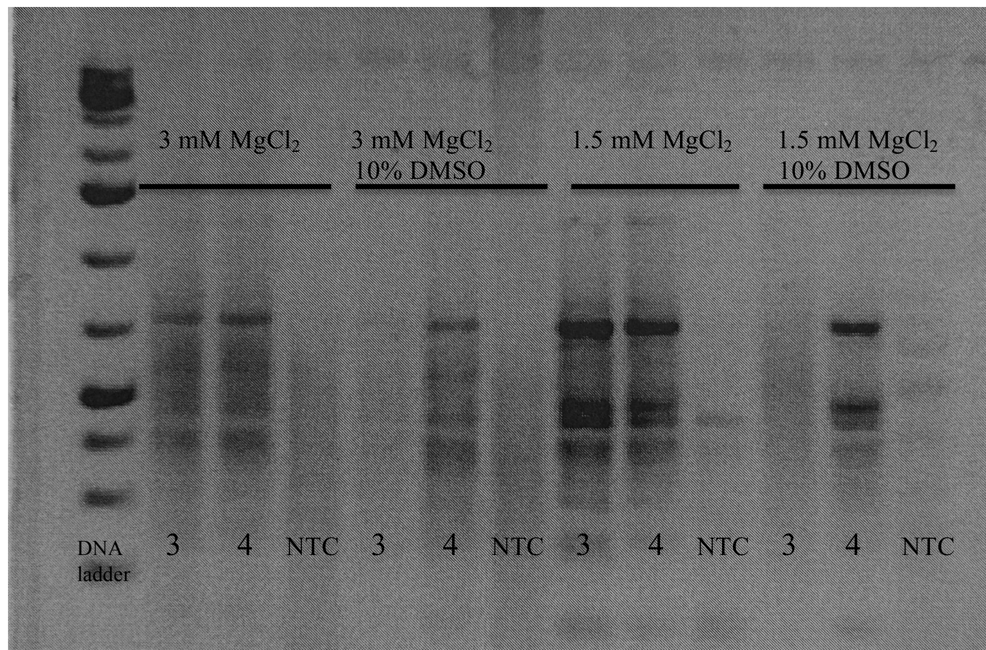


Fig. 6. PCR analysis of the control DNA samples (individuals 3 and 4), with different concentrations in the master mix due to the use of DMSO and lower concentration of $MgCl_2$. The PCR products were resolved in a 2% agarose gel stained with bromide ethidium. Different concentration of the reagents used 3 mM $MgCl_2$ (lanes 1-3), 3 mM $MgCl_2$ and 10% DMSO (lanes 4-6), 1.5 mM $MgCl_2$ (lanes 7-9), 1.5 mM $MgCl_2$, 10% DMSO (lanes 10-12), DNA ladder and non template control (NTC).

The results showed that lower concentrations of $MgCl_2$ (1.5 mM) improved the results of the PCR reaction, creating less non-specific bands and concentrating same size amplicons (Fig 6). Unlike Warner et al. (1993), DMSO wasn't effective for the disruption of the secondary structures created by the high concentration of C and G. This might be due to the effect of DMSO in the DNA chain. DMSO can cause mismatches in the primer-binding step and can affect the Taq polymerase binding (Hardjasa et al., 2010), this can hider the PCR reaction and alter the results. In lanes 4 and 10 can be seen that Individual 3's DNA didn't amplify when DMSO was used, this could be because of a mismatching on the primers with the DNA sequence, making impossible the identification and measurement of the length of the expected bands (Fig 6).

4.1.1.3 The effects of the treatment with betaine

Betaine was used instead of DMSO for the disruption of the DNA secondary structures. The concentrations and conditions of the PCR were the ones described in the Material and

methods section with 1.3 M of betaine, 1.5 mM of $MgCl_2$ and 58ºC in the second step of the second stage. Since individual 3's DNA wasn't able to amplify in presence of DMSO and previous PCR products showed same band size in individual 3 and 4's expected products, only individual 4's DNA was used in this assay.

