



FACULDADE DE MEDICINA DA UNIVERSIDADE DO PORTO

**Mestrado em Medicina e Oncologia Molecular  
2005-2006**

**CARACTERIZAÇÃO DE SETE VARIAÇÕES DE  
SEQUÊNCIA IDENTIFICADAS NO GENE *NOTCH3*  
EM DOENTES PORTUGUESES COM HIPÓTESE  
DIAGNÓSTICA DE DOENÇA DE CADASIL**

**MARIA MARGARIDA MARQUES NEVES VENÂNCIO**

**2007**

Faculdade de Medicina da Universidade do Porto



Mestrado em Medicina e Oncologia Molecular  
2005-2006

*Caracterização de sete variações de sequência  
identificadas no gene NOTCH3  
em doentes portugueses com hipótese  
diagnóstica de doença de CADASIL*

---

*Maria Margarida Marques Neves Venâncio*

2007



Caracterização de sete variações de sequência  
identificadas no gene *NOTCH3*  
em doentes portugueses com hipótese diagnóstica  
de doença de CADASIL

Dissertação apresentada à Universidade do Porto para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Medicina e Oncologia Molecular, realizada sob orientação da Professora Doutora Filipa Carvalho (Universidade do Porto)

---

*Maria Margarida Marques Neves Venâncio*

À **Professora Doutora Filipa Carvalho**, orientadora deste trabalho, agradeço pela orientação científica, apoio e incentivo incansáveis para a realização deste trabalho.

Ao **Professor Doutor Jorge M. Saraiva**, director do Serviço de Genética Médica do Centro Hospitalar de Coimbra, agradeço pela orientação, disponibilidade, generosidade e incentivo constantes para a realização deste Mestrado.

Ao **Professor Doutor Alberto Barros**, director do Serviço de Genética da Faculdade de Medicina do Porto, agradeço pela generosidade em me receber e disponibilizar o seu Serviço para a realização da parte experimental do Mestrado.

Ao **Dr. João Paulo Oliveira**, agradeço pela transmissão dos seus conhecimentos sobre a doença de CADASIL.

Às **Doutora Susana Fernandes** e **Dr.<sup>a</sup> Susana Ferreira**, agradeço toda a simpatia, disponibilidade e conselhos laboratoriais providenciais.

À **Dr.<sup>a</sup> Lina Ramos**, minha orientadora do Internato Complementar de Genética Médica, agradeço pela amizade e motivação diária para a realização deste trabalho.

À **Dr.<sup>a</sup> Carla Filipe**, muito obrigada pela sua amizade e generosidade em partilhar comigo todos os seus conhecimentos e o seu tempo.

Aos **meus queridos amigos** [Di, Fabi, Gi, Maria Teresa, Pipa, Tininha, Xana, Carlos, Jquim, Miguel e Sérgio (meu companheiro nesta *viagem*)], agradeço pelo amor e apoios incondicionais durante estes dois anos de grande sacrifício.

À minha Mãe que apenas me viu *sonhar*...  
Ao meu Pai e Irmão que me ajudaram a concretizar este *sonho*.

Índice de Figuras .....	vi
Índice de Quadros .....	viii
Resumo .....	ix
Abstract .....	xii
<b>1. Introdução .....</b>	<b>1</b>
1.1. Perspectiva histórica .....	2
1.2. Epidemiologia .....	3
1.3. Clínica .....	3
1.3.1. Acidentes isquémicos tipo-enfarte .....	4
1.3.2. Défice cognitivo .....	5
1.3.3. Perturbações psiquiátricas .....	5
1.3.4. Enxaqueca .....	6
1.3.5. Achados menos frequentes .....	6
1.3.5.1. Epilepsia .....	6
1.3.5.2. Quadros de encefalopatia aguda .....	6
1.3.5.3. Doença vascular cardíaca .....	6
1.3.5.4. Neuropatia periférica .....	6
1.3.5.5. Anomalias oculares .....	7
1.4. Evolução e prognóstico .....	7
1.5. Imagiologia .....	8
1.5.1. Ressonância magnética cerebral .....	8
1.5.1.1. Hiperintensidade na substância branca .....	9
1.5.1.2. Lesões lacunares .....	9
1.5.1.3. Enfarte lacunares subcorticais .....	9
1.5.1.4. Micro-hemorragias cerebrais .....	9
1.5.1.5. Atrofia cerebral .....	10
1.6. Técnicas de anatomia-patológica .....	10
1.6.1. Biopsia de pele .....	10
1.6.2. Estudos anátomo-patológicos .....	11
1.7. Diagnóstico .....	11
1.8. Diagnósticos diferenciais .....	12
1.9. Tratamento e seguimento .....	13

1.10.	Genética .....	14
1.10.1.	<i>NOTCH3</i> (MIM 600276) e Via de Sinalização Notch .....	14
1.10.2.	Características das mutações .....	15
1.10.3.	Estudo molecular na prática clínica .....	16
1.10.4.	Correlação genótipo-fenótipo .....	17
1.10.5.	Penetrância e expressividade .....	17
1.10.6.	Antecipação .....	17
1.11.	Aconselhamento genético .....	17
1.11.1.	Diagnóstico pré-sintomático .....	19
1.11.2.	Diagnóstico pré-natal .....	21
1.11.3.	Diagnóstico genético pré-implantação .....	22
<b>2.</b>	<b>Objetivos .....</b>	<b>23</b>
<b>3.</b>	<b>Material e Métodos .....</b>	<b>25</b>
3.1.	Amostra .....	26
3.2.	Variações de sequência analisadas .....	26
3.3.	Técnicas de biologia molecular .....	27
3.3.1.	Extracção de DNA .....	27
3.3.2.	Reacção de Polimerase em Cadeia (PCR) Convencional .....	27
3.3.3.	Sequenciação directa dos produtos de PCR .....	28
3.3.3.1.	Purificação dos Produtos de PCR .....	28
3.3.3.2.	Reacção de Sequenciação .....	28
3.3.3.3.	Precipitação dos produtos a sequenciar .....	29
3.3.3.4.	Sequenciação .....	29
3.4.	Análise das variações de sequência localizadas no exão 11 .....	30
3.5.	Análise das variações de sequência localizadas nos exões 18/19 e intrão 18 .....	30
3.6.	Análise da variação de sequência localizada no exão 20 .....	30
3.6.1.	Digestão com a enzima HpyCH4III .....	31
3.7.	Análise de dados .....	32
<b>4.</b>	<b>Resultados .....</b>	<b>34</b>
4.1.	Resultados da análise das variações de sequência localizadas no exão 11 ....	35
4.1.1.	Aplicação de ferramentas de bioinformática para avaliação do grau de associação desta variação com doença .....	37
4.2.	Resultados da análise da variação de sequência localizada no intrão 18 .....	38
4.3.	Resultados da análise das variações de sequência localizadas no exão 19 ....	39
4.3.1.	Aplicação de ferramentas de bioinformática para avaliação do grau de associação desta variação com doença .....	41
4.4.	Resultados da análise da variação de sequência localizada no exão 20 .....	41

4.4.1.	Aplicação de ferramentas de bioinformática para avaliação do grau de associação desta variação com doença .....	43
<b>5.</b>	<b>Discussão e Conclusões .....</b>	<b>44</b>
5.1.	Polimorfismos versus Mutações .....	45
5.2.	Mutações e polimorfismos descritos na doença de CADASIL .....	46
5.3.	MSRV <i>system</i> .....	48
5.4.	Discussão dos resultados obtidos .....	50
5.4.1.	Exão 11 – Variações de sequência C568Y, T575T e G594G .....	50
5.4.1.1.	Tipo de variações de sequência .....	50
5.4.1.2.	Diferenças entre as frequências alélicas determinadas nos controlos e nos doentes .....	50
5.4.1.3.	Avaliação do grau de associação da variação de sequência C568Y utilizando a base MSRV .....	51
5.4.2.	Intrão 18 – Variação de sequência IVS18+73A>G .....	51
5.4.2.1.	Tipo de variação de sequência .....	51
5.4.2.2.	Diferenças entre as frequências alélicas determinadas nos controlos e nos doentes .....	51
5.4.3.	Exão 19 – Polimorfismo A1020P e a variação de sequência W1028C .....	52
5.4.3.1.	Tipo de variações de sequência .....	52
5.4.3.2.	Diferenças entre as frequências determinadas nos controlos e nos doentes .....	52
5.4.3.3.	Avaliação do grau de associação das variações de sequência A1020P e W1028C utilizando a base MSRV .....	52
5.4.4.	Exão 20 – Variação de sequência C1099Y .....	53
5.4.4.1.	Tipo de variação de sequência .....	53
5.4.4.2.	Diferenças entre as frequências determinadas nos controlos e nos doentes .....	53
5.4.4.3.	Avaliação do grau de associação com doença da variação de sequência C1099Y utilizando a base MSRV .....	53
5.5.	Conclusões .....	53
<b>6.</b>	<b>Anexos.....</b>	<b>56</b>
<b>7.</b>	<b>Referências bibliográficas .....</b>	<b>87</b>



<b>Figura 1</b> – Sequência de investigação etiológica .....	12
<b>Figura 2</b> – Sequência do protocolo do diagnóstico pré-sintomático .....	20
<b>Figura 3</b> – Representação esquemática da acção da enzima HpyCH4III no produto de PCR do exão 20 na ausência de variação de sequência .....	31
<b>Figura 4</b> – Representação esquemática da acção da enzima HpyCH4III no produto de PCR do exão 20 na presença de variação de sequência .....	31
<b>Figura 5</b> – Distribuição das variações de sequência do exão 11 nos doentes e nos controlos .....	36
<b>Figura 6</b> – Electroforograma representando a variação de sequência T575T .....	36
<b>Figura 7</b> – Resultado da análise da variação de sequência C568Y utilizando a base MSR.V .....	37
<b>Figura 8</b> – Número de controlos e doentes onde foi identificada esta variação intrónica .....	38
<b>Figura 9</b> – Número de controlos e doentes onde foram identificadas as variações de sequência no exão 19 .....	39
<b>Figura 10</b> – Electroforograma representando as variações de sequência IVS18+73A>G e A1020P .....	40
<b>Figura 11</b> – Resultado da análise da variação de sequência A1020P utilizando a base MSR.V .....	41
<b>Figura 12</b> – Resultado da análise da variação de sequência W1028C utilizando a base MSR.V .....	41
<b>Figura 13</b> – Número de controlos e doentes onde foram identificadas as variações de sequência no exão 20 .....	42
<b>Figura 14</b> – Electroforese em gel de agarose Nusieve 3:1 a 3% durante 60' a 110V após digestão com a enzima HpyCH4II dos produtos de PCR do exão 20 de um grupo de amostras .....	42

<b>Figura 15</b> – Resultado da análise da variação de sequência C1099Y utilizando a base MSRV .....	43
<b>Figura 16</b> – Distribuição das mutações já descritas no gene <i>NOTCH3</i> .....	47
<b>Figura 17</b> – Esquemas representativo dos modelos tradicional e bioinformático .....	49

<b>Quadro 1</b> – Critérios de diagnóstico propostos por Davous .....	4
<b>Quadro 2</b> – Manifestações clínicas e complicações na fase terminal .....	7
<b>Quadro 3</b> – Idade de início dos AVC, imobilização e morte .....	8
<b>Quadro 4</b> – Diferenças entre as técnicas de diagnóstico pré-natal disponíveis .....	21
<b>Quadro 5</b> – Número de controlos estudados de acordo com os exões onde se verificaram as variações de sequência encontradas .....	26
<b>Quadro 6</b> – Variações de sequência estudadas e respectiva metodologia de análise.	26
<b>Quadro 7</b> – Sequências oligonucleotídicas utilizadas como <i>primers</i> para amplificar os fragmentos do gene <i>NOTCH3</i> a analisar .....	27
<b>Quadro 8</b> – Determinação das frequências das variações de sequência estudadas no exão 11, nos controlos e doentes .....	36
<b>Quadro 9</b> – <i>Odds ratio</i> das variações de sequência no exão 11 .....	37
<b>Quadro 10</b> – Frequência da variação IVS18+73A>G nos controlos e doentes .....	38
<b>Quadro 11</b> – Avaliação do <i>odds ratio</i> da variação IVS18+73A>G .....	39
<b>Quadro 12</b> – Frequências encontradas nos controlos e doentes para as alterações estudadas no exão 19 .....	40
<b>Quadro 13</b> – Determinação do <i>odds ratio</i> para as alterações A1020P e W1028C .....	40
<b>Quadro 14</b> – Determinação das frequências da variação de sequência C1099Y nos controlos e nos doentes .....	43
<b>Quadro 15</b> – <i>Odds ratio</i> da variação de sequência C1099Y .....	43
<b>Quadro 16</b> – Número de mutações descritas acordo com HGMD .....	46
<b>Quadro 17</b> - Alguns dos polimorfismos já descritos, onde se encontra mencionado o que foi analisado neste estudo – A1020P .....	48
<b>Quadro 18</b> – Resultados obtidos para as variações de sequência estudadas .....	54
<b>Quadro 19</b> – Avaliação quanto à probabilidade destas variações serem polimorfismos ou mutações .....	54

### Introdução

A arteriopatía cerebral autossómica dominante com enfartes subcorticais e leucoencefalopatia (CADASIL, MIM#125310) é uma doença vascular cerebral hereditária de manifestação tardia. Foi descrita pela primeira vez, em 1977, na Lancet por dois grupos independentes.

Como o próprio acrónimo define, trata-se de uma doença de hereditariedade autossómica dominante. A prevalência da doença de CADASIL é desconhecida. Num estudo populacional realizado, na área ocidental da Escócia, foi determinada uma prevalência de 1,98 casos por 100 000 habitantes. Um estudo recente elaborado pela Orphanet, refere que apenas haviam sido descritas 500 famílias com esta patologia. Em 1996, foi clonado o único gene associado à doença, o gene *NOTCH3*, localizado em 19p13.

O quadro clínico da doença de CADASIL é caracterizado fundamentalmente por episódios isquémicos tipo-enfarte, défice cognitivo, enxaqueca com ou sem aura e perturbações neuro-psiquiátricas. É uma doença de manifestação tardia, com início dos primeiros sintomas entre a terceira e a quarta décadas de vida.

O diagnóstico desta situação é sobretudo clínico com ênfase para a história familiar. Pode recorrer-se ao estudo molecular nas seguintes circunstâncias: para teste de confirmação do diagnóstico clínico; para teste preditivo; para diagnóstico pré-natal; e para diagnóstico pré-implantação.

### Objectivo

O objectivo principal deste trabalho foi a determinação da frequência alélica, num grupo controlo, das seguintes variações de sequência: no exão 11 - C568Y, T575T e G594G; no exão 19 - A1020P; W1028C e no intrão 18 - IVS18+73A>G e no exão 20 - C1099Y. Estas tinham sido identificadas no decorrer na avaliação molecular num grupo de indivíduos com hipótese diagnóstica de doença de CADASIL no Serviço de Genética da Faculdade de Medicina da Universidade do Porto. Todas as variações não haviam sido

previamente descritas nas bases de dados internacionais, à excepção do polimorfismo A1020P.

Foi também inferido o potencial de patogenicidade das variações de sequência encontradas, em particular das variações não sinónimas, não só dependendo das frequências encontradas mas também recorrendo a ferramentas de bioinformática.

## Material e Métodos

Foram utilizados amostras de DNA de um grupo controlo existente no Serviço de Genética da Faculdade de Medicina da Universidade do Porto (105 indivíduos na análise das variações de sequência nos exões 11 e 20 e 102 para as localizadas no produto de PCR do exão 19 e intrão 18).

Para as variações de sequência localizada nos exões 11 e 19 e intrão 18, o método utilizado foi a sequenciação directa. Foi realizada a digestão do produto de PCR do exão 20 com a enzima HpyCH4III, para a identificação da alteração localizada neste exão.

Para a análise dos resultados foram determinadas as frequências alélicas para as diferentes variações de sequências e respectivos *odds ratio*. Foi aplicado o método MSRV às variações de sequência não sinónimas.

## Resultados

Nas variações de sequência estudadas no exão 11, foram utilizados 105 controlos, sendo a alteração T575T a que apresentou uma maior frequência em ambos os grupos, apesar de não haver diferenças com significância estatística. De acordo com a base MSRV, a variação C568Y apresenta um valor elevado de estar associado à doença.

A variação de sequência intrónica estudada (IVS18+73A>G) apresentou maior frequência no grupo dos doentes, sem haver contudo diferença estatisticamente significativa. Verificou-se que no grupo controlo esta alteração encontra-se em *cis* com o polimorfismo A1020P.

Das variações estudadas no exão 19, a variação A1020P apresenta uma maior frequência em ambos os grupos (sem haver igualmente diferenças estatisticamente significativas). Na aplicação do sistema MSRV a W1028C apresenta uma maior probabilidade de estar associada a CADASIL.

A única variação de sequência estudada no exão 20 não foi identificada em qualquer controlo e apresenta uma elevada probabilidade de estar associada a esta patologia.

### **Discussão e conclusões**

Poder-se-á afirmar com algum grau de certeza que as variações de sequência C568Y e C1099Y, em igualdade de circunstâncias, e a W1028C são mutações.

As variações de sequência T575T, IVS18+73A>G e G594G necessitarão da realização de estudos adicionais (avaliação de segregação familiar, da conservação e interferência com o processamento de RNA) para se afirmar com maior convicção que são de polimorfismos.

## Introduction

Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL, MIM#125310) is an autosomal dominantly inherited angiopathy that the main signs and symptoms occur late in life. In 1977, two independent groups described it in *The Lancet*.

Its prevalence is still unknown. A population study performed in the west region of Scotland determined a prevalence of 1.98 cases in 100,000 inhabitants. A recent bibliographic study done by Orphanet revealed that only 500 families had been described. In 1996, the only gene, the *NOTCH3*, known associated to this disease was identified in 19p13.

The main clinical features are ischemic episodes, cognitive deficits, migraine with/without aura and psychiatric disturbances. The first signs and symptoms generally begin between the third and fourth decades of life.

The clinical picture mainly does the diagnosis without forgetting to look at the family history. However, one should ask for the molecular study of *NOTCH3* for confirming the clinical diagnosis, as test of at-risk asymptomatic adults, and prenatal and pre-implantation diagnosis.

## Aim

The aim of this work was to determine the allelic frequency in a control group of the following sequence variations: on exon 11 – C568Y, T575T and G595G; on exon 19 – A1020P and W1028C; on intron 18 – IVS18+73; and on exon 20: C1099Y. All these alterations have been identified during the diagnostic evaluation of several patients with the clinical suspicion of CADASIL disease in the Genetics Department of Medical Faculty from the University of Porto. All of these sequence variations had been previously described on the international mutation databases, except the polymorphism A1020P.

Depending on the allelic frequencies found, especially of the nonsynonymous ones, and also by using bioinformatics tools available it was tried to estimate their pathogenic potential.

## Materials and Methods

The DNA samples were obtained from the DNA controls bank from the Genetics Department of the Medical Faculty from the Porto University. For studying the sequence variations on exons 11 and 19 and intron 18, DNA sequencing was the method chosen and 105 individuals were studied for the exon 11 and 102 for the last two regions. The restriction enzyme HpyCH4III was used in 105 controls to detect the sequence variation on exon 20.

It was calculated for each one the frequency and the odds ratio. For the evaluation of the probability of these variations being associated with CADASIL, it was used a bioinformatics tool – MSRV system.

## Results

On exon 11, DNA samples from 105 controls were studied. The sequence variation T575T was the one with a higher frequency in both groups but this difference was not statistically relevant. By using the MSRV system, the variation C568Y showed an elevated score that suggest it could be associated with this disease.

The intronic sequence variation analysed (IVS18+73A>G) has a higher frequency in the patients, even though it was not statistically significant. On the controls, this sequence variation was always found in *cis* with the polymorphism A1020P.

From the sequence variations studied on exon 19, the polymorphism A1020P showed to have the higher frequency on either groups (though without statistical difference). When applying the MSRV system, the W1028C alteration has a greater probability of being associated with CADASIL.

The only variation investigated on exon 20 was not identified in any control. By using the bioinformatics tool, it was demonstrated that it has a high likelihood of being related with this disease.

## Discussion e conclusions

With some degree of certainty, the sequence variations C568Y and C1099Y (with the same likelihood) and W1028C are mutations.

The variants T575T, IVS18+73A>G and G594G will need additional studies (evaluation of the familial segregation, conservation and splicing



interference) in order to affirm with more conviction that they are polymorphisms.

## **1. Introdução**

---

## 1.1. Perspectiva histórica

A doença de CADASIL [MIM#125310] é também conhecida pelas designações de demência hereditária com multi-enfartes [1], encefalopatia vascular familiar crónica [2] e demência subcortical familiar [3].

Uma forma de demência com multi-enfartes foi descrita pela primeira vez, em 1977, por Sourander e Walinder [4]. Neste artigo, os autores apresentaram uma família com três gerações afectadas por doença neuropsiquiátrica com sinais e sintomas cerebelares, bulbares e piramidais com uma evolução progressiva, culminando numa forma de demência grave. Foram também publicados os resultados de estudos anátomo-patológicos de alguns indivíduos onde eram evidentes alterações a nível vascular. Esta descrição indiciou já que esta situação parecia ser uma doença de manifestação tardia, com os primeiros sinais a evidenciarem-se na vida adulta jovem. Os autores conjecturaram uma etiologia genética e presumiram ter um padrão de hereditariedade autossómico dominante.

Posteriormente, em 1987, Sonninen e Savontaus [5] descreveram mais dezasseis doentes pertencentes a uma mesma família. Tinham um quadro clínico semelhante à família descrita uma década antes por Sourander e Walinder. Estes autores reafirmaram tratar-se de uma doença de manifestação tardia com um padrão autossómico dominante, caracterizada, em termos imagiológicos, por enfartes cérebro-vasculares oclusivos a nível da substância branca.

Foi apenas no início da década de 90, que esta patologia volta a chamar a atenção da comunidade científica, com a descrição de novos casos [6]. Em 1993, o acrónimo CADASIL (que em inglês, corresponde à seguinte denominação: *cerebral, arteriopathy, autosomal dominant, with subcortical infarcts and leukoencephalopathy* [em português: arteriopatia cerebral autossómica dominante com enfartes subcorticais e leucoencefalopatia]) é introduzido quando é mapeada no cromossoma 19 [7]. Três anos depois o gene *NOTCH3* foi identificado como o gene causal da doença pelo mesmo grupo [8].

## 1.2. Epidemiologia

---

A doença de CADASIL tem uma distribuição mundial, existindo publicadas algumas centenas de famílias [9] sendo este número referido como de 500 num estudo bibliográfico da Orphanet (<http://www.orpha.net/>).

A prevalência da doença de CADASIL é desconhecida. Num estudo populacional realizado, na área ocidental da Escócia, foi determinada uma prevalência de 1,98 casos por 100 000 habitantes [10]. Como é característico das doenças autossómicas dominantes, tem uma expressão variável intra e inter-familiar muito grande. Este facto pode não só dificultar o diagnóstico como também, em determinadas circunstâncias, conduzir a um diagnóstico incorrecto.

## 1.3. Clínica

---

O quadro clínico da doença de CADASIL é caracterizado fundamentalmente por episódios isquémicos tipo-enfarte (acidentes vasculares cerebrais ou acidentes isquémicos transitórios), défice cognitivo, enxaqueca com ou sem aura e perturbações neuro-psiquiátricas. É uma doença de manifestação tardia, com início dos primeiros sintomas entre a terceira e a quarta décadas de vida [11].

Não existem critérios consensuais para o diagnóstico desta patologia. Em 1998, Davous [12] publicou uma proposta que nunca foi utilizada correntemente na prática clínica [**Quadro 1**].

A imagiologia do sistema nervoso central desempenha um papel importante no diagnóstico. A presença de anomalias do sinal na ponderação T2 na substância branca na região temporal e na cápsula externa são muito características [13, 14].

Para além destas características, há que investigar exaustivamente a história familiar. Um padrão de transmissão autossómico dominante é quase sempre verificado [10]. Contudo, há casos reportados de indivíduos com doença de CADASIL, com confirmação molecular por estudo do gene

*NOTCH3*, em que não foi identificada qualquer alteração nos progenitores [15, 16], ou seja, onde ocorreram mutações *de novo*.

**Quadro 1 – Critérios de diagnóstico propostos por Davous.**

**CADASIL PROVÁVEL**

- (1) Idade de início jovem (<50 anos)
- (2) Pelo menos dois dos seguintes achados clínicos:
  - Episódios tipo enfarte com sinais neurológicos permanentes
  - Alterações *major* do humor
  - Demência subcortical
- (3) Sem factores de risco vascular relacionados com o défice
- (4) Evidências de um padrão de transmissão autossómico dominante
- (5) Ressonância magnética da substância branca sem enfartes corticais

**CADASIL DEFINITIVO**

Critérios de CADASIL provável associados a estudo de ligação génica ao cromossoma 19 (mutações no gene *NOTCH3*) e/ou achados patológicos com arteriopatía dos pequenos vasos com material osmiofílico granular

**CADASIL POSSÍVEL**

- (1) Idade de início avançada (>50 anos)
- (2) •Episódios tipo enfarte sem sinais neurológicos permanentes
  - Alterações *minor* do humor
  - Demência global
- (3) Factores de risco vascular *minor*, tais como hipertensão ligeira, hiperlipidémia ligeira, hábitos tabágicos e uso de anti-concepcionais orais
- (4) História familiar incompleta ou desconhecida
- (5) Ressonância magnética da substância branca atípica

**CRITÉRIOS DE EXCLUSÃO**

- (1) Idade de início > 70 anos
- (2) Hipertensão grave ou doença vascular sistémica ou cardíaca complicada
- (3) Ausência de outro caso documentado na família
- (4) Ressonância magnética normal realizada > 35 anos

Adaptado de Davous, 1998.

**1.3.1. Acidentes isquémicos tipo-enfarte**

Os acidentes vasculares cerebrais (AVC) e os acidentes isquémicos transitórios (AIT) são a forma mais frequente de apresentação, ocorrendo em cerca de 70 a 85% dos indivíduos sintomáticos [17, 18]. Têm na maior parte dos indivíduos uma elevada probabilidade de recorrência, levando ao aparecimento de sequelas incapacitantes (perturbações da marcha, culminando na perda de deambulação; incontinência urinária e paralisia pseudobulbar).

Clinicamente, a maioria dos casos apresenta-se como uma síndrome lacunar clássica (AVC motor puro, hemiparésia atáxica, síndrome de mão-

desajeitada e disartria, AVC sensorial puro ou AVC motor e sensorial) ou como outra forma de síndrome lacunar [19].

A idade média de início destes episódios é por volta dos 46 anos, com intervalo que pode variar entre os 19 e os 67 anos [20]. O estudo retrospectivo realizado em 411 doentes por Opherck *et al.* não identificou diferenças relevantes entre o sexo feminino e masculino (homens 51 anos e mulheres 53 anos) [9].

### **1.3.2. Défice cognitivo**

O défice cognitivo é reconhecido em cerca de 60% dos doentes, podendo evidenciar-se desde idades jovens (a partir dos 35 anos). Atinge várias áreas e em cerca de 75% dos doentes pode progredir até à demência [9, 18]. Aparece cerca de dez a vinte anos após o estabelecimento dos episódios isquémicos silenciosos e recorrentes [9, 17, 21]. Tem uma evolução lenta mas progressiva [22].

A demência na doença de CADASIL é do tipo subcortical com disfunção do lobo frontal, apragmatismo, apatia disfunção executiva e alteração da memória declarativa. A linguagem é uma função bem preservada [23]. Inicialmente tem um perfil heterogéneo, mas após os 60 anos torna-se mais homogéneo afectando todas as áreas [24]. Manifesta-se inicialmente por défice de execução seguindo-se perturbações da memória visuo-constructiva [24, 25]. Com a idade outras disfunções cognitivas vão-se tornando cada vez mais evidentes [26].

Dois estudos publicados recentemente [27, 28] concluíram que o défice cognitivo está associado sobretudo com o volume das lesões lacunares, micro-hemorragias cerebrais, pressão sistólica, e com a idade mas não com as hemorragias da substância branca.

### **1.3.3. Perturbações psiquiátricas**

As perturbações do humor ocorrem em cerca de 20% dos casos [17]. Podem encontrar-se outros distúrbios, tais como doença bipolar, ataques de pânico, síndrome alucinatória, esquizofrenia e episódios de ilusões [29-31].

#### **1.3.4. Enxaqueca**

Cerca de 30-40% dos doentes, têm queixas de enxaqueca e, em cerca de 40% dos casos, pode ser este o primeiro sintoma [17, 18, 32]. Pode manifestar-se por volta dos vinte anos e, em algumas famílias, pode ser o sinal mais marcante e debilitante. Em cerca de 90% dos casos, há enxaqueca com aura [18]. Esta forma de enxaqueca pode apresentar desde uma aura típica (visual) até uma forma com uma maior envolvimento sensorial e motor; ser uma forma de enxaqueca hemiplégica ou uma aura prolongada [17, 33, 34].

#### **1.3.5. Achados menos frequentes**

##### **1.3.5.1. Epilepsia**

A epilepsia é descrita em 5 a 10% dos doentes com este diagnóstico. Podem ser uma manifestação primária da doença ou secundária às lesões isquémicas [18, 32].

##### **1.3.5.2. Quadros de encefalopatia aguda**

Quadros de encefalopatia aguda foram reportados em alguns doentes associados a confusão, cefaleias, epilepsia e coma e, em alguns casos, podendo mesmo levar à morte [35, 36].

##### **1.3.5.3. Doença vascular cardíaca**

A associação com doença de vascular cardíaca é controversa. Num estudo realizado na população holandesa doente, cerca de 25% dos indivíduos com diagnóstico confirmado por estudo molecular do gene *NOTCH3*, tinham tido história de enfarte agudo do miocárdio e/ou alterações patológicas da onda Q no electrocardiograma [37]. Um estudo posterior de 2006, mostra contudo resultados contraditórios. Nos 22 doentes estudados, não houve diferenças nos sinais de enfarte ou isquémia, distúrbios da condução ou arritmias relativamente ao grupo controlo [38].

##### **1.3.5.4. Neuropatia periférica**

Na maioria dos casos, são formas assintomáticas. Mas identificam-se sinais de degeneração dos pequenos vasos condicionando sinais de isquémia

crônica de baixo grau com hipoxidose secundária e alterações estruturais no músculo esquelético e nos nervos periféricos [39].

Noutros doentes foram encontrados sinais de neuropatia sensório-motora (em sete dos onze doentes estudados), com sinais axonais e desmielinização [40].

### 1.3.5.5. Anomalias oculares

As anomalias oculares observadas são sobretudo vasculares, a nível das arteríolas retinianas, conduzindo, no entanto, a perturbações funcionais mínimas contrastando com a gravidade das lesões arteriolas da substância branca cerebral [41].

## 1.4. Evolução e prognóstico

A doença de CADASIL tem uma evolução progressiva e uma fase final muito debilitante e incapacitante. Existe uma grande variabilidade inter e intra-familiar. Há a descrição de indivíduos que permaneceram assintomáticos até à sétima década de vida [9]. Na fase final da doença, a maioria dos doentes apresentam demência com paralisia pseudobulbar e incontinência esfinteriana [**Quadro 2**].

**Quadro 2** – Manifestações clínicas e complicações na fase terminal.

Sinais e sintomas	Número de doentes (%)
Tetra/hemiparésia espástica	55/62 (89)
Demência	49/64 (77)
Disartria	47/61 (77)
Disfagia	41/56 (73)
Com necessidade de gastrostomia	18/59 (31)
Riso ou choro patológicos	30/56 (54)
Incontinência urinária	48/60 (80)
Com necessidade de cateterização supra-púbica	12/58 (21)
Incontinência fecal	42/58 (72)
Úlceras de decúbito	18/55 (33)
Cuidados de enfermagem permanentes	51/65 (78)
Doentes acamados	41/65 (63)

Adaptado de Opherk *et al.*, 2004.



No estudo retrospectivo de avaliação do seguimento a longo prazo de 411 doentes, constatou-se que a idade da morte ou em que um indivíduo afectado fica acamado é significativamente inferior nos homens do que nas mulheres; a sobrevivência geral é inferior em relação à população em geral (mas apenas significativamente importante nos indivíduos do sexo masculino); à idade da morte 75% dos indivíduos estava dependente da assistência de terceiros [Quadro 3]. Também se verificou que a principal causa de morte foi a pneumonia (38%) [9].

**Quadro 3** – Idade de início dos AVC, imobilização e morte.

	Idade (IC 95%)	
	Homens	Mulheres
AVC	50,7 (48,2-53,1)	52,5 (50,0-54,9)
Incapacidade de andar sem assistência	58,9 (56,6-61,3)	62,1 (59,7-64,4)
Doentes acamados/ imobilizados	62,1 (59,7-64,4)	66,5 (63,9-69,1)
Morte	64,6 (61,7-67,6)	70,7 (74,7-82,4)

Adaptado de Opherck *et al.*, 2004.

A gravidez pode ser um período de agravamento da doença. Isso foi constatado noutro estudo retrospectivo com 25 mulheres, publicado em 2005, onde se verificou a existência de uma maior frequência de acidentes vasculares (em doze das 25), particularmente quando a idade era superior a 30 anos e, especialmente, durante o puerpério [42].

## 1.5. Imagiologia

### 1.5.1. Ressonância magnética cerebral

A ressonância magnética do sistema nervoso central (RM) tem um papel importante no diagnóstico e no seguimento na doença de CADASIL. As lesões do sistema nervoso central têm uma idade de início e um ritmo de progressão muito variável [17, 43]. Podem ser identificadas mesmo em indivíduos assintomáticos [11].

Na maioria dos doentes, estas lesões têm uma localização na substância branca subcortical, mas também podem ser encontradas noutras localizações do sistema nervoso central (tronco cerebral e substância cinzenta subcortical) [11].

### 1.5.1.1. Hiperintensidade na substância branca

Estas alterações podem ser uma consequência normal do envelhecimento, contudo há que as contextualizar de acordo com a idade e a história familiar. Outro factor a ter em consideração, é o seu padrão evolutivo [44, 45].

Têm uma distribuição simétrica e uma localização profunda e periventricular na substância branca [11]. O lobo frontal é onde se verificam o maior número de lesões, seguido dos lobos temporal e parietal [46]. Apesar de altamente sensível e específico, o envolvimento da cápsula externa e dos lobos temporais nem sempre se verifica [45]. Mesmo nos portadores assintomáticos podem ser identificadas pequenas hiper-intensidades em T2 na substância branca periventricular e subcortical [22].

### 1.5.1.2. Lesões lacunares

Os enfartes lacunares são pequenos enfartes cerebrais profundos, devido provavelmente à oclusão de ramos arteriais perfurantes. Na ponderação T1, apresentam-se como focos punctiformes e hipodensos. A probabilidade de serem diagnosticados aumenta com a idade [11], podendo ser encontrados a partir da quarta década de vida [44]. Localizam-se mais frequentemente na substância branca profunda, podendo também ser encontrados nos gânglios da base, tálamo, cápsulas interna e externa e no tronco cerebral [44].

Estas lesões podem não existir em cerca de 20 a 30% dos doentes com alterações da substância branca [11, 44].

### 1.5.1.3. Enfarte lacunares subcorticais

Os enfartes lacunares subcorticais são agrupamentos lineares de lesões circulares circunscritas que se localizam junto ao córtex, na transição substância branca-cinzenta. Apresentam um sinal na RM semelhante ao LCR [44].

### 1.5.1.4. Micro-hemorragias cerebrais

As micro-hemorragias cerebrais são diagnosticadas em cerca de 31 a 69% dos doentes. São pequenas lesões (dois a cinco milímetros de diâmetro) focais ou multifocais e correspondem a depósitos de hemossiderina

resultantes de hemorragias subclínicas. Na ponderação T2 ou em eco-gradiente são observadas como imagens arredondadas e escuras que são sensíveis ao ferro [47, 48].

Um estudo recente veio sugerir que estas micro-hemorragias estariam associadas a deterioração neurológica e que o tratamento da hipertensão arterial e da hiperglicémia seriam benéficos para o retardar do declínio funcional observado nesta doença [49].

#### **1.5.1.5. Atrofia cerebral**

A atrofia cerebral é um sinal de progressão de doença encontrado frequentemente e parece estar significativamente relacionada com o déficit cognitivo e funcional [50]. Um estudo do corrente ano parece demonstrar que a atrofia cerebral é uma consequência das lesões lacunares e de outras alterações micro-estruturais [51].

### **1.6. Técnicas de anatomia-patológica**

#### **1.6.1. Biopsia de pele**

A alteração paradigmática da doença de CADASIL é a microangiopatia caracterizada pela degeneração e eventual desaparecimento das células da musculatura lisa vascular e pelos depósitos de material osmiofílico granular (GOM) nas vizinhanças da membrana basal. A destruição grave destes vasos leva à hipotonia das arteríolas cerebrais e a danos hemodinâmicos responsáveis quer pela clínica, quer pelas alterações observadas na RM. A acumulação deste material também ocorre em vasos de outros tecidos e órgãos [52, 53].

Esta análise é feita por microscopia electrónica, geralmente, em biopsias cutâneas, mas também podem ser de biopsias musculares e de nervos periféricos [54]. Apesar de altamente específicos, têm uma sensibilidade que varia entre os 45 e 90% [52, 53, 55].

Estes depósitos ainda não foram completamente caracterizados. No entanto, sabe-se que não são depósitos de amilóide, ateroscleróticos, minerais ou metálicos. Os estudos de imunofluorescência realizados foram negativos.

Num estudo publicado recentemente foi demonstrada a presença de material imunorreactivo anti-notch3 (o ectodomínio da proteína Notch3) como um dos principais componentes dos GOM. Concluíram também que a presença de GOM em biopsias musculares era altamente específico de CADASIL e sugeriam fortemente a presença de uma mutação no gene *NOTCH3* [54]. Já no presente ano, um grupo italiano chegou a conclusões semelhantes, mas reforçou a importância do estudo molecular no estabelecimento de um diagnóstico definitivo [53].

Podem também realizar-se estudos de imuno-histoquímica, que estão descritos como altamente sensíveis e específicos e tecnicamente mais simples do que a microscopia electrónica. No entanto, ainda não estão disponíveis comercialmente [56].

### 1.6.2. Estudos anátomo-patológicos

Os estudos anátomo-patológicos de cérebros e de outros tecidos de doentes de CADASIL confirmaram a presença de enfartes cerebrais e degeneração das paredes arteriolas com depósitos de GOM nas proximidades das células da musculatura lisa vascular. Estes estudos indicaram que a doença de CADASIL tem um envolvimento multissistémico, atingindo também a pele, os nervos periféricos e o coração. Os depósitos de GOM no coração, músculo, pele e noutros órgãos precedem a destruição da musculatura lisa vascular [57-59].

## 1.7. Diagnóstico

O diagnóstico de doença de CADASIL deve ser considerado em qualquer doente com enxaqueca (com ou sem aura), acidentes vasculares ou distúrbios do humor e em quem a imagiologia do sistema nervoso central demonstre a existência de anomalias da substância branca (especialmente na cápsula externa e na zona anterior do lobo temporal). A história familiar pode na maior parte dos casos desempenhar um papel determinante [Figura 1].

Nos indivíduos com o fenótipo clássico e com achados imagiológicos compatíveis há que considerar a confirmação do diagnóstico quer pelo estudo molecular do gene *NOTCH3*, quer pela realização de uma biopsia de pele.

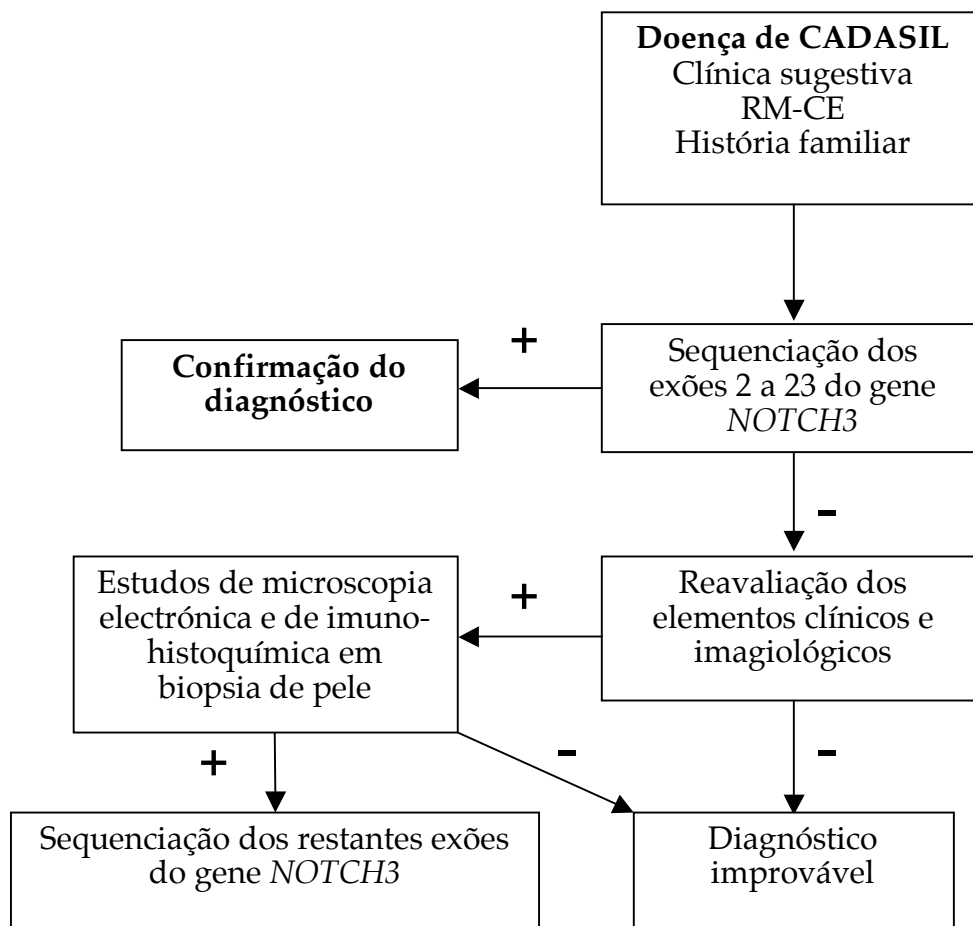


Figura 1 – Sequência de investigação etiológica.

### 1.8. Diagnósticos diferenciais

Nos diagnósticos diferenciais da doença de CADASIL, estão incluídas todas as doenças que apresentem semelhanças clínicas e das características na RM. Assim sendo, deve pensar-se em doença de Binswanger, esclerose múltipla e aneuria primária do sistema nervoso central [22, 32].

Devem ser também consideradas outras doenças menos frequentes, tais como a hiper-homocisteinémia, dislipoproteinémias hereditárias, doenças trombóticas hereditárias, doença de Fabry, adrenoleucodistrofias, MELAS

(*mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes*), [60-62], e doença de CARASIL (*cerebral autosomal recessive arteriopathy with subcortical leucoencephalopathy*) [63].

A clínica, o padrão de hereditariedade e o resultado dos exames auxiliares de diagnóstico podem ser determinantes no estabelecimento do diagnóstico correcto.

### **1.9. Tratamento e seguimento**

Não existe nenhuma terapêutica eficaz. Estes doentes são geralmente vigiados em consultas de neurologia, mas com a progressão da doença há necessidade de ter uma equipa multidisciplinar, onde o apoio psicológico do doente e dos seus familiares não deve ser esquecido.

Assim, o tratamento proporcionado é meramente o alívio sintomático. As enxaquecas são tratadas com analgesia. A eficácia e a segurança dos triptanos são incertas. As crises frequentes e incapacitantes podem necessitar de terapêutica farmacológica profiláctica a longo prazo. A depressão e os distúrbios de humor são tratados com os antidepressivos convencionais. As crises de labilidade emocional beneficiam com a administração de inibidores da recaptção da serotonina [64].

Há quem prescreva terapêuticas anti-coagulante ou anti-agregante plaquetar com o intuito de prevenir os acidentes isquémicos cerebrais, contudo a sua utilização é controversa. Alguns estudos parecem contraindicá-la, uma vez que mostraram serem potenciais factores de risco de hemorragias cerebrais [48]. Os estudos para avaliar a eficácia das estatinas na disfunção endotelial não mostraram quaisquer benefícios [23].