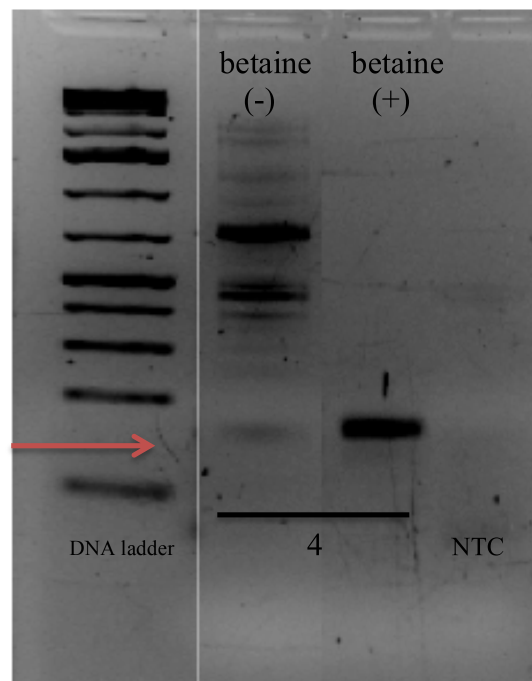


Fig. 7. PCR analysis of the Individual 4 with Htt primers, performed in a 2% agarose gel and stained with bromide ethidium. The PCR was made without betaine (lane 1), with betaine (lane 2), DNA ladder and non template control (NTC). Arrow shows the expected band size (>196 bp).

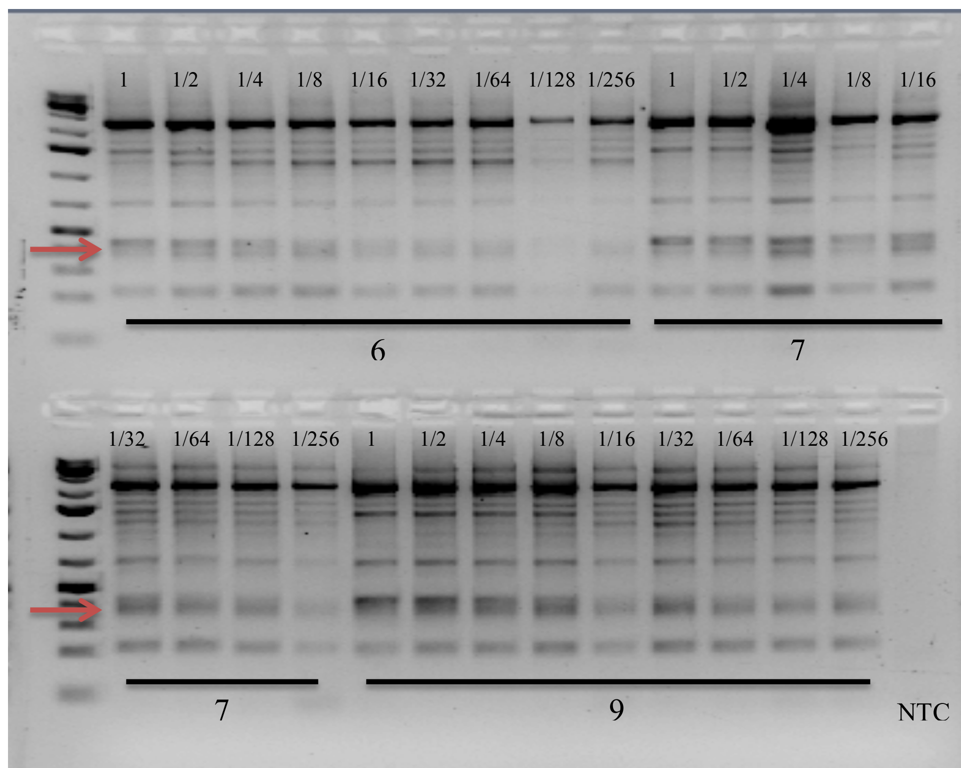
When betaine was used in the PCR reaction, the disruption of the secondary structures of the DNA was successful. Unlike DMSO, the results showed that betaine was more effective in the secondary structures disruption. When no betaine was used in the PCR nonspecific bands appeared in the PCR products (Fig 7). Also, betaine enhanced the PCR reaction better than DMSO, due to its isostabilizing activity on the DNA, whereas DMSO denatures the DNA sequence creating mismatches and hindering the PCR reaction. As Henke et al. (1997) and Frackman et al. (1998) showed in their respective works, betaine is a more effective PCR enhancer in presence of high concentrations of G and C DNA sequences.