A conduta actual tem como principal preocupação o controlo agressivo de outros factores de risco cérebro-vasculares como a hipertensão, diabetes, dislipidémias e dos hábitos tabágicos [22, 49].

## 1.10. Genética

### 1.10.1. *NOTCH3* (MIM 600276) e Via de Sinalização Notch

A doença de CADASIL, por estudos de ligação génica, foi mapeada no cromossoma 19q12 em 1993. Neste estudo, as melhores estimativas para a localização do gene afectado foram um intervalo de 14 centimorgan flanqueado pelos *loci* D19S221 e D19S222 [65]. Três anos mais tarde, o mesmo grupo descobriu o gene *NOTCH3* nesse *locus* como sendo o responsável por esta patologia [8]. Este gene localiza-se em 19p13.2-p13.1 e tem 33 exões. Nos adultos, o *NOTCH3* é expresso nas células do músculo liso [59]. A única patologia associada a mutações neste gene é a doença de CADASIL.

A via de sinalização Notch é possivelmente a via de sinalização mais utilizada no desenvolvimento animal. Este facto é comprovado também pelo grau de homologia que existe desde os nemátodos até aos seres humanos [59]. Existem pelo menos quatro genes *NOTCH*. São altamente conservados entre espécies e têm elevado grau de homologia. Codificam para receptores celulares e apresentam um domínio de ligação extra-celular, um domínio único trans-membranar e um domínio intra-celular para a transdução de sinal [66]. Dependendo da interacção com outras vias de sinalização e com outros elementos celulares podem influenciar processos de diferenciação, proliferação ou de apoptose [66, 67].

Disfunções nesta via foram já descritas na patogénese de inúmeras doenças humanas, incluindo na leucemia linfoblástica aguda de células T (*NOTCH1*) [68], na carcinogénese [69], na síndrome de Alagille (MIM#118450) (*Jagged 1 – JAG1*) [70], na disostose espondilocostal (MIM#277300) (*Delta-like 3 – DLL3*) [71, 72] e na doença de CADASIL (*NOTCH3*) [8].

O gene *NOTCH3* codifica para uma proteína trans-membranar que desempenha um papel fulcral de sinalização celular. Este receptor é uma proteína de 2700 aminoácidos. A sua constituição é muito semelhante aos outros receptores Notch, excepto no número de repetições tipo factor de crescimento epidérmico (EGF) na parte N-terminal do domínio extra-celular – tem 34 em vez das 36 caracteristicamente existentes nos outros - e por um

domínio citoplasmático mais pequeno [59]. O domínio extra-celular é ainda constituído por três repetições notch/lin-12. A nível citoplásmatico existem seis repetições anquirina [19].

A via de sinalização Notch inicia-se com uma reacção de proteólise desencadeada pelo ligando, condicionando que um fragmento intracelular do próprio receptor se transfira até ao núcleo onde regula a expressão de alguns genes. O processamento do receptor Notch sofre uma série de processamentos proteolíticos. Inicialmente ocorre uma clivagem (S1) ao nível do domínio extra-celular no aparelho de Golgi, resultando na heterodimerização do receptor. Na superfície celular, o receptor interage com ligandos, *Jagged* ou *Delta*, que são proteínas trans-membranares expressas nas células justapostas. A união ligando-receptor é regulada por eventos pós-tradução como a glicosilação e outras modificações envolvendo quer o domínio do receptor quer do ligando [73]. A interacção com os ligandos induz uma segunda clivagem (S2) perto da membrana plasmática na região extra-celular. Esta segunda clivagem condiciona uma terceira e última clivagem (S3) que ocorre na parte do receptor na própria membrana plasmática, que é controlada pelo complexo presenilina/ $\gamma$ -secretase. Verifica-se conseqüentemente a libertação do domínio intra-celular que se desloca até ao núcleo, onde actua conjuntamente, com a proteína ligadora de DNA RBP-Jk na regulação da transcrição dos genes alvo, principalmente os genes “*hairy/enhancer of split*” (*HES*) [74].

### 1.10.2. Características das mutações

Foram já descritas mais de 100 mutações diferentes e cerca de 95% são mutações do tipo *missense* [59, 75]. Estas mutações resultam na maioria dos casos em alterações do número de resíduos de cisteína no domínio extra-celular do receptor NOTCH3, indiciando haver a formação de pontes de dissulfeto e alteração da conformação proteica. São portanto mutações de natureza particularmente estereotípica. Há algumas mutações publicadas (cerca de cinco) que não afectam este equilíbrio de resíduos de cisteína. São cinco pequenas deleções (quatro *in frame* e uma *frameshift*) [75, 76]. Louvi *et. al* sugeriram que as mutações “atípicas” (todas aquelas que não afectam os resíduos de cisteína) devem ser rigorosamente estudadas antes que possam



ser designadas como patogénicas [59]. Muitos polimorfismos foram também já descritos na região codificante do gene [75].

O mecanismo causal que leva à patogénicidade destas mutações é desconhecido até ao momento. Especula-se quanto ao efeito sobre a função do receptor (perda ou ganho de função). Estudos na *Drosophila* pareceram demonstrar que a perda (haplo-insuficiência) ou a existência de mais cópias (duplicações) do gene condicionariam anomalias do desenvolvimento [59]. Foi também constatado que os ligandos se poderiam comportar de uma forma haplo-insuficiente. Outro aspecto que tem vindo a ser objecto de consideração é a possibilidade da alteração do número de cisteínas induzir modificações conformacionais do receptor, diminuindo a sua reactividade; os receptores mutados não serem eficazmente removidos da membrana e conseqüentemente haver uma interrupção da via de sinalização. Uma vez que estas mutações ocorrem em zonas do gene onde a diversidade entre espécies é menor, foi sugerido motivarem ganhos de função [59].

Na tentativa de esclarecer estes mecanismos, têm sido usados modelos animais. Contudo, a experiência tem sido pouco informativa, não tendo sido possível, na maioria dos casos, reproduzir um fenótipo [59].

### 1.10.3. Estudo molecular na prática clínica

Pode recorrer-se ao estudo molecular nas seguintes circunstâncias: para teste de confirmação do diagnóstico clínico; para teste preditivo; para diagnóstico pré-natal; e para diagnóstico pré-implantação.

O gene é relativamente grande e a sequenciação completa do mesmo é laboriosa e dispendiosa [75]. Em indivíduos com clínica muito sugestiva de doença de CADASIL ou com uma biopsia de pele com alterações específicas, o estudo molecular por sequenciação completa tem uma taxa de detecção que pode variar entre 57 e 96% [55, 77].

Os exões onde foram identificadas a maioria das mutações são, por ordem decrescente de frequência, os exões 4, 3, 5, 6 e 11 [77].

Estão descritas algumas variações geográficas. Na França, Reino Unido e Alemanha o exão 3 é o segundo mais frequente [55, 58, 77]. Na Holanda, identificaram-se muitas mutações no exão 11 [37]. Em 143 indivíduos portugueses com suspeita clínica de doença de CADASIL, os exões onde se diagnosticaram um maior número de mutações, foram no 4 e 11 [78].

#### 1.10.4. Correlação genótipo-fenótipo

Apesar de existirem algumas publicações que fazem a associação entre certas mutações e a clínica, não há evidências seguras para o estabelecimento de uma correlação genótipo-fenótipo [9, 48, 79].

A doença de CADASIL parece seguir também a definição clássica das doenças autossômicas dominantes, em que doentes homocigóticos e heterocigóticos são indistinguíveis clinicamente [80].

#### 1.10.5. Penetrância e expressividade

Tem provavelmente uma penetrância completa e uma expressividade nitidamente muito variável. A idade de início, gravidade da sintomatologia e a forma de progressão têm uma grande variabilidade intra e inter-familiar [20, 81].

#### 1.10.6. Antecipação

Há apenas um caso reportado em que parece haver fenómeno de antecipação. Numa família japonesa onde foi identificada a mutação S180C no exão 4, houve uma maior precocidade da idade de início da doença (acidentes vasculares cerebrais) e na documentação das anomalias do SNC nas gerações mais recentes [30].

### 1.11. Aconselhamento genético

Segundo o geneticista britânico Peter Harper [82]:

*“Genetic counselling is the process by which patients or relatives at risk of a disorder that may be hereditary are advised of the consequences of the disorder, the probability of developing or transmitting it and the ways in which this may be prevented, avoided or ameliorated”.*

O aconselhamento de genético deve ser oferecido a todos os indivíduos e respectivas famílias após o diagnóstico de doença de CADASIL, com o objectivo de cumprir todas as premissas incluídas na definição de Peter Harper [82].

A maioria dos casos publicados tinha história familiar compatível com o padrão de transmissão autossômico dominante verificado nesta patologia, em que pelo menos um dos progenitores era também afectado. Até ao momento, foram descritos apenas dois casos onde não foi encontrada qualquer mutação nos progenitores, ou seja, onde ocorreram mutações *de novo* nos indivíduos doentes [15, 16].

Perante estas evidências, sempre que haja a confirmação molecular pela identificação de uma mutação no gene *NOTCH3* num *probando*, a próxima etapa será o aconselhamento genético e oferecer a possibilidade de realização do estudo molecular dos seus progenitores, irmãos e filhos desde que estes tenham atingido a maioridade. O estudo de indivíduos assintomáticos deverá obedecer a um protocolo especial e deverá ser sempre solicitado por um médico com a especialidade de genética médica. No caso dos indivíduos que apresentem sintomas o estudo molecular - teste de diagnóstico – poderá ser solicitado por qualquer médico.

Quando no indivíduo afectado não se identifica qualquer mutação patogénica e, após a reavaliação do caso, o diagnóstico de CADASIL continuar a ser inequívoco, deve ser oferecida a realização de exames auxiliares de diagnóstico que permitam identificar e confirmar esta patologia nos familiares em risco – a RM crânio-encefálica e a biopsia de pele.

O risco de recorrência para descendentes de um indivíduo afectado (heterozigótico) é de 50%. No caso dos indivíduos homozigóticos [80], esse risco será de 100%.

O risco dos irmãos de um indivíduo doente serem portadores, dependerá do estado de heterozigotia dos seus progenitores. Quando não é possível identificar uma mutação em ambos os progenitores, concluir-se-á por uma mutação *de novo*. Esta será a conclusão mais óbvia, após a exclusão de possibilidade de falsa paternidade ou de adopção. Nestas circunstâncias, o risco de heterozigotia em irmãos será muito baixo, contudo não pode ser excluída a possibilidade de um mosaicismo gonadal em um dos progenitores (situação nunca descrita). Quando um indivíduo é homozigótico, ambos os progenitores serão heterozigóticos, excepto na situação nunca descrita nesta patologia de isodissomia uniparental, conferindo um risco de 50% para heterozigotia, de 25% para homozigotia para o alelo mutado ou heterozigotia composta e de 25% para homozigotia para o alelo normal aos outros irmãos.

### 1.11.1. Diagnóstico pré-sintomático

O protocolo habitualmente utilizado segue as recomendações internacionais para o diagnóstico pré-sintomático na doença de Huntington [83], segundo o diagrama apresentado na **Figura 2**. Deve ser realizado no contexto de uma consulta de aconselhamento genético e com apoio psicológico.

No nosso país, em Janeiro de 2005, foi publicada a Lei N.º 12/2005 que regulamenta o acesso a informação genética e informação de saúde, na qual estão definidas as regras para a realização de testes pré-sintomáticos.

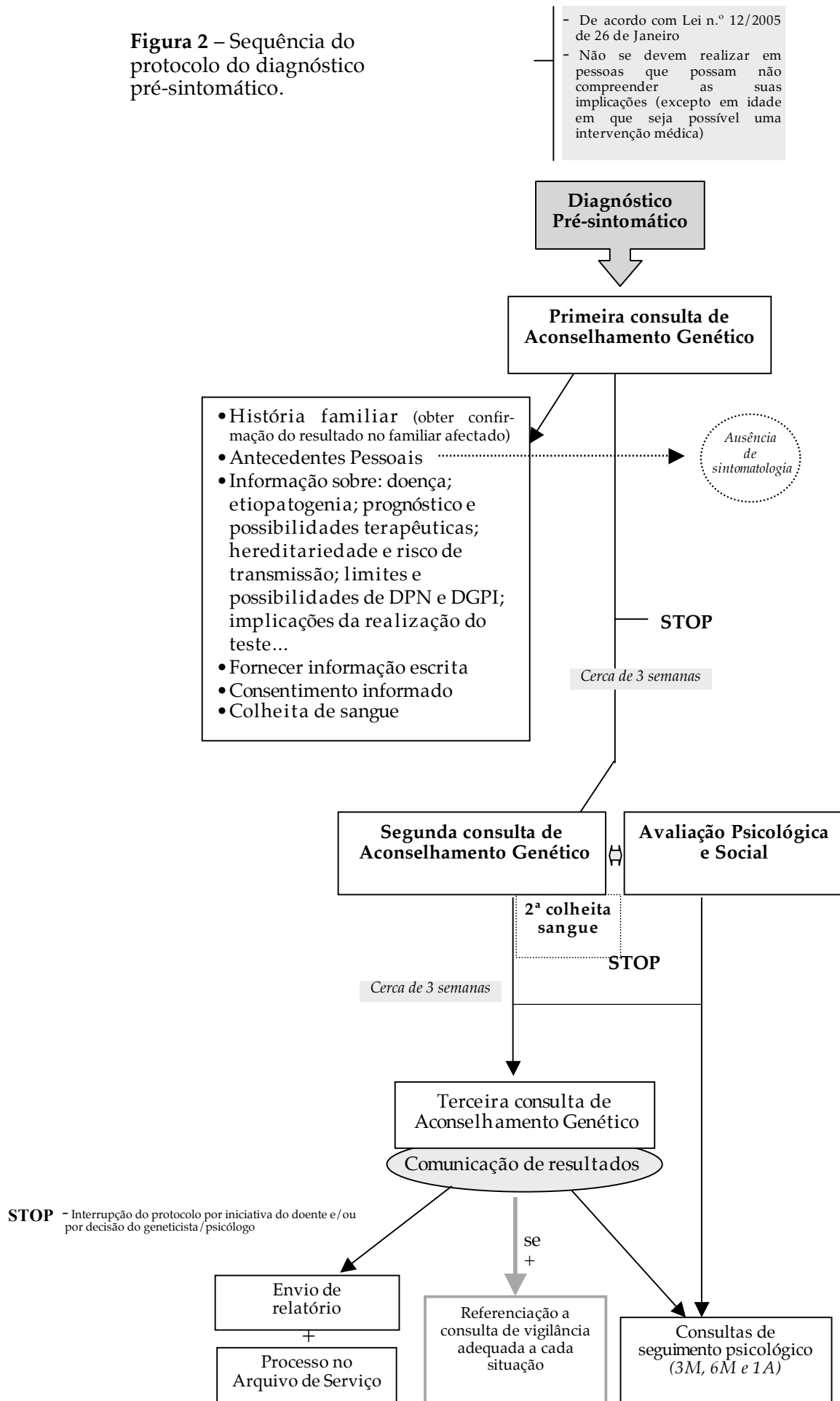
O Artigo 6.º estabelece que, nos serviços de genética, os processos clínicos que contenham informação genética sobre pessoas saudáveis, não poderão ser acedidos ou facultados em qualquer circunstância, mesmo a outro profissional de saúde, excepto com autorização por escrito do próprio.

O Artigo 9.º diz que o diagnóstico pré-sintomático de doenças monogénicas só pode ser efectuado com autorização do próprio e a pedido de um médico com a especialidade de genética na sequência de uma consulta de aconselhamento genético, após consentimento informado escrito. O resultado deste teste só deverá ser comunicado ao interessado e só poderá ser comunicado a terceiros com a sua autorização. Proíbe a realização deste tipo de estudos em pessoas com incapacidade mental que não possam compreender as suas implicações. Salienta ainda a importância de uma avaliação psicológica não só durante o processo de diagnóstico, mas também após a comunicação dos resultados.

No Artigo 11.º está estabelecido o princípio de não discriminação. Qualquer indivíduo que opte pela realização destes estudos não poderá sob forma alguma ser prejudicado (por exemplo: em termos laborais, seguros de vida ou de saúde). Nos artigos subsequentes, 12.º e 13.º, está ainda proibido o acesso a resultado de estudos genéticos por companhias de seguros ou entidades patronais.

A proibição da realização de estudos preditivos a menores, está contemplada no Artigo 17.º, salvo quando sejam efectuados para seu benefício, nomeadamente quando existe a possibilidade de intervenção preventiva ou terapêutica.

**Figura 2** – Sequência do protocolo do diagnóstico pré-sintomático.



### 1.11. 2. Diagnóstico pré-natal

O diagnóstico pré-natal (DPN) é possível pela realização do estudo molecular do gene *NOTCH3* em células fetais, com o objectivo de identificar a mutação presente no progenitor afectado. O material fetal pode ser obtido por intermédio de duas técnicas [Quadro 4]:

- biopsia das vilosidades coriônicas, realizada entre as 10 e as 12 semanas de gestação
- amniocentese, realizada entre as 16 e as 18 semanas de gestação.

**Quadro 4** – Diferenças entre as técnicas de diagnóstico pré-natal disponíveis.

<b>Biopsia das Vilosidades Coriônicas</b>	
<b>Vantagens</b>	<b>Desvantagens</b>
<ul style="list-style-type: none"> <li>- Resultados em idades gestacionais mais precoces</li> <li>- Possibilidade de o material ser estudado directamente sem necessidade de cultura</li> <li>- DNA de qualidade superior</li> </ul>	<ul style="list-style-type: none"> <li>- Experiência limitada em Portugal</li> <li>- Risco de abortamento superior (cerca de 5% dos casos)</li> <li>- Estudo citogenético de pior qualidade e menor fiabilidade</li> </ul>
<b>Amniocentese</b>	
<b>Vantagens</b>	<b>Desvantagens</b>
<ul style="list-style-type: none"> <li>- Estudo citogenético de maior qualidade e fiabilidade</li> <li>- Risco de abortamento (inferior a 1%)</li> </ul>	<ul style="list-style-type: none"> <li>- Obtenção de resultados em idades gestacionais mais tardias</li> </ul>

Estas duas técnicas apresentam as vantagens e desvantagens de que o casal deverá ser informado na consulta de aconselhamento genético, de forma a poder tomar uma decisão consciente e informada [84].

O pedido de interrupção médica de gravidez poderá ser feito por um casal até às 24 semanas de gestação, de acordo com a legislação portuguesa, sendo subsequentemente avaliado por uma Comissão Técnica de Interrupção Médica de Gravidez de um Centro de Diagnóstico Pré-Natal.

Há apenas um caso publicado na literatura de um casal que recorreu ao diagnóstico pré-natal molecular. Foi identificada no feto a mutação

presente no seu pai, tendo o casal pedido a interrupção médica da gravidez [85].

### **1.11.3. Diagnóstico genético pré-implantação**

O diagnóstico genético pré-implantação (DGPI) pode ser oferecido aos casais, desde que haja identificação de mutação no progenitor afectado. Foi recentemente publicado um caso por Konialis *et al.* (2007) [86], onde pela primeira vez foi realizada esta técnica a pedido de um casal com história prévia de uma interrupção médica da gravidez após DPN molecular.

## **2. Objetivos**

---



A doença de CADASIL é uma doença autossómica dominante de manifestação tardia muito debilitante e incapacitante. Como é característico de doenças com este padrão de hereditariedade, tem uma expressão intra e inter-familiar muito variável. Assim é impossível prever para cada indivíduo a idade de início, a sua evolução e o seu prognóstico. É reconhecido tratar-se de uma situação subdiagnosticada.

O estudo molecular para confirmação do diagnóstico está disponível desde 1997 e muito provavelmente só alargadamente utilizável após o ano 2000. Assim sendo ainda não há uma experiência acumulada na avaliação molecular destes doentes. Neste contexto, são identificadas variações de sequência cujo significado requer a realização de estudos adicionais para a clarificação da sua natureza patogénica – mutação *versus* polimorfismo (polimorfismo inclui uma valorização de frequência; uma variação de sequência pode não ser patogénica e não ser um polimorfismo *strictu sensu*). Neste processo há que ter em conta a avaliação de vários parâmetros, onde se incluem o estudo molecular dos progenitores [analisar a segregação familiar], verificar se a região onde ocorrem é conservada entre espécies [avaliar o grau de conservação], determinar a frequência destas alterações num grupo controlo [determinar a frequência alélica], analisar a alteração da proteína correspondente e a realização de estudos funcionais.

O objectivo principal deste trabalho foi determinar a frequência alélica, num grupo controlo, das seguintes variações de sequência: no exão 11: C568Y, T575T e G594G; no exão 19: A1020P; W1028C e no intrão 18: IVS18+73 e no exão 20: C1099Y. Estas tinham sido identificadas no decorrer na avaliação diagnóstica num grupo de indivíduos com hipótese diagnóstica de doença de CADASIL. Todas as variações não haviam sido previamente descritas nas bases de dados internacionais, à excepção do polimorfismo A1020P [75, 87].

Foi também inferido o potencial de patogenezidade das variações de sequência encontradas, em particular das variações não sinónimas, não só dependendo das frequências encontradas mas também recorrendo a ferramentas de bioinformática [MSRV *system* [88]].

### **3. Material e Métodos**

---

### 3.1. Amostra

Sob consentimento informado escrito, foram utilizadas neste estudo amostras de DNA de um banco de controlos existente no Serviço de Genética da Faculdade de Medicina da Universidade do Porto. Todos estes indivíduos eram saudáveis, de raça caucasiana e estudantes daquela Faculdade que se voluntariaram a doar uma amostra de sangue periférico para extracção de DNA [Quadro 5].

**Quadro 5** – Número de controlos estudados de acordo com os exões onde se verificaram as variações de sequência encontradas.

Regiões analisadas	N.º de controlos estudados (N.º total – Sexo Masculino / Sexo Feminino)	N.º de alelos estudados
Exão 11	105 – 41 / 64	210
Exão 19 e intrão 18	102 – 36 / 66	204
Exão 20	105 – 35 / 70	210

### 3.2. Variações de sequência analisadas

No **Quadro 6** expõem-se as variações de sequência investigadas e respectivas metodologias utilizadas na sua análise.