4.2 Analysis of the control DNA of affected individuals and test DNA

4.2.1 Quality control and minimum concentration of DNA for PCR

Before using the specific primers (Htt primers) with the mutated DNA, the nonspecific primers were used to amplify the mutated DNA, in order to check for the quality of the mutated DNA. Also, serial dilutions of the mutated DNA were performed to determine the best concentration of the DNA, for the later amplification with the Htt primers.

The dilutions performed were 1, 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128 and 1/256 for all the mutated samples. The PCR was performed with the concentrations and conditions described in the Material and methods, with 1.5 mM of $MgCl_2$ and 58ºC on the second step of the second stage of the PCR.



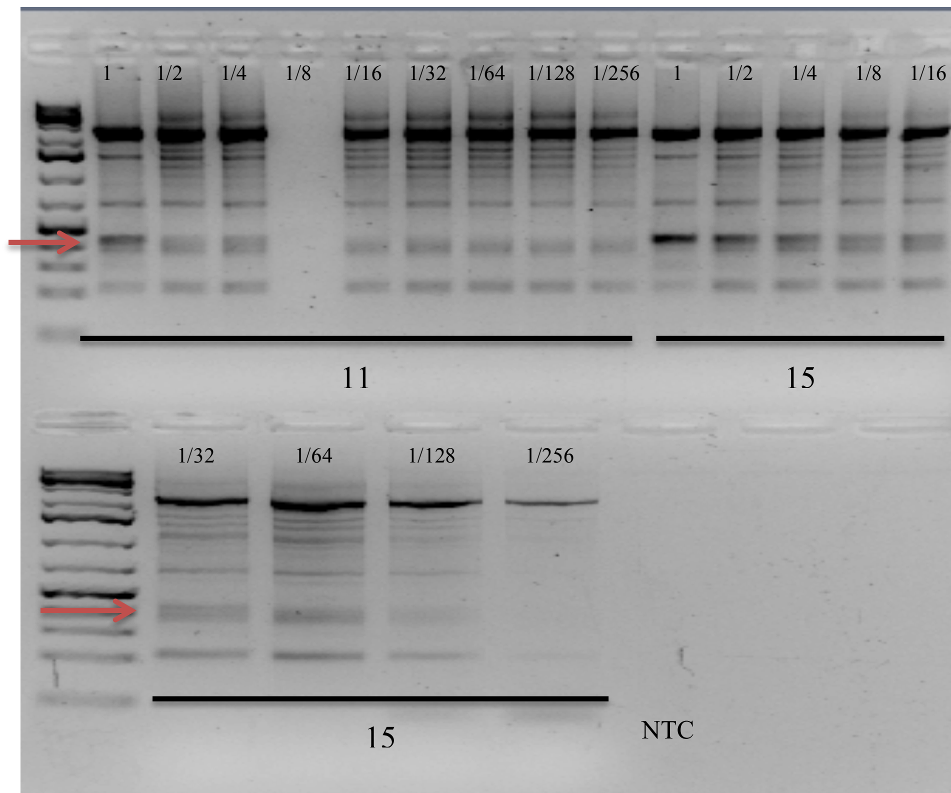


Fig. 8. PCR analysis of the mutated DNA samples individuals 6 (lanes 1-9), 7 (lanes 10-18), 9 (lanes 19-27), 11 (lanes 28-36) 15 (lanes 37-46), DNA ladder and non template control (NTC); with non-specific primers in a 2% agarose gel, stained with bromide ethidium. Above each lane the respective dilution factor 1, 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128 and 1/256. Arrow indicates the expected 429 bp bands.