**Quadro 6** – Variações de sequência estudadas e respectiva metodologia de análise.

NOTCH3	Posição DNA		Variação na proteína	Metodologias utilizadas
	Posição relativa (NM_000435)	Variação sequência		
Exão 11	C 1779 G/A C 1801 G/A C 1858 C/T	TGT>TAT ACG>ACA GGC>GGT	C568Y T575T G594G	1.Extracção DNA 2.PCR convencional 2.Sequenciação
Exão 19 e intrão 18	C 3134 G/C C 3160 G/T -	GCC>CCC TGG>TGT A>G	A1020P W1028C IVS18+73	1.Extracção DNA 2.PCR convencional 2.Sequenciação
Exão 20	C 3372 G/A	TGC>TAC	C1099Y	1.Extracção DNA 2.PCR convencional 3.Enzima de restrição HpyCH4III 4.Electroforese 5.Sequenciação

Foi utilizada como referência a sequência do GenBank com número de acesso NM\_000435 [Anexo 1].

### 3.3. Técnicas de Biologia Molecular

#### 3.3.1. Extração de DNA

Procedeu-se ao isolamento dos leucócitos de amostras de sangue periférico colhidas em ácido etilenodiaminotetraacético (EDTA) a 5% por centrifugação (dez minutos a 2200 rpm a 20°C).

O DNA genómico foi extraído usando o método de precipitação salina/clorofórmio previamente descrito por Mullenbach *et al.* em 1989 [89].

#### 3.3.2. Reacção de Polimerase em Cadeia (PCR) Convencional

As condições de PCR foram optimizadas de modo a obter um produto de PCR com ausência de produtos inespecíficos.

Os *primers* usados (Thermo Electron Corporation, Dreieich, Alemanha) [Quadro 7] permitiram amplificar os respectivos exões do gene *NOTCH3* e as regiões flangeadoras, contendo a base implicada na variação de sequência a analisar. Para a apresentação das sequências de bases dos *primers*, utilizou-se como referência a sequência do GenBank com o número de acesso AC004663 [Anexo 2].

**Quadro 7** – Sequências oligonucleotídicas utilizada como *primers* para amplificar os fragmentos do gene *NOTCH3* a analisar.

Designação utilizada	Sequências de bases	Referência AC004663		Tamanho do produto de PCR
CAD11F	5'-ACCATTCCTCAAACCCTCTGT-3'	30199	30218	425bp
CAD11R	5'-TATTGAGCTCGCCTCCTGAC-3'	30624	30605	
CAD18/19F	5'-CCCCATTCGGCTCACACTA-3'	23754	23772	720bp
CAD18/19R	5'-AGGGACCTGATTGGCTTCTG-3'	24474	24455	
CAD20F	5'-ATGCCACCTCCTCTTCC-3'	23187	23204	254bp
CAD20R	5'-GTTCTGTGCCACTCTCCTCT-3'	23440	23420	

A amplificação do DNA genómico por PCR foi realizada utilizando um termociclador GeneAmp® PCR System 9700 (GE Applied Biosystems, Foster City, CA, U.S.A.; n.º de catálogo 4310899), usando 25µl de mistura de reacção

contendo: 1,0µl de DNA (100ng/µl); 2,5µl de *buffer* 10x; 1µl de MgCl<sub>2</sub> (25mM); 0,5µl de uma *mix* de dNTP a 10mMol (100mM dNTP Set, PCR grade, Invitrogen™, Carlsbad, CA, U.S.A.; n.º de catálogo 10297-018); 0,5µl do *primer* directo (F- *forward*) e 0,5µl do *primer* reverso (R-*reverse*); e 0,2µl de Taq DNA Polimerase recombinante (MBI Fermentas, Vilnius, Lituânia; 5 U/µl, n.º de catálogo EP0402). Prefez-se o volume desta mistura com 18,8µl de água destilada (B Braun®). Todos os *primers* foram usados em concentrações de 25pmol/µl.

Para as reacções de PCR envolvendo os exões 11 e 18/19, as misturas foram aquecidas inicialmente a 94°C durante 5 minutos. Posteriormente foram submetidas a 35 ciclos com desnaturação a 94°C durante um minuto, *annealing* a 60°C durante um minuto, extensão a 72°C durante um minuto e trinta segundos e uma extensão final a 72°C durante dez minutos. As condições da reacção de PCR para o exão 20, apenas diferiam das dos outros exões na temperatura de *annealing* utilizada, que foi 62°C.

Os produtos de PCR resultantes das reacções acima descritas foram colocados num gel de agarose a 2% para electroforese e visualizados com brometo de etídeo a 5%.

### 3.3.3 Sequenciação directa dos produtos de PCR

#### 3.3.3.1. Purificação dos Produtos de PCR

Os produtos de PCR obtidos (após verificação da correcta amplificação em gel de agarose como acima descrito) foram purificados, removendo por ultrafiltração o excesso de *primers* e de dNTPs, usando o GFX™ PCR DNA and Gel Band Purification Kit (Amersham Biosciences AB, Uppsala, Sweden; n.º de catálogo 27-9602-01), de acordo com as instruções do fabricante. Dois microlitros de cada produto de PCR purificado são submetidos a electroforese num gel de agarose a 0,8% e visualizados com brometo de etídeo.

#### 3.3.3.2. Reacção de Sequenciação

As reacções de sequenciação foram preparadas utilizando reagentes comerciais (ABI PRISM BigDye Terminator v1.1 Cycle Sequencing Kit, Applied Biosystems, Foster City, CA, U.S.A.), de acordo com as recomendações do fabricante. Para um volume final de solução de 10µl, foram

adicionadas entre 0,5 e 4,0µl de produto de PCR purificado (dependendo da respectiva intensidade da banda no gel de agarose), 2,0µl da solução premix, 1,0µl do *primer* de sequenciação (com uma concentração de 3,2pmol/µl), prefazendo-se, por último, o volume desta mistura com água destilada (B Braun®).

A reacção de sequenciação consistia numa desnaturação inicial de três minutos a 94°C, seguida de 24 ciclos de dez segundos a 96°C, cinco segundos a 50°C e de extensão final de 4 minutos, num termociclador GeneAmp® PCR System 9700 (GE Applied Biosystems, Foster City, CA, U.S.A.; n.º de catálogo 4310899).

### 3.3.3.3. Precipitação dos produtos a sequenciar

Após a reacção de sequenciação, efectuou-se a fase de precipitação que consistia em adicionar ao produto a sequenciar: 25µl de etanol a 100%, 1,0µl de uma solução EDTA a 125mM e 1,0µl de uma solução de acetato de sódio a 3M. Esta mistura ficava a incubar, à temperatura ambiente, durante trinta minutos, seguindo-se uma centrifugação a 14000rpm durante o mesmo tempo, a 4°C. De seguida o sobrenadante era retirado e procedia-se à lavagem do *pellet* com 400µl de etanol a 70%. No passo seguinte, realizava-se nova centrifugação durante 20 minutos a 14000rpm. Por último, o sobrenadante era retirado cautelosamente e procedia-se à secagem do *pellet* à temperatura ambiente ou a 37°C num estufa.

### 3.3.3.4. Sequenciação

Cada *pellet*, no final da fase de precipitação, foi ressuspenso em 20µl de formamida desionizada. Após incubação, durante 30 minutos à temperatura ambiente, ao abrigo de luz directa e antes da transferência da referida mistura para tubos de sequenciação, esta era homogeneizada em vórtex durante 10 segundos.

O sequenciador utilizado foi o ABI PRISM® 310 Genetic Analyser® (GE Applied Biosystems, Foster City, CA, U.S.A.). Utilizou-se para análise das sequências o programa Sequencing Analysis 3.4.1. (fornecido pelo mesmo fabricante).

### 3.4. Análise das variações de sequência localizadas no exão 11

As três variações de sequência estudadas que se localizam neste exão são amplificáveis no produto de PCR deste exão [Quadro 7].

A sequência metodológica aplicada para a sua análise encontra-se esquematizada no Quadro 6 e descrita na Secção 3.2..

### 3.5. Análise das variações de sequência localizadas nos exões 18/19 e intrão 18

As duas variações de sequência em análise neste exão foram estudadas empregando as mesmas técnicas utilizadas para o exão 11. Com os *primers* utilizados também amplificava a zona em estudo do intrão 18.

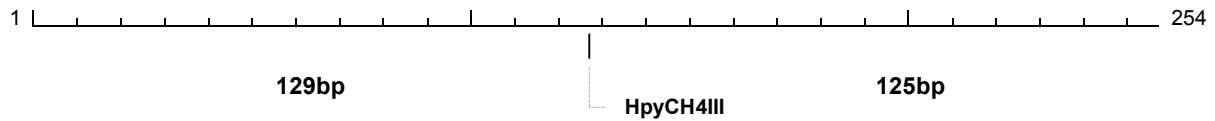
### 3.6. Análise da variação de sequência localizada no exão 20

Na metodologia usada para o estudo da variação de sequência encontrada no exão 20, para além de se terem empregues os dois processos iniciais para os exões anteriores, foi também utilizada a enzima de restrição HpyCH4III (New England Biolabs® Inc, Ipswich, MA, U.S.A., código do produto: R0618L).

A variação de sequência induz um local de corte adicional reconhecido por esta enzima, de acordo com o seguinte sequência palindrômica:

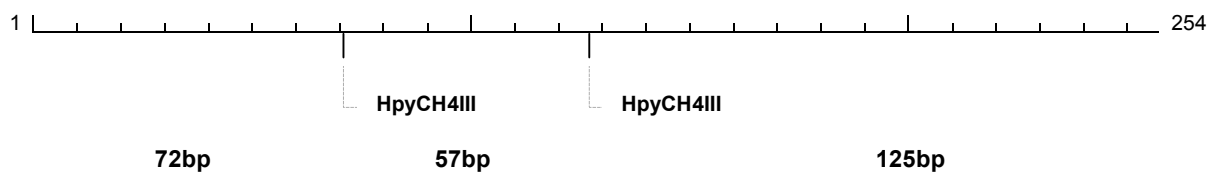


O produto de PCR resultante tem um tamanho de 254bp, que na sequência normal tem um local de corte reconhecido por esta enzima o que origina dois fragmentos um de 129bp e outro de 125bp [Figura 3].



**Figura 3** – Representação esquemática da acção da enzima HpyCH4III no produto de PCR do exão 20 na ausência de variação de sequência.  
[Adaptado de NEBcutter V2.0 da New England Biolabs® Inc]

Na presença da variação de sequência, há o reconhecimento de dois locais de corte para a enzima HpyCH4III, levando à formação de três fragmentos de 125bp, 72bp e 57bp. Para análise dos controlos com o intuito de investigar esta variação de sequência, é portanto exequível diferenciar a sua presença ou ausência [Figura 4].



**Figura 4** – Representação esquemática da acção da enzima HpyCH4III no produto de PCR do exão 20 na presença de variação de sequência.  
[Adaptado de NEBcutter V2.0 da New England Biolabs® Inc]

O processo foi previamente testado e otimizado para o doente com a hipótese de diagnóstico de CADASIL, onde havia sido encontrada esta variação de sequência nunca antes descrita.

Posteriormente, como intuito de confirmar e clarificar alguns dos casos estudados por esta metodologia onde foi constatada ausência de digestão por esta enzima ou resultados que suscitassem dúvidas, foi feita a análise por sequenciação directa.

### 3.6.1. Digestão com a enzima HpyCH4III

Os produtos de PCR foram obtidos das misturas iniciais de reacção de PCR convencional usando os *primers* acima referidos para o exão 20 [Quadro 6] (e após verificação da correcta amplificação em gel de agarose como acima descrito).



O volume total da reacção de digestão foi de 25µl, sendo 20µl de produto de PCR e 5µl de uma mistura composta por: 1µl da enzima HpyCH4III (New England Biolabs® Inc, Ipswich, MA, U.S.A., código do produto: R0618L); 2,5µl do tampão fornecido pelo fabricante – 1x NEBuffer 4 ((New England Biolabs® Inc, Ipswich, MA, U.S.A.); e 1,5µl de água destilada (B Braun®).

A mistura total da reacção foi incubada a 37°C durante no mínimo 12 horas. Foi sempre usado como controlo de digestão o produto de PCR de um indivíduo previamente conhecido como heterozigoto para a variação de sequência C1099Y. Além do controlo de digestão e do marcador de peso molecular, em cada linha de poços foi colocado um fragmento não digerido.

Os fragmentos de DNA foram submetidos a electroforese em gel de agarose Nusieve®3:1 (BMA Products, Rockland, ME, USA) a 3% durante cerca de 45 a 60 minutos a 110V e visualizados usando brometo de etídeo.

### **3.7. Análise de dados**

Para a análise das variações de sequência encontradas foi determinado para cada uma delas a frequência no grupo controlo estudado e nos casos que foram referenciados para diagnóstico ao Serviço de Genética da Faculdade de Medicina da Universidade de Porto. Foi também realizado o cálculo do *odds ratio* com um intervalo de confiança de 95%.

Para a análise das variações de sequência não sinónimas foi ainda aplicada a ferramenta de bioinformática (MSRV) apresentada no American Journal of Human Genetics por Jiang *et. al.* e que se encontra disponível na internet (<http://msms.usc.edu/msrv/index.html>) [88].

Este modelo baseia-se nos conhecimentos biofísicos e bioquímicos associados a métodos estatísticos para estabelecer escalas de probabilidade de causalidade de variações de sequência e determinadas patologias. Para isso foram utilizadas 26 características que incluem as informações das sequências proteicas, alinhamentos de sequências múltiplos conjugados com um modelo [*multiple selection rule voting (MSRV)*] projectado para capturar características subtis das mutações de cada tipo de aminoácido. Estas características baseiam-se em três propriedades físico-químicas (peso molecular, valor do

ponto isoelétrico (pI) e escala de hidrofobicidade) dos aminoácidos, nas três frequências relativas da ocorrência dos aminoácidos na estrutura secundária das proteínas com informação estrutural conhecida e nos dois valores de conservação evolutiva.

## **4. Resultados**

---

Os resultados das experiências efectuadas no grupo controlo encontram-se seguidamente discriminados de acordo com a localização do exão e/ou intrão onde foram identificadas as variações de sequência em estudo.

#### **4.1. Resultados da análise das variações de sequência localizadas no exão 11**

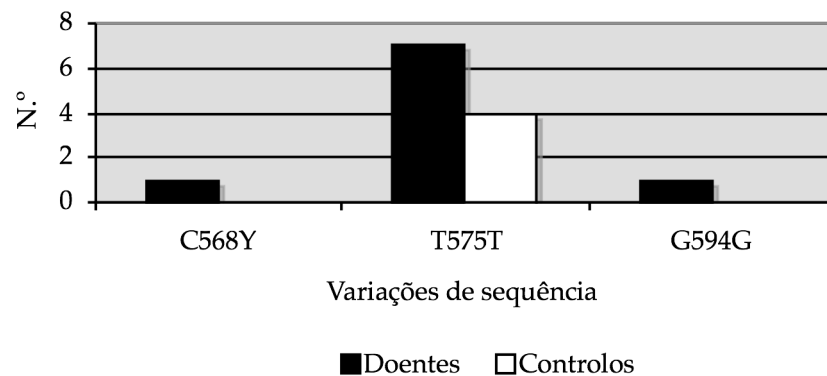
No produto de PCR deste exão, foram estudadas três variações de sequência: C568Y, T575T e G594G. As variações de sequência T575T e G594G são do tipo *silence*, enquanto que a C568Y é *missense*. Destas variações estudadas, apenas a C568Y envolve a substituição de um resíduo de cisteína por um de tirosina na proteína.

Utilizaram-se amostras de DNA de 105 indivíduos [64 (61%) do sexo feminino e 41 (39%) do sexo masculino] como controlos, tendo-se consequentemente estudado 210 alelos normais.

O Serviço de Genética da Faculdade de Medicina da Universidade do Porto na sua actividade diagnóstica identificou num conjunto de 200 doentes com hipótese de diagnóstico de CADASIL, um doente com a variação de sequência C568Y, 7 doentes com a alteração T575T e um com G594G.

Nos 210 alelos controlo avaliados, as variações de sequência C568Y e G594G não foram identificadas em nenhum, mas a variação T575T foi detectada em quatro dos controlos (C9, C55, C56 e C69) [Anexo 3]. Destes controlos três eram do sexo masculino e apenas um do sexo feminino.

A variação de sequência T575T foi identificada em maior número, quer no grupo dos indivíduos doentes, quer nos controlos [Figura 5].



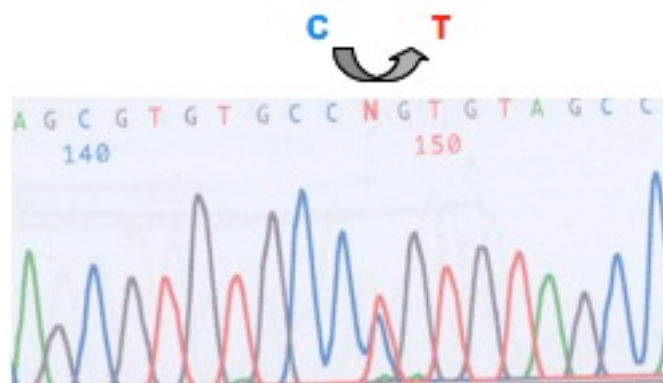
**Figura 5** – Distribuição das variações de sequência do exão 11 nos doentes e nos controlos.

No **Quadro 8**, apresentam-se as frequências das variações de sequência encontradas.

**Quadro 8** – Determinação das frequências das variações de sequência estudadas no exão 11, nos controlos e doentes.

Variações de sequência no exão 11	Frequências alélicas dos controlos	Frequências alélicas dos doentes
C568Y	0,00	0,01
T575T	0,02	0,04
G594G	0,00	0,01

A variação de sequência com maior frequência, em ambos os grupos, é a T575T (0,02 nos Controlos e 0,04 nos Doentes 0,04) [Figura 6].



**Figura 6** – Electroforograma representando a variação de sequência T575T (amostra C56, onde foi usado o *primer* CAD11 Reverse para sequenciação).

No **Quadro 9**, expõem-se os *odds ratio* (com um intervalo de confiança de 95%) obtidos para as variações de sequência estudadas. Pode-se constatar que para qualquer das alterações estudadas se obteve um valor inferior a 3.

**Quadro 9** – *Odds ratio* das variações de sequência no exão 11.

Variação de sequência	<i>Odds ratio</i>
C568Y	0,00
T575T	0,54
G594G	0,00

#### 4.1.1. Aplicação de ferramentas de bioinformática para avaliação do grau de associação desta variação com doença

Para a variação C568Y, que é uma variação não sinónima, foi possível aplicar a ferramenta bioinformática (MSRV) descrita por Jiang et. al (2007) [88] com o objectivo de avaliar se esta variação poderia estar associada a doença. Essa probabilidade é definida por um resultado que varia entre 0 e 1, correspondendo este último a uma probabilidade máxima de estar associada a doença. Para a utilização desta base é necessário saber qual a designação correspondente da proteína NOTCH3 na Swiss-prot database [<http://expasy.org/sprot/sp-docu.html>] – NOTC3\_HUMAN.

Aplicando esta ferramenta (que está disponível na internet, em <http://msms.usc.edu/msrv/about.html>) obteve-se um resultado de 0,9976 [Figura 7]. Este resultado significa que esta variação de sequência apresenta uma elevada probabilidade de estar associada a doença.

PRIORITIZATION RESULTS					
Rank	Score	Protein	Position	Original	Substitution
1	0.9976	NOTC3_HUMAN	568	C	Y

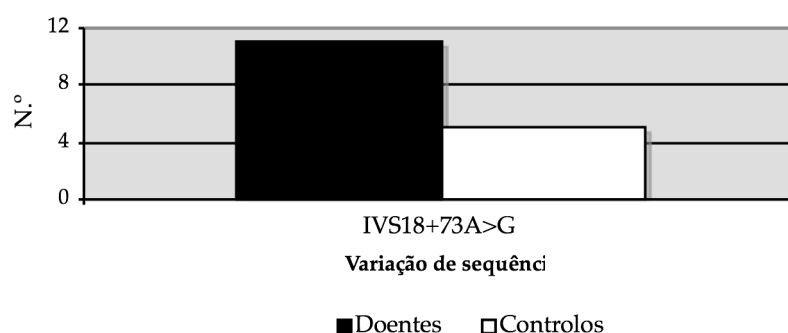
**Figura 7** – Resultado da análise da variação de sequência C568Y utilizando a base MSRV.

#### 4.2. Resultados da análise da variação de sequência localizada no intrão 18

No produto de PCR resultante da amplificação dos exões 18/19, foi possível analisar a variação IVS18+73A>G, que se encontra no intrão 18. Para tal foram usadas amostras de DNA de 102 controlos [66 (65%) do sexo feminino e 36 (35%) do sexo masculino].

Em 200 casos referenciados por suspeita de diagnóstica de CADASIL, foram identificados 11 casos esta alteração intrónica [Figura 8].

Nos 102 controlos, consequentemente, nos 204 alelos normais analisados foram identificados 5 indivíduos (dois do sexo feminino e três do sexo masculino) com esta variação (C15, C66, C80, C106 e C116) [Anexo 4].



**Figura 8** – Número de controlos e doentes onde foi identificada a variação intrónica IVS18+73A>G.

No **Quadro 10**, apresentam-se os números de casos onde foi identificada esta alteração intrónica em ambos os grupos.

**Quadro 10** – Frequências da variação IVS18+73A>G nos controlos e doentes.

Varição de sequência no intrão 18	Frequências alélicas dos controlos	Frequências alélicas dos doentes
IVS18+73A>G	0,02	0,06

Esta alteração intrónica encontra-se em maior frequência nos doentes.

No **Quadro 11**, descreve-se o *odds ratio* (com um intervalo de confiança de 95%) obtido para a variação IVS18+73A>G, também sendo um valor inferior a 3.

**Quadro 11** – Avaliação do *odds ratio* da variação IVS18+73A>G.

Varição de sequência	Odds ratio
IVS18+73A>G	0,45

### 4.3. Resultados da análise das variações de sequência localizadas no exão 19

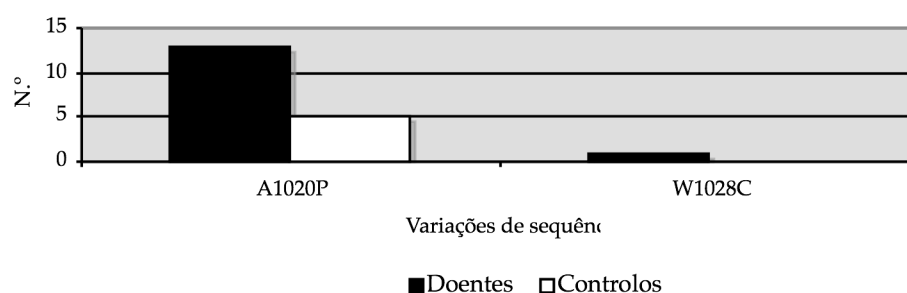
Neste exão foram analisados o polimorfismo A1020P e a variação de sequência W1028C. Ambas são variações do tipo *missense*, mas apenas a última envolve a substituição de um resíduo de triptofano por um de cisteína na proteína NOTCH3.

Para tal foram usadas amostras de DNA de 102 [66 (65%) do sexo feminino e 36 (35%) do sexo masculino].

Em 200 casos referenciados por suspeita de diagnóstica de CADASIL, foram identificados 13 casos com o polimorfismo A1020P e um com a variação de sequência W1028C.

Nos 102 controlos estudados, ou seja, em 204 alelos normais analisados o polimorfismo A1020P foi identificado em 5 indivíduos (três do sexo masculino e dois sexo feminino) (C15, C66, C80, C106 e C116), enquanto a alteração W1028C não foi encontrada em nenhum dos controlos [**Anexo 5**].

A variação de sequência A1020P foi identificada em maior número, quer no grupo dos indivíduos doentes, quer nos controlos [**Figura 9**].



**Figura 9** – Número de controlos e doentes onde foram identificadas as variações de sequência no exão 19.

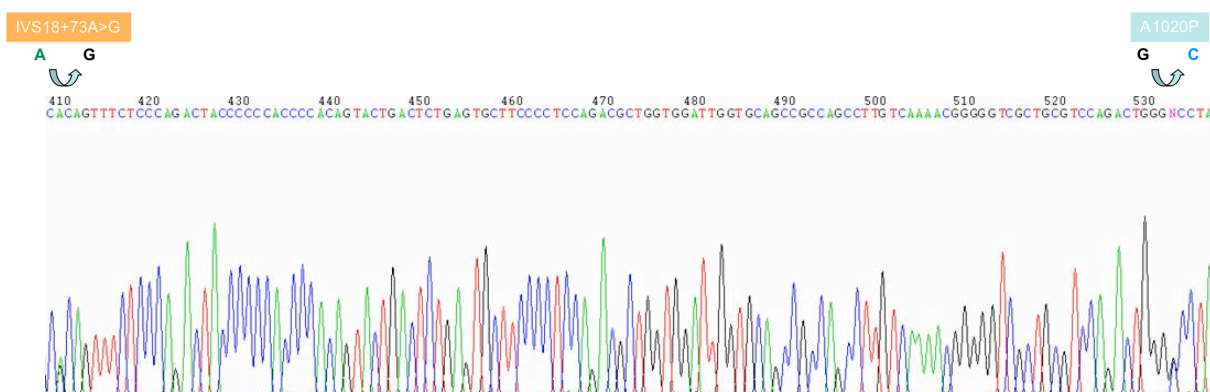


No **Quadro 12** mostram-se as frequências das variações estudadas neste exão.

**Quadro 12** – Frequências encontradas nos controlos e doentes para as alterações estudadas no exão 19.

Variações de sequência no exão 19	Frequências alélicas dos controlos	Frequências alélicas dos doentes
A1020P	0,02	0,07
W1028C	0,00	0,01

Em ambos os grupos avaliados, a variação de sequência com maior frequência foi a A1020P (0,02 nos Controlos e nos 0,07 Doentes) [Figura 10].



**Figura 10** – Electroforeograma representando as variações de sequência IVS18+73A>G e A1020P (amostra C80, onde foi usado o *primer* CAD18/19F para sequenciação).

**NOTA:** As variações de sequência possíveis de analisar no produto de PCR deste exão, foram sempre encontradas em *cis* nos cinco indivíduos controlos onde foram verificadas.

No **Quadro 13**, expõem-se os *odds ratio* (com um intervalo de confiança de 95%) obtidos para as variações de sequência estudadas. Igualmente aqui foi obtido um valor inferior a 3 para qualquer das duas variações.

**Quadro 13** – Determinação do *odds ratio* para as alterações A1020P e W1028C.

Variação de sequência	<i>Odds ratio</i>
A1020P	0,38
W1028C	0,00

#### 4.3.1. Aplicação de ferramentas de bioinformática para avaliação do grau de associação desta variação com doença

Utilizando a base MSRV já referida anteriormente obtiveram-se os seguintes resultados para a variação A1020P – 0,5537 e W1028C – 0,8759 [Figuras 11 e 12]. Ou seja, a variação de sequência W1028C apresenta uma maior probabilidade de estar associada a doença do que a A1020P que apresenta um valor intermédio nesta escala de avaliação.

PRIORITIZATION RESULTS					
Rank	Score	Protein	Position	Original	Substitution
1	0.5537	NOTC3_HUMAN	1020	A	P

**Figura 11** – Resultado da análise da variação de sequência A1020P utilizando a base MSRV.

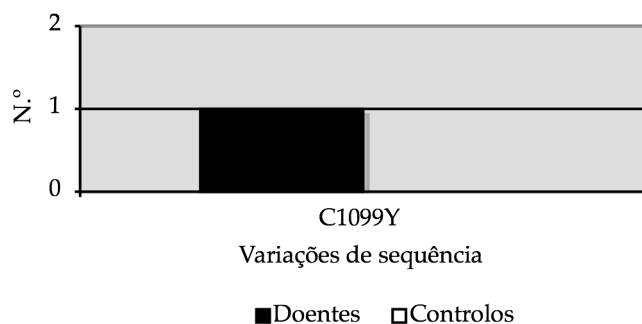
PRIORITIZATION RESULTS					
Rank	Score	Protein	Position	Original	Substitution
1	0.8759	NOTC3_HUMAN	1028	W	C

**Figura 12** – Resultado da análise da variação de sequência W1028C utilizando a base MSRV.

#### 4.4. Resultados da análise da variação de sequência localizada no exão 20

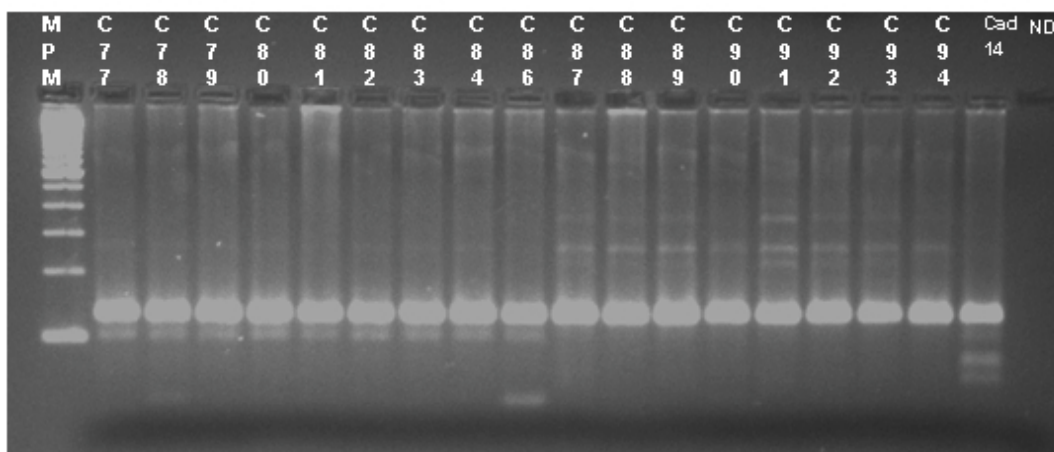
Para análise da variação de sequência C1099Y, foram estudadas amostras de DNA de 105 controlos [70 (67%) do sexo feminino e 35 (33%) do sexo masculino]. Esta alteração é do tipo *missense* e induz a substituição de um resíduo de cisteína por um de tirosina na proteína.

Em 200 casos referenciados por suspeita de diagnóstico de CADASIL, esta variação de sequência foi encontrada em apenas um caso [Figura 13].



**Figura 13** – Número de controlos e doentes onde foram identificadas as variações de sequência no exão 20.

Os 105 controlos, ou seja os 210 alelos normais, foram estudados por digestão com a enzima de restrição HpyCH4III [Figura 14] [Anexo 6].



**Figura 14** – Electroforese em gel de agarose Nusieve 3:1 a 3% durante 60' a 110V após digestão com a enzima HpyCH4II dos produtos de PCR do exão 20 de um grupo de amostras (MPM – marcador peso molecular; Cad14 – Controlo heterozigotia para C1099Y; ND – fragmento não digerido).

Em todos os controlos estudados não foi demonstrada a existência dessa alteração por este método. Contudo nos indivíduos C3, C12, C27, C66 e C74 houve a necessidade de confirmar os resultados por outra técnica de biologia molecular. Procedeu-se à sua sequenciação, não tendo sido, igualmente, identificada a variação de sequência C1099Y.

No **Quadro 14** apresentam-se as frequências desta alteração em ambos os grupos.

**Quadro 14** – Determinação das frequências da variação de sequência C1099Y nos controlos e nos doentes.

Variação de sequência no exão 20	Frequência alélica dos controlos	Frequência alélica dos doentes
C1099Y	0,00	0,01

O grupo dos doentes apresenta uma frequência de 0,01 para esta alteração.

No **Quadro 15**, descrevem os *odds ratio* obtidos (com um intervalo de confiança de 95%) para a variação C1099Y, identicamente ao constatado nas outras variações o valor obtido é inferior a 3.

**Quadro 15** – *Odds ratio* da variação de sequência C1099Y.

Variação de sequência no exão 20	<i>Odds ratio</i>
C1099Y	0,00

#### 4.4.1. Aplicação de ferramentas de bioinformática para avaliação do grau de associação desta variação com doença

Aplicando a base MSRV, a variação de sequência C1099Y tem um valor de 0,9976, isto é, uma probabilidade elevada de estar associada com doença de CADASIL [**Figura 15**].

PRIORITIZATION RESULTS					
Rank	Score	Protein	Position	Original	Substitution
1	0.9976	NOTCH3_HUMAN	1099	C	Y

**Figura 15** – Resultado da análise da variação de sequência C1099Y utilizando a base MSRV.

## **5. Discussão e Conclusões**

---

### 5.1. Polimorfismos *versus* Mutações

Em 1998, Cotton and Scriver [90] definiram polimorfismo como “a neutral or “harmless” mutation”. De acordo com a definição apresentada na página da internet da Human Genome Research Institute’s [91], um polimorfismo é simplesmente “a common variation in the sequence of DNA among individuals”, o que é geralmente quantificado como uma alteração com uma frequência superior a 1% na população [92].

Os mesmos autores que, em 1998, propuseram a definição de polimorfismo também não conseguiram definir com clareza se uma mutação teria um efeito neutro ou modificador sobre o fenótipo. A própria página da Human Genome Research Institute’s [91] apresenta uma definição que é em si própria contraditória – “a permanent structural alteration in DNA. In most cases, DNA changes either have no effect or cause harm, but occasionally a mutation can improve an organism’s chance of surviving and passing the beneficial change on to its descendants”.

Mesmo dentro da classificação dos tipos de mutações houve conceitos que foram sofrendo evoluções. As mutações silenciosas eram previamente classificadas como polimorfismos benignos, contudo actualmente sabe-se que podem induzir alterações nos locais de *splicing* do processamento do RNA e consequentemente terem um efeito patológico [93]. Também à luz dos conhecimentos actuais, nem sempre as mutações *missense* produzirão efeitos nocivos no funcionamento das suas proteínas [94].

A vulgarização e a maior disponibilidade de estudos moleculares para estudo de doenças veio dar a conhecer um número crescente de variações de sequência de significado desconhecido. No processo de clarificação destas variações há um longo processo que deve ser percorrido. Neste contexto, quando são identificadas variações de sequência cujo efeito se desconhece há que proceder à realização de estudos adicionais para a clarificação da sua natureza patogénica – mutação *versus* polimorfismo (polimorfismo inclui uma valorização de frequência; uma variação de sequência pode não ser patogénica e não ser um polimorfismo *strictu sensu*). Neste processo há que ter em conta a avaliação de vários parâmetros, onde se incluem o estudo

molecular dos progenitores [analisar a segregação familiar], verificar se a região onde ocorrem é conservada entre espécies [avaliar o grau de conservação], determinar a frequência destas alterações num grupo controlo [determinar a frequência alélica], analisar a alteração da proteína correspondente e a realização de estudos funcionais.

## 5.2. Mutações e polimorfismos descritos na doença de CADASIL

Como foi referido na **Secção 1.10.2.** quanto à natureza das mutações no gene NOTCH3, estas alteram, na maior parte dos casos, o equilíbrio de resíduos de cisteína no domínio extra-celular do receptor proteico [59]. De acordo com Louvi et al., todas as mutações que não conduzem a esta alteração (mutações “atípicas”) devem ser consideradas com algumas reservas. Consequentemente necessitarão de uma avaliação mais rigorosa previamente a serem consideradas como patogénicas [59].

Pela análise do **Quadro 16**, pode-se verificar que as mutações do tipo missense/nonsense correspondem à quase totalidade das mutações publicadas até ao momento na Human Mutation Database (HGMD)[95]. A mesma conclusão pode ser obtida pela observação da **Figura 20** [75].

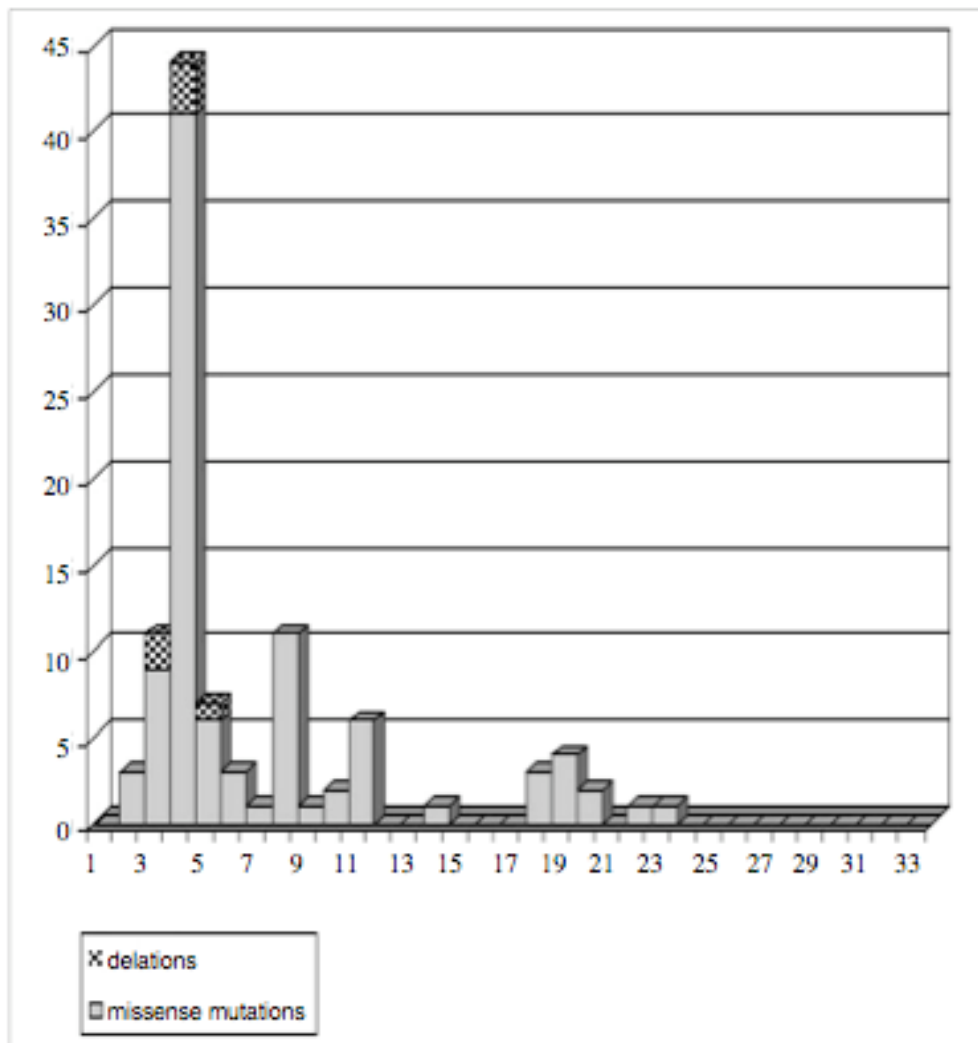
**Quadro 16** – Número de mutações descritas de acordo com HGMD.

<i>NOTCH3</i>	
Tipo de Mutação	Número de Mutações
Missense/Nonsense	75
Splicing	1
Regulatory	0
Small deletions	4
Small insertions	0
Small indels	0
Gross deletions	1
Gross insertions	0
Complex rearrangments	0
Repeat variations	0
<b>Total</b>	<b>81</b>

Os exões onde foram identificadas a maioria das mutações são, por ordem decrescente de frequência, os exões 4, 3, 5, 6 e 11. Na literatura, encontram-se descritas variações populacionais, por exemplo a segunda

localização mais frequente, na França, Reino Unido e Alemanha, é o exão 3 [75, 77] [Figura 16]. Em Portugal, parece haver evidências de que os exões 4 e 11 corresponde àqueles onde são detectadas um maior número de mutações [78].

O estudo das diferenças populacionais é obviamente importante, não só para tentar compreender diferenças fenotípicas que possam existir como também para estabelecer estratégias mais rentáveis na orientação do estudo molecular do gene *NOTCH3*.



**Figura 16** – Distribuição das mutações já descritas no gene *NOTCH3* [75].

Muitos polimorfismos foram também já descritos na região codificante do gene [75] [Quadro 17].



**Quadro 17** - Alguns dos polimorfismos já descritos, onde se encontra mencionado o que foi analisado neste estudo – A1020P.

No	Exon	Notch3 nucleotide change	Amino acid change
1	3	caG360caA	No
2	3	acC381acT	No
3	4	ccT423ccC	No
4		gcC474gcA	No
5		cgA546cgG	No
6		cAt587cGt	H170R
7		ggT591ggC	No
8		ggC594ggA	No
9		gcG684gcA	No
		taC738taT	No
10	7	ccT1218ccC	No
11	9	cCc1565cTc	P496L
12	11	acG1698acT	No
13	16	aCc2585aTc	T836I
14		tgT2616tgC	No
15	17	ccG2820ccA	No
16	19	Gcc3136Ccc	A1020P
17	21	caC3477caA	H1133Q
18	22	Gtg3625Atg	V1183M
19	23	cGc3803cAc	R1242H
20	25	ccG4641ccA	No
21	33	ccG6300ccA	No
22		gcG6516gcA	No
23		gTg6746gCg	V2223A

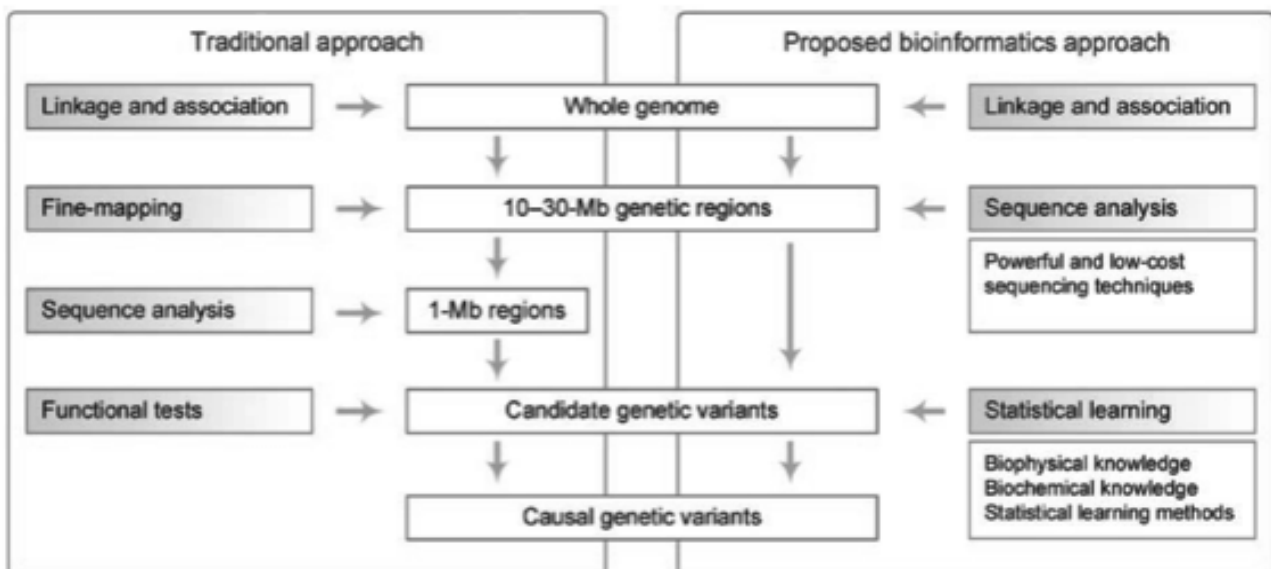
Adaptado de Federico *et al.*, 2005

### 5.3. MSR<sub>V</sub> system

Este modelo foi desenvolvido por Jiang *et al.* [88] perante a necessidade de clarificar as inúmeras variações genéticas que se têm vindo a descobrir com a maior disponibilidade de acesso à realização de estudos moleculares. Havendo esta identificação de possíveis mutações causadores de doença e de variantes genéticas torna-se imperioso verificar a sua causalidade e relação com patologia para que estas informações possam ser aplicados na prevenção, diagnóstico e, futuramente, no tratamento. As mutações patogénicas geralmente induzem alterações significativas na estrutura e função da

proteína, em contraste, com os *single nucleotide polymorphisms* (SNP) não sinónimos neutrais que levam a alterações diminutas ou mesmo negligenciáveis, quase não produzindo qualquer efeito sobre a proteína.