All mutated DNA samples were able to amplify in presence of the non-specific primers (Fig 8). In all the individuals, the selected dilution for the amplification of the DNA with Htt primers was 1/16 of the original concentration. The concentrations of DNA for each individual (1/16 dilution) were, for individual individual 6: 18.33 ng/µl; individual 7: 19.99 ng/µl; individual 9: 19,17; individual 11: 18.95 ng/µl and 15: 18.47 ng/µl.

4.2.2 Detection of the affected Htt alleles with the optimized methodology

Finally, to calculate the number of CAG repeats, the mutated DNA was amplified with the Htt primers and a semilogarithmic regression was made. Mutated DNA, was expected to have >196 bp (36 CAG repeats), bands smaller than <196 bp would indicate a normal allele; but, bands between 190 and 202 bp would indicate an incomplete penetrance, which can be meiotically unstable and develop Huntington's disease.

The PCR was made with the concentrations and conditions mentioned in the Materials and methods. 1.3 M of betaine was used for the disruption of the secondary structures and 1.5 mM of $MgCl_2$ for the optimal functioning of the Htt primers. The second step of the second stage of the PCR was performed at 58ºC.

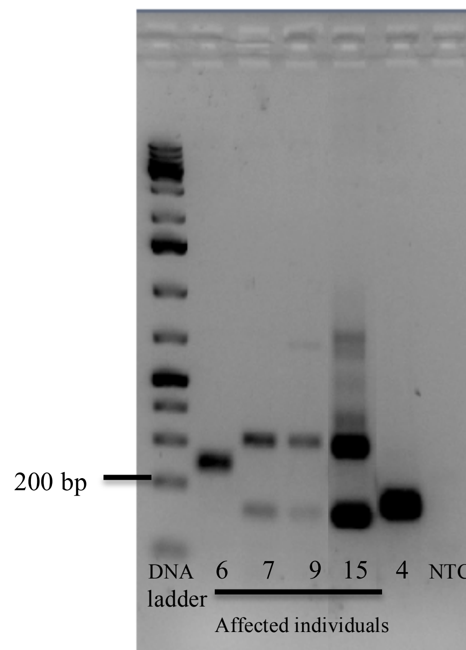


Fig. 9. PCR analysis of the mutated DNA samples (individuals 6, 7, 9 and 15; lanes 1, 2, 3 and 4 respectively), normal allele (individual 4; lane 5), DNA ladder and non template control (NTC); performed in a 2% agarose gel and stained with bromide ethidium. Bands over 200 bp indicate more than 38 CAG repeats in the DNA sequence.

As expected, the betaine used in the PCR reaction made the electrophoresis viable; the respective bands sizes for each individual were observed (Fig 9). The DNA ladder and the

bands distance were used to create a semilogarithmic regression and estimate the CAG repeats in each individual.The distance of the DNA ladder bands and the $\log_{10}$ of the length of those bands (in bp) were used to make the two semilogarithmic regression (Fig 10). The already known band size of each individual was used to compare with the estimated size (Table 2).