No modelo tradicional de mapeamento de doenças [Figura 17], após a definição de regiões de 10-30Mb estatisticamente significativas por estudos de haplotipagem ou de associação, estas são reduzidas até regiões inferiores a 1Mb. Finalmente, é feita a sequenciação e estudos funcionais para inferir quais as alterações genéticas associadas a patologia. No modelo informático que estes autores propõem a fase de mapeamento fino é transposta para a fase de sequenciação directa da região de 10-30Mb, aplicando-se a abordagem bioinformática que integra não só dados bioquímicos e físicos como também metodologias estatísticas.



**Figura 17** – Esquemas representativo dos modelos tradicional e bioinformático [88].

Uma das limitações que apresenta é o facto de apenas ser possível aplicar quando a alteração ocorre num domínio proteico conhecido, tendo-se que recorrer a outros métodos de alinhamento de sequência múltiplas. Também só pode ser aplicado quando as alterações são não sinónimas e localizadas em regiões génicas previamente já conhecidas como relacionadas com a patologia. Contudo, é claramente reconhecido que há outras regiões genómicas que quando alteradas podem também induzir alterações

patogénicas, tais como locais de ligação de factores de transcrição e regiões promotoras.

#### 5.4. Discussão dos resultados obtidos

Os resultados apresentados irão ser discutidos e analisados tendo por base as considerações acima apresentadas.

##### 5.4.1. Exão 11 – Variações de sequência C568Y, T575T e G594G

###### 5.4.1.1. Tipo de variações de sequência

Das três variações de sequência estudadas, apenas a primeira – C568Y – é do tipo *missense*, enquanto que as outras duas são silenciosas. Como já foi comentado previamente esta associação com a patogenicidade já não é tão linear quanto antigamente se supunha. Ou seja, nem todas as variações que conduzem a substituição de aminoácidos nas proteínas têm um efeito nocivo na sua actividade. E, por outro lado, as supostas variações de sequência silenciosas podem alterar o processamento do RNA e, conseqüentemente, a proteína. Mas à partida poderemos considerar que a variação C568Y terá maior potencial patogénico que as outras duas.

A variação de C568Y traduz a substituição, no codão 568, de um resíduo de cisteína por um de resíduo de tirosina. Este facto é mais um dado a favor de se tratar de uma variação patogénica, uma vez que caracteristicamente as mutações da CADASIL envolvem resíduos de cisteína.

###### 5.4.1.2. Diferenças entre as frequências alélicas determinadas nos controlos e nos doentes

A determinação do *odds ratio* também permitiu concluir que não existe uma diferença estatisticamente significativa entre os dois grupos.

A variação com maior frequência em ambos os grupos foi a T575T (0,02 nos controlos e 0,04 nos doentes). As outras duas variações, C568Y e G594G, apresentaram frequências de 0,00 e 0,01 nos controlos e doentes respectivamente. Tendo em conta a definição de polimorfismo, actualmente aceite, podemos afirmar que haverá uma maior probabilidade de que a

variação T575T se enquadre neste conceito pela sua frequência ser superior a 0,01 ou a 1%. Da mesma forma, pela ausência de identificação no grupo controlo, as variações de sequência, C568Y e G594G, estarão associadas a um maior risco de patogenicidade, ou seja, de se tratarem de verdadeiras mutações.

#### **5.4.1.3. Avaliação do grau de associação da variação de sequência C568Y utilizando a base MSR**

Utilizando esta bioferramenta estimou-se um risco elevado (0,9976) de estar associado com CADASIL, de acordo com esta base de dados. Portanto, é mais um factor a ter em conta na classificação desta variação como patogénica – mutação.

Esta base não permite a sua aplicação para variações de sequência sinónimas, como é o caso das T575T e G594G.

#### **5.4.2. Intrão 18 – Variação de sequência IVS18+73A>G**

##### **5.4.2.1. Tipo de variação de sequência**

Esta variação de sequência é intrónica, logo é necessário avaliar se interfere com o processamento do RNA. Constatou-se, pelo menos nos controlos, que esta variação foi sempre identificada em *cis* com o polimorfismo A1020P no exão 19.

##### **5.4.2.2. Diferenças entre as frequências alélicas determinadas nos controlos e nos doentes**

A determinação do *odds ratio* não permitiu concluir que existe uma diferença estatisticamente significativa entre os dois grupos. Contudo, apresentou uma maior frequência nos doentes (0,06 para 0,02 nos controlos).

Observa-se assim uma frequência que permite afirmar uma maior tendência para a não patogenicidade desta variação de sequência – polimorfismo.

### 5.4.3. Exão 19 – Polimorfismo A1020P e a variação de sequência W1028C

#### 5.4.3.1. Tipo de variações de sequência

A variação A1020P, apesar de ser do tipo *missense* (resultando na substituição no codão 1020 de uma alanina por prolina), é um polimorfismo já conhecido, como foi anteriormente descrito. Consequentemente, não tem qualquer efeito patogénico e, portanto, não está associada a doença de CADASIL [75].

A alteração W1028C é do tipo *missense*, e as repercussões na proteína resultam na substituição de um resíduo de triptofano no codão 1028 por um de cisteína, estando em concordância com a maior parte das mutações descritas para esta doença.

#### 5.4.3.2. Diferenças entre as frequências determinadas nos controlos e nos doentes

O *odds ratio* também permitiu concluir que não existe uma diferença estatisticamente significativa entre os dois grupos.

O polimorfismo A1020P apresenta uma frequência na nossa população de controlos de 0,02, ou seja, uma percentagem de 2%. Nos doentes, tem uma frequência de 0,07.

A variação de sequência *missense* W1028C tem uma maior frequência no grupo dos doentes. Nos controlos, não foi identificado qualquer indivíduo com esta alteração, assim poder-se-á afirmar que apresenta um maior risco de patogenicidade, ou seja, de se tratar de uma verdadeira mutação.

#### 5.4.3.3. Avaliação do grau de associação das variações de sequência A1020P e W1028C utilizando a base MSR/V

Apesar de o polimorfismo A1020P ser já reconhecido como tal, foi submetido à análise por este sistema. O valor obtido de 0,5537 não permite estabelecer com clareza umnexo de causalidade com a CADASIL. Por outro lado, a variação W1028C tem um elevado valor 0,8759.

#### 5.4.4. Exão 20 – Variação de sequência C1099Y

##### 5.4.4.1. Tipo de variação de sequência

A alteração C1099Y é do tipo *missense*, conduzindo na proteína a uma substituição, no codão 1099, do aminoácido cisteína por tirosina. Alterando, portanto, o equilíbrio de cisteínas como acontece na maior parte das mutações apresentadas na bibliografia.

##### 5.4.4.2. Diferenças entre as frequências determinadas nos controlos e nos doentes

Mesmo não havendo uma diferença estatisticamente significativa entre os dois grupos, podemos afirmar esta variação apresenta uma maior frequência nos doentes (0,01 e nos controlos 0,00). Assim sendo, parece haver uma tendência para que esta variação de sequência se encontre associada a doença, isto é, tenha um efeito patogénico.

##### 5.4.4.3. Avaliação do grau de associação com doença da variação de sequência C1099Y utilizando a base MSR/V

Pela utilização da base MSR/V, o valor que se obteve para esta alteração foi de 0,9976, traduzindo uma elevada probabilidade de estar associada a doença.

## 5.5. Conclusões

---

O objectivo principal deste estudo foi conseguido, ou seja, foi possível determinar as frequências das variações de sequência no grupo controlo e inferir com algum grau de certeza o potencial patogénico de algumas das variações de sequência estudadas.

O **Quadro 18** sumariza os resultados obtidos.

**Quadro 18** – Resultados obtidos para as variações de sequência estudadas.

Variações de sequência	Tipo	Alteração resíduo de cisteína	Frequência grupo controlo	MSRV
C586Y	Missense	+	0,00	0,9976
T575T	Silenciosa	-	0,02	NA
G594G	Silenciosa	-	0,00	NA
IVS18+73A>G	Intrónica	-	0,02	NA
A1020P	Missense	-	0,02	0,5537
W1028C	Missense	+	0,00	0,08759
C1099Y	Missense	+	0,00	0,09776

+ com alteração; - sem alteração; NA- não aplicável

As variações de sequência com maior frequência no grupo controlo são a T575T, no exão 11, e a variação intrónica IVS18+73A>G. O polimorfismo A1020P apresenta um frequência de 0,02 na população estudada e foi sempre encontrado em *cis* com a variação intrónica IVS18+73A>G.

No que diz respeito à avaliação das variações de sequência não sinónimas (excluindo o polimorfismo já conhecido), em que foi possível aplicar a ferramenta bioinformática MSRV, as que apresentam uma maior probabilidade de estarem associadas a CADASIL são por ordem crescente W1028C e as C568Y e C1099Y (com o mesmo grau de probabilidade).

Quanto à probabilidade de estas variações de sequência serem polimorfismos ou mutações, de acordo com os resultados deste estudo, o **Quadro 19** apresenta essa avaliação.

**Quadro 19** – Avaliação quanto à probabilidade destas variações serem polimorfismos ou mutações.

Variações de sequência	Probabilidade Polimorfismo <i>versus</i> Mutação
C568Y	Mutação>Polimorfismo
T575T	Polimorfismo>Mutação
G594G	Polimorfismo>Mutação
IVS18+73A>G	Polimorfismo>Mutação
A1020P	Polimorfismo
W1028C	Mutação>Polimorfismo
C1099Y	Mutação>Polimorfismo

De acordo com o que foi exposto anteriormente, poder-se-á afirmar com algum grau de certeza que as variações de sequência C568Y e C1099Y, em igualdade de circunstâncias, e a W1028C são mutações.

As variações de sequência T575T, IVS18+73A>G e G594G necessitarão da realização de estudos adicionais (avaliação de segregação familiar, da conservação e interferência com o processamento de RNA) para se afirmar com maior convicção que são polimorfismos.

Apesar da sua aparente simplicidade é o ponto de partida para a realização de outros estudos futuros, para se obterem conclusões mais definitivas. Neste contexto, seria importante a execução de estudos que permitissem analisar com certeza a alteração que induzem na proteína NOTCH3 e a realização de estudos funcionais.

A caracterização molecular de doentes com CADASIL é de extrema importância. Trata-se de uma doença incapacitante. Nesse sentido, o estudo molecular e sobretudo a clarificação com certeza de certas variações de sequência ainda identificadas *de novo* poderão a vir a constituir uma valiosa ajuda à prática clínica com vista a proporcionar terapêuticas mais eficazes, melhorar estratégias para prevenção de complicações, oferecer a realização do estudo dos indivíduos assintomáticos e estabelecer protocolos de vigilância naqueles em que se confirme molecularmente o diagnóstico.



## **6. Anexos**

---

**Anexo 1** - GenBank – *NOTCH3* (NM\_004663) – Páginas 57 a 63.

[<http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucore&id=134244284>]

Search  for

[Limits](#)

[Preview/Index](#)

[History](#)

[Clipboard](#)

[Details](#)

Display  Show   Hide:  sequence  all but gene, CDS and mRNA features

Range: from  to   Reverse complemented strand Features:  SNP  STS  Exon

1: [NM\\_000435](#). Reports Homo sapiens Notc...[gi:134244284]

[Links](#)

[Comment](#) [Features](#) [Sequence](#)

LOCUS NM\_000435 8089 bp mRNA linear PRI 18-NOV-2007  
 DEFINITION Homo sapiens Notch homolog 3 (Drosophila) (NOTCH3), mRNA.  
 ACCESSION NM\_000435  
 VERSION NM\_000435.2 GI:134244284  
 KEYWORDS .  
 SOURCE Homo sapiens (human)  
 ORGANISM [Homo sapiens](#)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
 Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 8089)  
 AUTHORS Konishi,J., Kawaguchi,K.S., Vo,H., Haruki,N., Gonzalez,A.,  
 Carbone,D.P. and Dang,T.P.  
 TITLE Gamma-secretase inhibitor prevents Notch3 activation and reduces  
 proliferation in human lung cancers  
 JOURNAL Cancer Res. 67 (17), 8051-8057 (2007)  
 PUBMED [17804716](#)  
 REMARK GeneRIF: Notch pathway plays an important role in lung cancer  
 biology.  
 REFERENCE 2 (bases 1 to 8089)  
 AUTHORS Gramantieri,L., Giovannini,C., Lanzi,A., Chieco,P., Ravaioli,M.,  
 Venturi,A., Grazi,G.L. and Bolondi,L.  
 TITLE Aberrant Notch3 and Notch4 expression in human hepatocellular  
 carcinoma  
 JOURNAL Liver Int. 27 (7), 997-1007 (2007)  
 PUBMED [17696940](#)  
 REMARK GeneRIF: Suggest an aberrant expression of Notch3 and Notch4 in HCC  
 and allow the hypothesis of an activation of Notch signalling by  
 Notch3.  
 REFERENCE 3 (bases 1 to 8089)  
 AUTHORS Coleman,M.L., McDonough,M.A., Hewitson,K.S., Coles,C.,  
 Mecinovic,J., Edelmann,M., Cook,K.M., Cockman,M.E., Lancaster,D.E.,  
 Kessler,B.M., Oldham,N.J., Ratcliffe,P.J. and Schofield,C.J.  
 TITLE Asparaginyl hydroxylation of the Notch ankyrin repeat domain by  
 factor inhibiting hypoxia-inducible factor  
 JOURNAL J. Biol. Chem. 282 (33), 24027-24038 (2007)  
 PUBMED [17573339](#)  
 REMARK GeneRIF: NOTCHES N1, N2, and N3 all bound to FIH; results suggest  
 the possibility that Notch ICDs are FIH substrates.  
 REFERENCE 4 (bases 1 to 8089)  
 AUTHORS Kusaba,T., Hatta,T., Kimura,T., Sonomura,K., Tanda,S.,  
 Kishimoto,N., Kameyama,H., Okigaki,M., Mori,Y., Ishigami,N.,  
 Mizuno,T., Nakagawa,M. and Matsubara,H.  
 TITLE Renal involvement in cerebral autosomal dominant arteriopathy with  
 subcortical infarcts and leukoencephalopathy (CADASIL)  
 JOURNAL Clin. Nephrol. 67 (3), 182-187 (2007)  
 PUBMED [17390743](#)  
 REMARK GeneRIF: Genetic analysis of Notch3 revealed an R141C missense  
 mutation and she was diagnosed with CADASIL complicated with IgA  
 nephropathy.  
 REFERENCE 5 (bases 1 to 8089)  
 AUTHORS Zang,S., Ji,Ch., Qu,X., Dong,X., Ma,D., Ye,J., Ma,R., Dai,J. and  
 Guo,D.  
 TITLE A study on Notch signaling in human breast cancer  
 JOURNAL Neoplasma 54 (4), 304-310 (2007)  
 PUBMED [17822320](#)  
 REMARK GeneRIF: Overexpression of notch3 is associated with breast cancer

REFERENCE 6 (bases 1 to 8089)  
AUTHORS Lewis,J.  
TITLE Neurogenic genes and vertebrate neurogenesis  
JOURNAL Curr. Opin. Neurobiol. 6 (1), 3-10 (1996)  
PUBMED [8794055](#)  
REMARK Review article

REFERENCE 7 (bases 1 to 8089)  
AUTHORS Larsson,C., Lardelli,M., White,I. and Lendahl,U.  
TITLE The human NOTCH1, 2, and 3 genes are located at chromosome positions 9q34, 1p13-p11, and 19p13.2-p13.1 in regions of neoplasia-associated translocation  
JOURNAL Genomics 24 (2), 253-258 (1994)  
PUBMED [7698746](#)

REFERENCE 8 (bases 1 to 8089)  
AUTHORS Sugaya,K., Fukagawa,T., Matsumoto,K., Mita,K., Takahashi,E., Ando,A., Inoko,H. and Ikemura,T.  
TITLE Three genes in the human MHC class III region near the junction with the class II: gene for receptor of advanced glycosylation end products, PBX2 homeobox gene and a notch homolog, human counterpart of mouse mammary tumor gene int-3  
JOURNAL Genomics 23 (2), 408-419 (1994)  
PUBMED [7835890](#)

REFERENCE 9 (bases 1 to 8089)  
AUTHORS Tournier-Lasserre,E., Joutel,A., Melki,J., Weissenbach,J., Lathrop,G.M., Chabriat,H., Mas,J.L., Cabanis,E.A., Baudrimont,M., Maciazek,J. et al.  
TITLE Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12  
JOURNAL Nat. Genet. 3 (3), 256-259 (1993)  
PUBMED [8485581](#)

REFERENCE 10 (bases 1 to 8089)  
AUTHORS Beleil,O.M., Mickey,M.R. and Terasaki,P.I.  
TITLE Comparison of male and female kidney transplant survival rates  
JOURNAL Transplantation 13 (5), 493-500 (1972)  
PUBMED [4557798](#)

COMMENT REVIEWED [REFSEQ](#): This record has been curated by NCBI staff. The reference sequence was derived from [U97669.1](#) and [AC004663.1](#). On Mar 28, 2007 this sequence version replaced gi:[4557798](#).

Summary: This gene encodes the third discovered human homologue of the *Drosophila melanogaster* type I membrane protein notch. In *Drosophila*, notch interaction with its cell-bound ligands (delta, serrate) establishes an intercellular signalling pathway that plays a key role in neural development. Homologues of the notch-ligands have also been identified in human, but precise interactions between these ligands and the human notch homologues remains to be determined. Mutations in NOTCH3 have been identified as the underlying cause of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL).

Sequence Note: This RefSeq record was created from transcript and genomic sequence data because transcript sequence consistent with the reference genome assembly was not available for all regions of the RefSeq transcript. The extent of this transcript is supported by transcript alignments.

Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Entrez Gene record to access additional publications.  
COMPLETENESS: full length.

PRIMARY	REFSEQ_SPAN	PRIMARY_IDENTIFIER	PRIMARY_SPAN	COMP
	1-681	U97669.1	3-683	
	682-682	AC004663.1	35190-35190	c
	683-999	U97669.1	685-1001	
	1000-1000	AC004663.1	34693-34693	c
	1001-2613	U97669.1	1003-2615	
	2614-2614	AC004663.1	27480-27480	c
	2615-2817	U97669.1	2617-2819	
	2818-2818	AC004663.1	24783-24783	c
	2819-4638	U97669.1	2821-4640	
	4639-4639	AC004663.1	17398-17398	c
	4640-6513	U97669.1	4642-6515	
	6514-7932	AC004663.1	2929-4347	c
	7933-8089	U97669.1	7935-8091	

FEATURES  
source

Location/Qualifiers

1..8089  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19p13.2-p13.1"

gene

1..8089  
/gene="NOTCH3"  
/note="Notch homolog 3 (Drosophila); synonyms: CASIL, CADASIL"  
/db\_xref="GeneID:4854"  
/db\_xref="HGNC:7883"  
/db\_xref="HPRD:02607"  
/db\_xref="MIM:600276"

exon

1..194  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=1

CDS

77..7042  
/gene="NOTCH3"  
/GO\_component="[integral to membrane](#); [integral to plasma membrane](#) [PMID 7835890]; [membrane](#)"  
/GO\_function="[calcium ion binding](#); [receptor activity](#)"  
/GO\_process="[cell differentiation](#); [morphogenesis](#) [PMID 7698746]; [Notch signaling pathway](#); [regulation of development](#); [regulation of transcription, DNA-dependent](#); [transcription](#)"  
/codon\_start=1  
/product="Notch homolog 3"  
/protein\_id="NP\_000426.2"  
/db\_xref="GI:134244285"  
/db\_xref="CCDS:CCDS12326.1"  
/db\_xref="GeneID:4854"  
/db\_xref="HGNC:7883"  
/db\_xref="HPRD:02607"  
/db\_xref="MIM:600276"

/translation="MGPGARGRRRRRPMSPPPPPVVRALPLLLLLLAGPGAAAPPCL  
DGSPCANGGRCTQLPSREAAACLP PGWVGERCQLEDPC HSGPCAGRGVCQSSV VAGTA  
RFSCRCPRGFRGPDCSLPDPCLSSPCA HGARCSVGP DGRFLCSCPPGYQGRSCRSDVD  
ECRVGEP CRHGGTCLNTPG SFRQC PAGYTG PLCENPAVPCAPSPCRNGGTCRQSGDL  
TYDCACLPGFEGQNC EVNVDDCPGHRCLNGGTCVDGVNTYNCQCPPEWTGQFC TEDVD  
ECQLQP NACHNGGTCFN TLGGHSCVCVNGWTGESCSQNI DDCAVCFHGATCHDRVA  
SFYCACPMGKTGLLCHLDDACVSNPCHEDAICDTNPVNGRAICTCPPGFTGGACDQDV  
DECSIGANPCEHLGRCVNTQGSFLCQCGRGYTGPRCETDVNECLSGPCR NQATCLDRI  
GQFTICMAGFTGT YCEVDIDECQSSPCVNGGVCKDRVNGFSCTCPSGFSGSTCQLDV  
DECASTPCRNGAKCVDQPDGYECRAEGFEGTLCDRNVDDCSPDPCHHGRCVDGIASF  
SCACAPGYTGTRCESQVDECRSQPCR HGGKCLDLVDKYLRCRPSGTTGVNCEVNIDDC  
ASNPTFGVCRDGINRYDCVCQPGFTG PLCNVEINECASSPCGEGGSCVDGENGFRCL  
CPPGSLPPLCLPPSHPCAHEPCSHGICYDAPGGFRCVCEPGWSGPRCSQSLARDACES  
QPCRAGGTCSSDGMGFHCTCPPGVQGRQCELLSPCTPNPCEHGGRCESAPGQLPVCSC  
PQGWQGPCRQDQVDE CAGPAPCGPHGICTNLAGSF SCTCHGGYTGPS CDQDINDCDPN  
PCLNGGSCQDGVGSFSCSCLPGFAGPRCARDVDECLSNPCGPGTCTDHVASFTCTCPP  
GYGFHCEQDL PDCSPSSCFNGGTCVDGVNSFSCLCRPGYTG AHQCHEADPCLSRPCL  
HGGVCSAAHPGFRCTCLESFTGPQCQTLVDWCSRQPCQNGGRCVQTGAYCLCPPGWSG  
RLCDIRSLPCREAAAQIGVRLEQLCQAGGQCVDEDS SHYVCPEGRTGSHCEQEVDP  
LAQPCQHGGTCRGMGGYMCCLPGYNGDN CEDDVCASQPCQHGGS CIDLVARYLC  
SCPFGTLGVLC EINEDDCGPGPPLDSGPRCLHNGTCDLVGGFRCTCPPGYTGLRCEA  
DINECRSGACHAAHTRDCLQDPGGFRCLCHAGFSGPRCQTVLSPCESQPCQHGGQCR  
PSPGPGGLTFTCHCAQPFWGP RCERVARSCRELQCPVGVPCQOTPRGPRCACPPGLS  
GPSCRFPSPGASNASCAAAPCLHGGSCR PAPLAPFFRCACAQGWTPRCEAPAAA  
PEVSEEP RCPRAACQAKRGDQRCDRECN SPGCGWDGDCSLSVGDPWRQCEALQCWRL  
FNNSRCDPACSSPACLYDNFDCHAGGRERTCNPVYKEYCADHFADGRCDQGCNTEECG  
WDGLDCASEVPALLARGV LVTVLLPPEELLRSSADFLQRLSAILRTSLRFR LDAHGQ  
AMVFPYHRPSPGSEPRARREL APEVIGSVVMLEIDNRLCQSPENDHCFPDAQSAADY  
LGLSAVERLDFPYPLRDV RGEPELEPPESVPLLLLVAGAVLLLVILV LGMVARRK  
REHSTLWFPEGFSLHKDVASGHKGRREPVGQDALGMKNMAKGESLMGEVATDWM DTEC  
PEAKRLKVEEPMGAEAEAVDCRQWTOHHLVAADIRVAPAMALTPPQGDADADGMDVNV  
RGPDGFTPLMLASFCGGALEPMPTEEDEADDT SASIISDLICQGAQLGARTDR TGETA  
LHLAARYARADA AKRLLDAGADTNAQDHSGRTP LHTAVTADAQGVFQILIRNRSTDL D  
ARMADGSTALIL AARLAVEGMVEELIASHADVNAVDELGKSALHWAAVNNVEATLAL  
LKNGANKMDQDSKEETPLFLAAREGSYEA AKLLLDHFNREITDHLDR LPRDVAQERL  
HQDIVRLLDQPSGPRSPPGPHGLGPLLCPPGAFLPGLKAAQSGSKSRPPGKAGLGP  
QGP RGRGK KLT LACPGPLADSSVTLSPVDSLDSRPF GGPASP GGFLEGP YAAATA

TAVSLAQLGGPGRAGLGRQPPGGCVLSLGLLNPNVAVPLDWARLPPAPPGPSFLLPLA  
PGPQLLNPGTPVSPQERPPPYLAVPGHGEEYPAAGAHSSPPKARFLRVPSEHPYLTPS  
PESPEHWASPPPSLSDWSESTPSPATATGAMATTTGALPAQPLPLSVPSSLAQAQTQ  
LGPQPEVTPKRQVLA"

[exon](#) 195..273  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=2

[exon](#) 274..416  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=3

[exon](#) 417..755  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=4

[exon](#) 756..878  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=5

[exon](#) 879..1112  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=6

[exon](#) 1113..1268  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=7

[exon](#) 1269..1454  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=8

[exon](#) 1455..1568  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=9

[exon](#) 1569..1682  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=10

[exon](#) 1683..1916  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=11

[exon](#) 1917..2027  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=12

[exon](#) 2028..2220  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=13

[exon](#) 2221..2372  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=14

[exon](#) 2373..2486  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=15

[exon](#) 2487..2642  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=16

[exon](#) 2643..2868  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=17

[exon](#) 2869..3070  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=18

[exon](#) 3071..3218  
/gene="NOTCH3"

[exon](#) /inference="alignment:Splign"  
/number=19  
3219..3403  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=20  
[exon](#) 3404..3536  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=21  
[exon](#) 3537..3794  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=22  
[STS](#) 3786..3910  
/gene="NOTCH3"  
/standard\_name="SHGC-57293"  
/db\_xref="UniSTS:[52587](#)"  
[exon](#) 3795..3913  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=23  
[exon](#) 3914..4479  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=24  
[exon](#) 4480..4812  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=25  
[exon](#) 4813..4967  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=26  
[exon](#) 4968..5190  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=27  
[exon](#) 5191..5275  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=28  
[exon](#) 5276..5438  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=29  
[exon](#) 5439..5743  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=30  
[exon](#) 5744..5891  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=31  
[exon](#) 5892..5989  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=32  
[exon](#) 5990..8071  
/gene="NOTCH3"  
/inference="alignment:Splign"  
/number=33  
[STS](#) 7055..7292  
/gene="NOTCH3"  
/standard\_name="PMC122602P1"  
/db\_xref="UniSTS:[270400](#)"  
[STS](#) 7783..7943  
/gene="NOTCH3"  
/standard\_name="G62127"  
/db\_xref="UniSTS:[139187](#)"  
[STS](#) 7828..7942  
/gene="NOTCH3"  
/standard\_name="D19S1009"  
/db\_xref="UniSTS:[20959](#)"

[polyA\\_signal](#)

8049..8054

/gene="NOTCH3"

[polyA\\_site](#)

8071

/gene="NOTCH3"

ORIGIN

```
1 gcggcgcgga ggctgcccc ggacgcgcc ggagcccagg gaaggagga ggaggggag
61 gtgcgggcc gccccatgg gccgggggc cctggccgc cgccgccgc gtcgcccgat
121 gtgcgcgcca ccgccaccgc caccctgctg ggcgctgcc ctgctgctgc tgcctagcggg
181 gccgggggt gcagccccc cttgctgga cgaagcccc tgtgcaaat gaggtcgtg
241 caccagctg ccctcccgg aggtgcctg cctgtgccg cctggctgg tgggtgagc
301 gtgtcagctg gaggacccct gtcactcagg cccctgtgct ggccgtggtg tctgccagag
361 ttcagtgggt gctggcacc cccgattctc atgcgggtgc cccctggct tccgaggccc
421 tgactgctcc ctgccagatc cctgcctcag cagcccttgt gccacggtg cccgctgctc
481 agtggggccc gatggacgt tcctctgctc ctgccacct ggctaccagg gccgcagctg
541 ccgaagcgac gtgatgagt gccgggtgg tgagccctgc cgccatggtg gcacctgcct
601 caacacacct ggctccttc gctgccagt tccagctggc tacacaggg cactatgtga
661 gaaccccgcg gtcctctgt caccctcacc atgcogtaac gggggcacct gcaggcagag
721 ttggcgacct gcttacgact gtcctgctc tctggggtt gaggtcaga attgtgaagt
781 gaacgtggac gactgtccag gacaccgat tctcaatgg gggacatgc tggatggct
841 caacacctat aactgccagt gccctcctga gtggacagg cagtctgca cggaggact
901 ggatgagtgt cagctgcagc ccaacgcct ccacaatgg ggtacctgt tcaacacct
961 ggggtggccc agctgcgtg gtgtcaatg ctggacagg gagagtgca gtcagaatat
1021 cgatgactgt gccacagcc tgtgctcca tggggccacc tgccatgacc gcgtggcttc
1081 tttctactgt gccctgcccc tgggcaagc tggcctcctg tctcactgg atgacgctg
1141 tgtcagcaac ccctgccagc aggtatgat ctgtgacaca aatccggtg acggccgggc
1201 catttgacc gctctccc gcttcagg tggggcatg gaccagatg tggacagtg
1261 ctctatcgcc gccaacccct gcgagcactt gggcaggtg gtgaacacgc agggctcctt
1321 cctgtgccag tgcggtcgtg gctacactgg acctcgtgt gagaccgat tcaacgagt
1381 tctgtcgggg ccctgccga accaggccac gtgctcgac cgcataggcc agttcactg
1441 tatctgtatg gcagcctca caggaaccta ttgcgaggtg gacattgac agtgtcagag
1501 tagcccctgt gtaacgggt gggctgcaa ggaccgagtc aatggctca gctgcacct
1561 cccctcgggc ttcagcggct ccacgtgtca gctggacgtg gacgaatgc ccagcacgcc
1621 ctgcaggaat ggcgcaaat gcgtggaca gcccgatggc tacgagtgc gctgtgccga
1681 gggctttgag ggcagctgt gtgatcgca cgtggacgac tgctcccctg acctatgcca
1741 ccatggtcgc tgcgtgatg gcatcgccag cttctcatg gcctgtgctc ctggctacac
1801 gggcacacgc tgcgagagc aggtggacga atgcgcgac cagccctgcc gccatggcgg
1861 caaatgccta gacctggtg acaagtacct ctgcccgtgc cctctggga ccacaggtgt
1921 gaactgcgaa gtgaacattg acgactgtgc cagcaacccc tgcaccttg gagtctgcc
1981 tgatggcatc aaccgctac actgtgtctg ccaacctggc ttcacaggg ccctttgtaa
2041 cgtggagatc aatgagtgt cttccagccc atgcggcgag ggaggttct gtgtggatgg
2101 ggaaaatggc tcccgtgcc tctgccgccc tggtccttg ccccactct gcctcccccc
2161 gagccatccc gtcgcccatt agccctgcag tcacggcatc tgctatgat cacctggcgg
2221 gttccgctgt gtgtgtgagc ctggctggag tggccccgc tgcagccaga gcctggccc
2281 agacgcctgt gactcccagc cgtgcagggc cgggtggaca tgcagcagc atggaatggg
2341 tttccactgc acctgcccgc ctggtgtcca gggacgtcag tgtgaactc tctccccctg
2401 caccgccaac ccctgtgagc atgggggccc ctgocgagct gccctggcc agctgcctgt
2461 ctgctcctgc cccaggggt gccaaaggccc acgatgccag caggatgtgg acgagtgtg
2521 tggccccgca ccctgtggcc ctcatggtat ctgcaccaac ctggcaggg gtttcagctg
2581 cactgccat gcagggtaca ctggcccttc ctgocgatg gacatcaat actgtgacc
2641 caaccatgc ctgaacggtg gctcgtcca agacggcgtg ggctccttt cctgtcctg
2701 cctcccctgt ttcgccggcc cacgatgctc ccgcatgtg gatgagtgc tgagcaacc
2761 ctgcccggcg ggcacctgta ccgaccact ggccctctc acctgcacct gcccgccagg
2821 ctacggaggc ttcactgctg aacaggacct gcccgactgc agcccagct cctgcttcaa
2881 ttggcgggacc tgtgtggac gcgtgaactc gttcagctgc ctgtgccgtc ccggctacac
2941 aggagcccac tgccaacatg aggcagacc ctgcctctc cggcctgcc tacacgggg
3001 cgtctgcagc gcccccacc ctggtctccg ctgcacctgc ctcgagagct tcacgggccc
3061 gcagtgcagc acgtggtgg attggtgac ccgccagcct tgcataaac ggggtcgtg
3121 cgtccagact ggcctcatt gctctgtcc cctggatgg agcggacgc tctgtgacat
3181 ccgaagcttg ccctgcaggg aggcgcgac ccagatcgg gtgcgctgg agcagctgtg
3241 tcaggcgggt gggcagtggt tggatgaaga cagctcccac tactgctgt gccagaggg
3301 ccgtactggt agccactgtg agcaggaggt ggaccctgc ttggcccagc cctgccagca
3361 tggggggacc tgccgtggct atatggggg ctacatgtgt gagtgtctc ctggctaaa
3421 tgggtgataa tgtgaggac acgtggacga gtgtgctcc cagccctgcc agcacgggg
3481 ttcattgcatt gacctcgtg cccgctatct ctgctcctgt ccccaggaa cgtgggggt
3541 gctctgcgag ataatgag atgactgcg cccaggccca ccgctggact caggggccc
3601 gtgcctacac aatggcacct gcgtggacct ggtgggtgg ttcggctgca cctgtcccc
3661 aggatacact ggtttgcgt gcgaggcaga catcaatgag tctcgtcag gtgctgcca
3721 cgggcacac acccgggact gcctgcagga cccaggcgg ggtttccgt gccctttgta
3781 tgctggcttc tcaggtcctc gctgtcagac tgtcctgtct ccctgcgag cccagccatg
3841 ccagcatgga gccagtgcc gtcctagccc gggctcctgg ggtggctga ccttcacctg
3901 tcactgtgcc cagccgttc ggggtccgc ttgcgagcgg gtggcgcgt cctgccggga
3961 gctgcagtg ccggtgggc tcccattgca gcagacgcc cgcgggccc gctgcctg
4021 cccccagggt ttgtcgggac cctcctgccc cagcttccc ggtcgcgc cggggccag
4081 caacgccagc tgcgcggcc cccctgtct ccacgggggc tctgcgcgc ccgcgcctg
```



4141 cgcgcccttc ttcgcgtgcg cttgcgcgca gggetggacc gggccgcgct gcgagggcgc  
4201 cgcccgggca cccgaggtct cggaggagcc gcggtgcccg cgcgcccgct gccaggccaa  
4261 gcgcggggac cagcgtgcg acccgcgagt caacagccca ggctgcggct gggacggcgg  
4321 cgactgctcg ctgagcgtgg ggcaccctg cgggcaatgc gaggcgtgc agtgctggcg  
4381 cctcttcaac aacagccgct ggcaccctgc ctgcagctcg cccgctgccc tctacgaaa  
4441 cttcgactgc cacgcccgtg gccgcgagcg cacttgcaac ccggtgtacg agaagtactg  
4501 cgccgaccac tttgccgacg gccgctgca ccagggctgc aacacggagg agtgcggtg  
4561 ggatgggctg gattgtgcca gcgaggtgcc ggccctgctg gcccgcggcg tgctgggtgct  
4621 cacagtgtcg ctgccgccag aggagctact cgttccagc gccgactttc tgcagcggct  
4681 cagcggcatc ctgcgcacct cgtcgcgctt ccgctggac gcgcacggcc aggcctagg  
4741 cttcccttac caccggccta gtcctggctc cgaaccccgg gccctcggg agctggcccc  
4801 cgaggtgatc ggctcggtag taatgctgga gattgacaac cggctctgccc tgcatgccc  
4861 tgagaatgat cactgcttcc ccgatgccc gagcggcgt gactacctgg gaggctgtc  
4921 agcgggtggag cgcctggact tcccgtacc actgcgggac gtgcggggg agccgctgga  
4981 gcctccagaa cccagcgtcc cgtcgtgccc actgctagtg gcggcgctg tcttgctgct  
5041 ggatcattctc gtcctgggtg tcatggtggc ccggcgcaag cgcgagcaca gcaccctctg  
5101 gttccctgag ggcttctcag tgcacaaggc cgtggcctct ggtcacaagg gccggcgagg  
5161 acccgtgggc cagcagcgc tgggcatgaa gaacatggcc aagggtgaga gcctgatgg  
5221 ggaggtggcc acagactgga tggacacaga gtgccagag gccaaagcggc taaaggtaga  
5281 ggagccaggc atgggggctg aggaggctgt ggattgccc cagtggactc aacaccatct  
5341 gggtgctgct gacatccgag tggcaccagc catggcactg acaccaccac agggcgacgc  
5401 agatgctgat ggcatggatg tcaatgtgcg tggcccagat ggcttcccc cgctaagtct  
5461 ggcttccctc tgtggggggg ctctggagcc aatgccaaact gaagaggatg aggcagatga  
5521 cacatcagct agcatcatct ccgacctgat ctgccagggg gctcagcttg gggcacggac  
5581 tgaccgtact ggcgagactg ctttgacct ggctgcccgt tatgcccgtg ctgatgcagc  
5641 caagcaccag cactgctgag gggcagacac caatgccag gaccactcag gccgcactcc  
5701 cctgcacaca gctgtcacag ccgatgccc ggggtgtcttc cagatttctca tccgaaaccg  
5761 ctctacagac ttggatgccc gcatggcaga ttgctcaacg gactgatcc tggcgccccg  
5821 cctggcagta gaggcatgg tggaaagact catcggcagc catgctgatg tcaatgctgt  
5881 ggatgagctt gggaaatcag ccttacctg ggctgcggct gtgaaacaac tggaaagccac  
5941 tttggccctg ctcaaaaatg gagccaataa ggacatgcag gatagcaagg aggagacccc  
6001 cctattcctg gccgcccgcg agggcagcta tgaggctgcc aagctgctgt tggaccactt  
6061 tgccaaccgc gagatcaccg accacctgga caggctgccg cgggacgtag cccagagag  
6121 actgcaccag gacatcgtgc gcttgctgga tcaaccagt gggccccgca gccccccgg  
6181 tccccacggc ctggggcctc tgetctgtcc tccaggggccc tctctccctg gcctcaaagc  
6241 ggcacagtgc gggccaaga agagcaggag gccccccggg aaggcggggc tggggccgca  
6301 ggggcccccg gggcggggca agaagctgac gctggcctgc ccgggcccc tggctgacag  
6361 ctgggtcacg ctgtcggccc tggactcgtc ggactcccc cggccttctg gtgggcccc  
6421 tgcttccccct ggtgcttcc cccttgagg gccctatgca gctgcccactg cactgcagt  
6481 gtctctggca cagcttggg gcccaggccc ggccgggtcta gggcggcagc ccctggagg  
6541 atgtgtactc agcctgggccc tgetgaacc tgtggctgtg ccctcgatt gggcccggct  
6601 gccccacct gcccctccag gcccctgctt cctgctgcca ctggcggcgg gaccacctg  
6661 gctcaaccaca gggacccccg tctccccgca ggagcggccc ccgcttacc tggcagtccc  
6721 aggacatggc gaggagtacc cggcggctgg ggcacacagc agcccccaa agggccgctt  
6781 cctgcggggtt cccagtgagc acccttacct gacccatccc cccgaatccc ctgagcactg  
6841 ggccagcccc tcacctccc ccctctcaga ctggtecgaa tccacgcta gccagccac  
6901 tgccactggg gccatggcca ccaccactgg ggcactgccc gccagcccac tcccttgtc  
6961 tgttcccagc tcccttgcct agggccagac ccagctgggg ccccagccc aagttacccc  
7021 caagaggcaa gtgttggcct gagacgctcg tcagtctta gatcttgggg gcctaaagag  
7081 acccccgtcc tgctcctttt ctttctctgt ctcttcttc cttttagtct ttttcatcct  
7141 cttctctttc caccaaccct cctgcatcct tgccttgcag cgtgaccgag ataggctatc  
7201 agcccagggc ttcagtcttc ctttatttat aatgggtggg ggctaccacc caccctctca  
7261 gtcttgtgaa gactctggga cctccttctt cccacttct ctcttcccctc attcctttct  
7321 ctctccttct ggcctctcat ttccttacac tctgacatga atgaattatt attattttta  
7381 ttttctttt tttttttaca ttttgtatag aaacaaatc atttaaaaaa acttattatt  
7441 attattttt acaaaatata tataatggaga tgctcccctc ccctgtgaa cccccagtgc  
7501 ccccgctggg ctgagtctgt gggcccattc ggccaagctg gattctgtgt acctagtaca  
7561 cagcagatgac tgggatcccg tgtaccagat acacgaccca ggtatgtacc aagtaggcac  
7621 ccttgggcgc acccactggg gccaggggtc gggggagtgt tgggagcctc ctccccacc  
7681 cacctccctc acttactgc attccagatg ggacatgttc catagccttg ctggggaagg  
7741 gccactgcc aactccctct gcccagccc cacccttggc catctcccct tgggaactag  
7801 ggggctgctg gtgggaaatg ggagccagg cagatgtatg cattcctttg tgtccctgta  
7861 aatgtgggac tacagaaga ggagctgccc gagtggctact ttcttctct ggtaatcctc  
7921 tggcccagcc tcatggcaga atagaggtat ttttaggcta tttttgtaat atggcttctg  
7981 gtcaaaaatc ctgtgtagct gaattcccaa gccctgcatt gtacagcccc ccactcccct  
8041 caccacctaa taaaggaata gttaacactc aaaaaaaaaa aaaaaaaaaa

//

**Anexo 2** - GeneBank – NOTCH3 (AC004663) – Páginas 65 a 81.

[<http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&id=310818>]

Search  for

[Limits](#)
[Preview/Index](#)
[History](#)
[Clipboard](#)
[Details](#)

Display  Show   Hide:  sequence  all but gene, CDS and mRNA features

Range: from  to   Reverse complemented strand Features:  SNP

**1:** [AC004663](#). Reports Homo sapiens chro...[gi:3108186]

[Links](#)

[Features](#) [Sequence](#)

**LOCUS** AC004663 41150 bp DNA linear PRI 04-MAY-1998  
**DEFINITION** Homo sapiens chromosome 19, cosmid R29783, complete sequence.  
**ACCESSION** AC004663  
**VERSION** AC004663.1 GI:3108186  
**KEYWORDS** HTG.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** [Homo sapiens](#)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
 Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 41150)  
**AUTHORS** Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,  
 Burkhart-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stilwagen,S.,  
 Phan,H., Velasco,N., Garnes,J., Danganan,L., Poundstone,P.,  
 Christensen,M., Georgescu,A., Avila,J., Liu,S., Attix,C.,  
 Andreise,T., Trankheim,M., Amico-Keller,G., Coefield,J., Duarte,S.,  
 Lucas,S., Bruce,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A.,  
 Montgomery,M., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S.  
 and Carrano,A.V.  
**TITLE** Sequence analysis of an ~1.5 Mb olfactory receptor (OLFR) cluster  
 in 19p13.1  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 41150)  
**AUTHORS** Lamerdin,J.E.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (04-MAY-1998) Joint Genome Institute, Lawrence Livermore  
 National Laboratory, 7000 East Ave., Livermore, CA 94551, USA  
**FEATURES**  
**Location/Qualifiers**  
**source** 1..41150  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:[9606](#)"  
 /chromosome="19"  
 /map="19p13.1 from OLFR to MEL"  
 /clone="R29783"  
 /cell\_line="5HL2-B"  
 /clone\_lib="LL19NC03 R chosome 19-specific cosmid  
 library"  
 /note="Cosmid library constructed at LLNL from flow-sorted  
 chromosomes from hybrid 5HL2-B, which carries chromosome  
 19 as its only human chromosome."  
[misc\\_feature](#) 1..41150  
 /note="Map and sequence oriented from p telomere to  
 centromere. Cosmid R29783 overlaps cosmid R34078 to the  
 left by 1.506 kb and cosmid R33209 to the right by 5.302  
 kb. Additional chr 19 map and sequence information is  
 available at  
<http://www-bio.llnl.gov/bbrp/genome/genome.html>."  
[repeat\\_region](#) 860..965  
 /rpt\_family="LINE2"