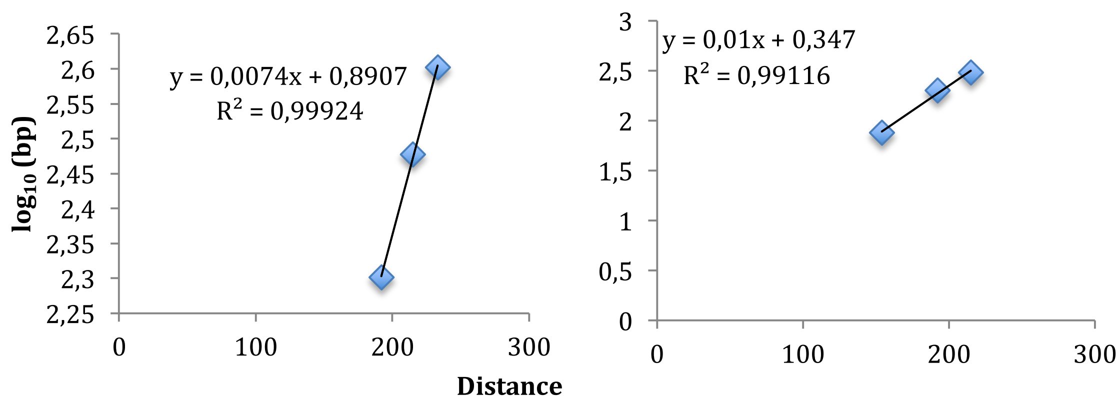


Fig. 10. The semilogarithmic regression and its equations, used estimate the number of bp of the different individuals. The regression on the left was used to calculate bands with >200 bp. The regression on the right was used to calculate bands with <200 bp.

Table 2. The expected length and the calculated length (in bp) of the mutated individuals. The length of each band was calculated with the equations of the semilogarithmic regression in Fig 10.

Individuals	Expected length (bp)	Calculated length (bp)
6	229	230.83
7	250 & 130	293.02 & 127.93
9	295 & 130	293.02 & 127.93
15	283 & 133	283. 20 & 133.97

Using the equations calculated in Fig 10 the length of each band was estimated. The estimated length size of each band was almost exact (±1.398 bp). Besides individual 7's mutated band

(>196 bp) couldn't be measured correctly due to a wrong migration (Table 2). It was proved that the method was effective to distinguish normal and mutated alleles (Fig 9), and that it was posible to measure the bp number of those alleles (Fig 10). Using the equations of Fig 10, individual 11's bp number was going to be measured to identify if it was affected by Huntington's disease.

4.3 Molecular diagnosis of individual 11

Because the developed method was effective, individual 11 was diagnosed to check if he had HD and to measure the number of bp of its alleles. The reason why individual 11 is tested with this method, is because, as explained in the Materials and method section, this individual could be affected by Huntington's disease due to familiar record.

Fig 9 was edited for the better understanding of the results, indeed Fig 9 and Fig 11 are the same gel. The reason why it has to be the same gel is to apply the equations of Fig 10, its regression is specific for each gel, because the migrated distance vary each time the electrophoresis is performed. This means that a new regression has to be done each time a test is performed.

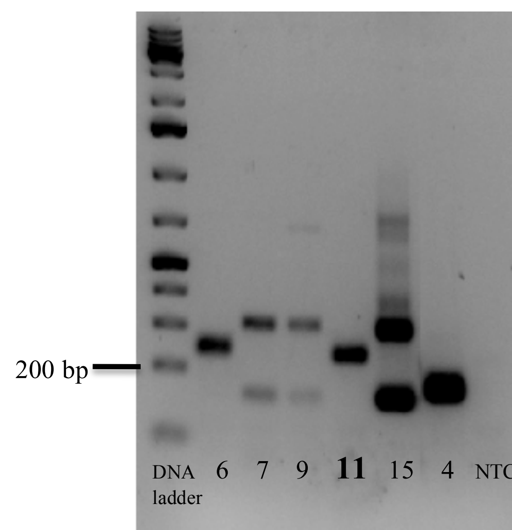


Fig. 11. PCR analysis of the mutated DNA samples (individuals 6, 7, 9 and 15; lanes 1, 2, 3 and 5 respectively), normal allele (individual 4; lane 6), test DNA (individual 11; lane 4), DNA ladder and non template control (NTC); performed in a 2% agarose gel and stained with bromide ethidium. The results showed that individual 11 was affected by HD. Bands over 200 bp indicate more than 38 CAG repeats in the DNA sequence.

The results showed that individual 11 was affected by Huntington's disease, because its band was over 200 bp (Fig 11). For the exact measurement of the number of bp, the equations of the Fig 10 were used, the calculations showed that he had 219,33 bp which is 45 (CAG)n repeats, resulting in a affected individual of the HD.

4.4 Relevant outcomes

The molecular diagnosis developed by Warner et al. (1993) was effective but required the use of DMSO, which can cause problems in the binding sides of the primers, hindering the amplification of the products in the PCR. Also, a denaturing polyacrylamide gel is required for the estimation of the length of the bands. With the method developed in this work the estimation of the length of the bands only require a 2% agarose gel. It has to be said that some complications may appear, for example a wrong migration can cause an erroneous estimation of the length of the bands (individual 7, Table 2). Also an individual with two alleles, with difference of ± 1-2 CAG repeats might not be distinguish if it is heterozygote or homozygote. Diluting the DNA and creating thinner band sizes can overcome these problems. The method developed in this work can be an easy and cheap way to find individuals with Huntington's disease.

5. Conclusion

1. Regarding to the PCR concentrations and conditions, high concentration of $MgCl_2$ create more nonspecific bands due to the different sizes of the amplicons, also, 58ºC was the optimal temperature for the amplification.

2. When no betaine or DMSO was used, nonspecific bands appear because of the high concentration of the bases C and G.

3. Betaine works better than DMSO, in the disruption of the secondary structures formed by the high concentration of the bases C and G.

4. The developed method works and it´s suitable for molecular diagnosis of the Huntington´s disease. A rough observation of the results is enough to predict the outcome of the individuals, since individuals with over 200 bp have a mutated allele.

5. In order to perform an accurate diagnosis, creating a semilogarithmic regression the exact number of bp can be estimated, knowing the size of the amplicon, the exact number of CAG repeats can be estimated.

Acknowledgements

I would like to thank Ainjoa Iglesias, Ana Zubiga and the AZ group for support that they offered me to develop this work. Also an especial thank to Naiara Zorilla for the technical help during the work.

6. References

Bhagya, C. H. W. M. R. C., Wijesundera Sulochana, W. S., & Hemamali, N. P. 2013. Polymerase chain reaction optimization for amplification of Guanine-Cytosine rich templates using buccal cell DNA. *Indian Journal of Human Genetics*, *19*(1), 78–83.

Corriel Institute of Medical Research (www.corriel.org)

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

Genetic Alliance. 2010. *Understanding Genetics: A New England Guide for Patients and Health Professionals*. The New England Public Health Genetic Education Collaborative, Washington (DC).

Hardjasa A., Ling M., Ma K. & Yu H. 2010. Investigating the Effects of DMSO on PCR Fidelity Using a Restriction Digest-Based Method. *Journal of Experimental Microbiology and Inmunology* (JEMI). 14: 161-164.

Henke W., Herdel K., Jung K., Schnorr D. & Loening S.A. 1997. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acid Research* 25 (19): 3957-3958.

National Institute of Neurological Disorders and Strokes (www.ninds.nih.gov)

Roos R.A.C. 2010. Huntington's disease: a clinical review. *Orphanet Journal of Rare Diseases*, 5:40.

Rosales-Reynoso M.A. & Barros-Nuñez P. 2008. Diagnóstico molecular de la enfermedad de Huntington. *Gac Med Mex*., 144 (3): 271-273.

Tasset I., Sánchez F. & Tunez I. 2009. Bases moleculares de la enfermedad de Huntington: papel del estrés oxidativo. *Revista de Neurologia*, 49 (8): 424-429.

Vazquez-Cerdas M., Campos-Ramirez D., Gutiérrez B., Fernández-Morales H., Morales-Montero F. & Cuenca-Berger P. 2011. Integral management of Costa Rican Patients with Huntington's Disease and their Families. *Acta med. costarric.,* 53.

Warner J.P., Barron L. H. & Brock D.J.H. 1993. A new polymerase chain reaction (PCR) assay for the trinucleotide repeat that is unstable and expanded on Huntington's disease chromosomes. *Molecular and Cellular Probes*, 7: 235-239.