[repeat\\_region](#) 1184..1389  
 /rpt\_family="MER20"  
[repeat\\_region](#) 1445..1745  
 /rpt\_family="AluSq"  
[gene](#) complement(2790..>40735)  
 /gene="NOTCH3"  
 /note="NEUROGENIC LOCUS NOTCH HOMOLOG 3"  
[misc\\_feature](#) 2791..3873  
 /note="DDS similarity to overlapping ESTs:  
 (2791..3277) AA348213 EST54656 Hippocampus I Homo sapiens  
 cDNA 3' end; Score: 908 Identity: 473/485 (97%).  
 (2791..3279) AA557201 nl74h09.s1 NCI\_CGAP\_Br2 Homo sapiens  
 cDNA clone IMAGE:1056449; Score: 966 Identity: 486/489  
 (99%).  
 (2791..3285) AA495983 zw06d07.s1 Soares NhHMPu S1 Homo  
 sapiens cDNA clone 768493 3'; Score: 986 Identity:  
 494/495 (99%).  
 (2795..3310) AA007598 zh99b03.s1 Soares fetal liver spleen  
 lNFLS S1 Homo sapiens cDNA clone 429389 3'; Score: 831  
 Identity: 479/521 (91%).  
 (3280..2819) W92249 ze14c12.r1 Soares fetal heart NbHH19W  
 Homo sapiens cDNA clone 358966 5'; Score: 780 Identity:  
 445/472 (94%).  
 (2880..3286) AA481651 zv06b10.r1 Soares NhHMPu S1 Homo  
 sapiens cDNA clone 752827 5'; Score: 802 Identity:  
 404/407 (99%).  
 (3312..2802) H04572 yj16c08.r1 Homo sapiens cDNA clone  
 148910 5'; Score: 729 Identity: 465/488 (95%).  
 (3409..2991) AA484871 ne8ld05.s1 NCI\_CGAP\_Ew1 Homo sapiens  
 cDNA clone IMAGE: 910665; Score: 797 Identity: 413/418  
 (98%).  
 (3523..3037) AA195359 zr33h05.r1 Soares NhHMPu S1 Homo  
 sapiens cDNA clone 665241 5'; Score: 928 Identity:  
 478/485 (98%).  
 (3463..3075) AA481418 zv06b10.s1 Soares NhHMPu S1 Homo  
 sapiens cDNA clone 752827 3'; Score: 766 Identity:  
 386/389 (99%).  
 (3621..3269) T79722 yd71g05.r1 Homo sapiens cDNA clone  
 113720 5'; Score: 634 Identity: 349/361 (96%).  
 (3707..3365) AA348214 EST54657 Hippocampus I Homo sapiens  
 cDNA 5' end; Score: 674 Identity: 340/343 (99%)  
 (3873..3616) AA301460 EST14422 Testis tumor Homo sapiens  
 cDNA 5' end; Score: 504 Identity: 255/258 (98%).  
 Additional EST matches:  
 T47069 and T47068."  
[repeat\\_region](#) 3396..3428  
 /rpt\_family="AT\_rich"  
[repeat\\_region](#) 3457..3495  
 /rpt\_family="POLY\_A"  
[repeat\\_region](#) 3685..3776  
 /rpt\_family="(GAAA)n"  
[CDS](#) complement(join(3819..4871,5622..5719,8525..8672,  
 8944..9248,10406..10568,13243..13327,13488..13710,  
 13828..13982,17225..17557,20682..21247,21980..22098,  
 22182..22439,22521..22653,23229..23413,23838..23985,  
 24118..24319,24733..24958,27452..27607,28063..28176,  
 28414..28565,28644..28836,30035..30145,30262..30495,  
 31038..31151,31392..31505,32146..32331,32430..32585,  
 34581..34814,34902..35024,35117..35455,35534..35676,  
 40657..>40735))  
 /gene="NOTCH3"  
 /note="Notch protein homolog 3- Human; Residues 1..41 of  
 Notch3 protein can be found in cosmid clone R33209  
 (AC004257) which overlaps to the right of cosmid R29783"  
 /codon\_start=3  
 /product="Notch 3"  
 /protein\_id="[AAC15789.1](#)"

/db\_xref="GI:3108187"  
/translation="PPCLDGSPCANGGRCTQLPSREAACLPPGWVGERCQLEDPCHS  
GPCAGRGVCQSSVVAGTARFSCRCPRGFRGPDCSLPDPCLSSPCAAGARCSVGPDRF  
LCSCPPGYQGRSCRSDVDECRVGEPCRHHGGTCLNTPGSFRCQCPAGYTGPLCENPAVP  
CAFPSRCRNGGTCRQSGDLTYDCACLPGFEGQNCNEVNVDDCPGHRCLNGGTCVDGVNTY  
NCQCPPEWTGQFCTEDVDECQLQPNACHNGGTCFNLTGGHSCVCVNGWTGESCSQNI  
DCATAVCFHGATCHDRVASFYCACPMGKTGLLCHLDDACVSNPCHEDAICDTNPVNGR  
AICTCPPGFTGGACDQDVDECSIGANPCEHLGRCVNTQGSFLCQCGRGYTGPRCETDV  
NECLSGPCRNOATCLDRIGQFTICMAGFTGTyceVDIDECQSSPCVNGGVCKDRVNG  
FSCTCPSGFSGSTCQLDVDECASTPCRNGAKVDQPDGYECRAEGFEGTLCDRNVDD  
CSPDPCHHGRCVDGIASFSCACAPGYTGTRCESQVDECRSOPCRHGGKCLDLVDKYL  
RCPSTGTGVNCEVNIDDCASNPCFTGVCRDGINRYDCVCQPGFTGPLCNVEINECASS  
PCGEGGSCVDGENGFRCCLPPGSLPPLCLPPSHPCAHEPCSHGICYDAPGGFRCVCEP  
GWSGPRCSQSLARDACESQPCRAGGTCSSDGMGFHCTCPPGVQGRQCELLSPCTPNPC  
EHGGRCESAPGQLPVCSCPQWQGPCRQDQVDECAGPAPCGPHGICTNLGSGFSTCH  
GGYTGPSCDQDINDCDPNPCLNGGSCQDGVGFSFSCSLPGFAGPRCARDVDECLSNPC  
GPGTCTDHVASFTCTCPPGYGGFHCEQDLPCSPSSCFNGGTCVDGVNSFSLCRPGY  
TGAHCQHEADPCLSRPCLHGGVCSAAHPGFRCTCLESFTGPOCQTLVDWCSRQPCQNG  
GRCVQTGAYCLCPPGWSGRLCDIRSLPCREAAAQIGVRLEQLCQAGGQCVDEDESSH  
VCPEGRTGSHCEQEVDPCLAQPCQHGGTCRGMGGYMCECLPGYNGDNCEDDVEDECAS  
QPCQHGGSCIDLVARLYLCSPPGTLGVLCINEDDCGPGPPLDSGPRCLHNGTCDVLDV  
GGFRCTCPPGYTGLRCEADINECRSGACHAAHTRDCLQDPGGGFRCLCHAGFSGPRCQ  
TVLSPCESQPCQHGGQCRPSPGGGGLTFTCHCAQPFWGPCRERVARSCRELQCPVGV  
PCQOTPRGPRCACPPGLSGPSCRSFPGSPPGASNASCAAAPCLHGGSCRPAFLAPFFR  
CACAQGTGPRCEAPAAAPEVSEEPRCPRAACQAKRGDQRCDRECNPSGCGWDGGDCS  
LSVGDWPWRQCEALQCWRLFNNRCDPACSSPACLYDNFDCHAGGRERTCNPVYKEYCA  
DHFADGRCDQGCNTEECGWDGLDCASEVPALLARGVLVLTVLLPPEELLRSSADFLQR  
LSAILRTSLRFRDLAHGQAMVFPYHRPSPGSEPRARRELAPEVIGSVVMLEIDNRLCL  
QSPENDHCFPDAQSAADYLGALSVERLDFPYPLRDVVRGEPELPEPVSPLPLLVAG  
AVLLLVILVLGVMVARRKREHSTLWFPEGFSLHKDVASGHKGRREPVSQDALGMKNMA  
KGESLMGEVATDWMDETECEAKRLKVEEPMGAEAEAVDCRQWTOHHLVAADIRVAPAM  
ALTTPQGDADADGMDVNVVRGPDGFTPLMLASFCGGALEPMPTTEEDEADTSASII  
ICQGAQLGARTDRTGETALHLAARYARADAAKRLLDAGADTNAQDHSGRTPHHTAVTA  
DAQGVFQILIRNRSTDLDARMADGSTALILAAARLAVEGMVEELIASHADVNAVDELGK  
SALHWAAAVNNVEATLALLKNGANKMDQDSKEETPLFLAAREGSYEAALKLLLDHFANR  
EITDHLDRLPDVAQERLHQDIVRLDQPSGPRSPPGPHGLGPLLCPGAFPLPGLKAA  
QSGSKSRPPGKAGLGPQGRGRGKLTACPGPLADSSVTLSPVDSLSPRPFGGP  
PASPGGFPLEGPYAAATATAVSLAQLGGPGRAGLGRQPPGGCVLSLGLLNPVAVPLDW  
ARLPPPAPPGPSFLLPLAPGPQLLNPGTPVSPQERPPPYLAVPGHGEYPAAGAHSSP  
PKARFLRVPSEHPYLTPSPESPEHWASPPPSLSDWSESTPSPATATGAMATTTGALP  
AQPLPLSVPSSLAQAQTQLGPQPEVTPKRQVLA"

- [repeat\\_region](#) complement(4542..4591)  
/rpt\_family="GC\_rich"
- [repeat\\_region](#) 4954..5036  
/rpt\_family="MIR"
- [repeat\\_region](#) complement(5140..5438)  
/rpt\_family="AluSx"
- [repeat\\_region](#) complement(5516..5578)  
/rpt\_family="Alu"
- [repeat\\_region](#) complement(5911..5969)  
/rpt\_family="(CA)n"
- [repeat\\_region](#) complement(6066..6361)  
/rpt\_family="AluJb"
- [repeat\\_region](#) 6397..6700  
/rpt\_family="L1"
- [repeat\\_region](#) 6706..7007  
/rpt\_family="AluJo"
- [repeat\\_region](#) 7009..7378  
/rpt\_family="L1MB8"
- [repeat\\_region](#) 7419..7718  
/rpt\_family="AluSq"
- [repeat\\_region](#) complement(7737..7875)  
/rpt\_family="MER3"
- [repeat\\_region](#) 7886..7950  
/rpt\_family="MER20"
- [repeat\\_region](#) complement(7953..8233)  
/rpt\_family="AluJb"

[misc\\_feature](#) complement(8944..9248)  
/gene="NOTCH3"  
/note="DDS similarity to N23491 yx13d12.r1 Homo sapiens  
cDNA clone 261623 5' similar to SP :NOTC\_XENLA P21783  
XOTCH (328..23); 98% identity."

[misc\\_feature](#) complement(10406..10593)  
/gene="NOTCH3"  
/note="DDS similarity to N23491 yx13d12.r1 Homo sapiens  
cDNA clone 261623 5' similar to SP :NOTC\_XENLA P21783  
XOTCH (512..329); 92% identity."

[repeat\\_region](#) complement(10715..10991)  
/rpt\_family="AluSx"

[repeat\\_region](#) complement(11042..11071)  
/rpt\_family="POLY\_A"

[repeat\\_region](#) complement(11094..11243)  
/rpt\_family="AluJo/FRAM"

[repeat\\_region](#) 11276..11367  
/rpt\_family="MIR"

[repeat\\_region](#) 11709..12008  
/rpt\_family="AluSx"

[repeat\\_region](#) complement(12081..12233)  
/rpt\_family="FRAM"

[repeat\\_region](#) 12235..12327  
/rpt\_family="MLT1B"

[repeat\\_region](#) complement(12340..12393)  
/rpt\_family="(GA)n"

[repeat\\_region](#) 12399..12566  
/rpt\_family="MLT1B"

[repeat\\_region](#) complement(12618..12773)  
/rpt\_family="AluSq"

[repeat\\_region](#) complement(14211..14488)  
/rpt\_family="AluSx"

[repeat\\_region](#) complement(14507..14819)  
/rpt\_family="AluJb"

[repeat\\_region](#) complement(14828..15039)  
/rpt\_family="AluSc"

[repeat\\_region](#) complement(15051..15352)  
/rpt\_family="AluSp"

[repeat\\_region](#) complement(15821..16109)  
/rpt\_family="AluSx"

[repeat\\_region](#) complement(16121..16261)  
/rpt\_family="(GAA)n"

[repeat\\_region](#) complement(16262..16285)  
/rpt\_family="POLY\_A"

[repeat\\_region](#) 16297..16598  
/rpt\_family="AluY"

[repeat\\_region](#) complement(16607..16689)  
/rpt\_family="FLAM\_A"

[repeat\\_region](#) 16724..16901  
/rpt\_family="AluSg/x"

[repeat\\_region](#) 16902..16968  
/rpt\_family="(GAA)n"

[misc\\_feature](#) 17228..17330  
/note="DDS similarity to R48628 yj68b06.r1 Homo sapiens  
cDNA clone 153875 5' (1..103); Score: 190 Identity:  
99/103 (96%)"

[repeat\\_region](#) complement(17729..18030)  
/rpt\_family="AluSx"

[repeat\\_region](#) complement(18041..18205)  
/rpt\_family="FRAM"

[repeat\\_region](#) complement(18220..18339)  
/rpt\_family="LINE2"

[repeat\\_region](#) 18408..18710  
/rpt\_family="AluSq"

[repeat\\_region](#) complement(18711..18786)  
/rpt\_family="(TA)n"

[repeat\\_region](#) complement(18840..19075)

[repeat\\_region](#) /rpt\_family="LINE2"  
19133..19250  
/rpt\_family="(CAGA)n"  
[repeat\\_region](#) complement(19321..19618)  
/rpt\_family="AluSp"  
[repeat\\_region](#) 19636..19939  
/rpt\_family="AluJb"  
[repeat\\_region](#) 19949..20248  
/rpt\_family="AluSg"  
[repeat\\_region](#) 20254..20422  
/rpt\_family="AluSg/x"  
[repeat\\_region](#) 20494..20615  
/rpt\_family="(CAGA)n"  
[repeat\\_region](#) 20589..20675  
/rpt\_family="(TGGA)n"  
[repeat\\_region](#) complement(20912..21165)  
/rpt\_family="GC\_rich"  
[misc\\_feature](#) complement(21109..21247)  
/gene="NOTCH3"  
/note="DDS similarity to overlapping ESTs:  
(21247..21109) AA284113 zc40b04.T7 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 3'  
similar to TR:G483581 G483581 NOTCH 3 (353..488); 95%  
identity.  
(21247..21110) W47159 zc40b04.s1 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 3'  
similar to PIR:S45306 S45306 notch 3 protein - mouse  
(136..1); 89% identity."  
[repeat\\_region](#) 21265..21384  
/rpt\_family="(GGAA)n"  
[repeat\\_region](#) complement(21547..21840)  
/rpt\_family="AluSx"  
[misc\\_feature](#) complement(21962..22098)  
/gene="NOTCH3"  
/note="DDS similarity to overlapping ESTs:  
(22098..21962) W47280 zc40b04.r1 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 5'  
similar to PIR:S45306 S45306 notch 3 protein - mouse  
(237..383); 87% identity.  
(22098..21980) AA284113 zc40b04.T7 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 3'  
similar to TR:G483581 G483581 NOTCH 3 (234..352); 100%  
identity.  
(22098..21980) W47159 zc40b04.s1 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 3'  
similar to PIR:S45306 S45306 notch 3 protein - mouse  
(254..137); 98% identity."  
[misc\\_feature](#) complement(22182..22418)  
/gene="NOTCH3"  
/note="DDS similarity to overlapping ESTs:  
(22418..22182) W47280 zc40b04.r1 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 5'  
similar to PIR:S45306 S45306 notch 3 protein - mouse  
(1..236); 99% identity.  
(22355..22182) W47159 zc40b04.s1 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 3'  
similar to PIR:S45306 S45306 notch 3 protein - mouse  
(422..255); 88% identity.  
(22416..22182) AA284113 zc40b04.T7 Soares senescent  
fibroblasts NbHSF Homo sapiens cDNA clone 324751 3'  
similar to TR:G483581 G483581 NOTCH 3 (1..233); 99%  
identity."  
[repeat\\_region](#) 22781..23057  
/rpt\_family="AluSx"  
[misc\\_feature](#) complement(24789..24958)  
/gene="NOTCH3"  
/note="DDS similarity to overlapping ESTs:



(24958..24789) emb|F13534|HSC32B031 H. sapiens partial cDNA sequence; clone c-32b03 (91..260); 98% identity.  
(24958..24845) T65500 yc73f03.r1 Homo sapiens cDNA clone 21534 5' similar to SP:NOTC\_XENLA P21783 XOTCH (79..191); 95% identity."

[repeat\\_region](#) 25258..25380  
/rpt\_family="MER33"

[repeat\\_region](#) 25382..25683  
/rpt\_family="AluJb"

[repeat\\_region](#) 25684..25796  
/rpt\_family="MER33"

[repeat\\_region](#) complement(25797..26098)  
/rpt\_family="AluSp"

[repeat\\_region](#) complement(26100..26141)  
/rpt\_family="AT\_rich"

[repeat\\_region](#) 26157..26427  
/rpt\_family="AluJo"

[repeat\\_region](#) 26429..26536  
/rpt\_family="MER33"

[repeat\\_region](#) complement(26555..26855)  
/rpt\_family="AluSp"

[repeat\\_region](#) 26926..27214  
/rpt\_family="AluJo"

[repeat\\_region](#) 27234..27348  
/rpt\_family="MIR"

[misc\\_feature](#) complement(28063..28142)  
/gene="NOTCH3"  
/note="DDS similarity to overlapping ESTs:  
(28412..28063) T65500 yc73f03.r1 Homo sapiens cDNA clone 21534 5' similar to SP:NOTC\_XENLA P21783 XOTCH (1..78); 94% identity.  
(28150..28063) emb|F13534|HSC32B031 H. sapiens partial cDNA sequence; clone c-32b03 (1..90); 92% identity."

[repeat\\_region](#) complement(29062..29232)  
/rpt\_family="AluJ"

[repeat\\_region](#) complement(29294..29421)  
/rpt\_family="AluJ"

[repeat\\_region](#) complement(29523..29576)  
/rpt\_family="(CA)n"

[repeat\\_region](#) complement(29577..29861)  
/rpt\_family="AluSx"

[repeat\\_region](#) complement(32674..32975)  
/rpt\_family="AluSq"

[repeat\\_region](#) complement(33027..33103)  
/rpt\_family="L1MC4"

[repeat\\_region](#) 33150..33449  
/rpt\_family="AluJo"

[repeat\\_region](#) complement(33739..34043)  
/rpt\_family="AluJo"

[repeat\\_region](#) complement(34051..34177)  
/rpt\_family="FLAM\_C"

[repeat\\_region](#) complement(34186..34489)  
/rpt\_family="AluSx"

[repeat\\_region](#) 35791..36091  
/rpt\_family="AluSx"

[repeat\\_region](#) complement(36099..36294)  
/rpt\_family="LINE2"

[repeat\\_region](#) 36324..36397  
/rpt\_family="MLT1D"

[repeat\\_region](#) complement(36398..36559)  
/rpt\_family="AluSg/x"

[repeat\\_region](#) 36655..36947  
/rpt\_family="AluSx"

[repeat\\_region](#) 36959..37256  
/rpt\_family="AluSg"

[repeat\\_region](#) 37257..37418  
/rpt\_family="AluJo/FRAM"



[repeat\\_region](#) 37442..37681  
 /rpt\_family="MLT1C"  
[repeat\\_region](#) complement(37737..37825)  
 /rpt\_family="LINE2"  
[repeat\\_region](#) 37833..38131  
 /rpt\_family="AluSp"  
[repeat\\_region](#) 38165..38229  
 /rpt\_family="(CA)n"  
[repeat\\_region](#) complement(38859..39051)  
 /rpt\_family="GC\_rich"  
[repeat\\_region](#) complement(39329..39355)  
 /rpt\_family="GC\_rich"  
[repeat\\_region](#) complement(39911..39977)  
 /rpt\_family="MIR"  
[repeat\\_region](#) complement(40126..40201)  
 /rpt\_family="MIR"  
[repeat\\_region](#) complement(40890..40916)  
 /rpt\_family="GC\_rich"

ORIGIN

```

1 gatcagtgca gatctggagg caaaaagcag acactcagct gttttatcat tgtcctttca
61 gtcaaaaaac caactaggat gtcaattcca aagtcccca cccttcccaa cttcagcatg
121 ccaatgaaat cacatacata tgaatccttg aatctcatta gttaattttc ctccatcacc
181 agtgcaggct tcctagtcct caacctaaaca tcctttctgc aatcattcca gggagctccc
241 agtgtgggtg agtctcacat gtgcttacia ataagtcagc tccaccagggt acgaaagtgg
301 gttcccgtg tgggacactc ttcttcaata ttctgttgac atgaagccaa ggttgcatga
361 tagagtggac aaaatgggta gggagaagga ctttctcctc atttgtacag cacttgggaa
421 gaagaatacc caaacttgct taagccccag tagatatttt attttttatt cccaagaatc
481 tactcatgca taaagctaag atggacttct cttgggcca tgattcaagt gcaaggaatt
541 tattcgcaag atcattccag aaagctctag taggagttag gaagagttag ttgctgtgtg
601 gggcaactgg ggcttaatcc caaagagcat ccttggggag actctgtgga tcatatccat
661 gaattcccac agagagtggc gaggaagtct gggattttat ccaccatctc tcatcctacc
721 ctcatatggt gaggggtgct cgtgccccac acaagagaaa gttcccagggt ggagagtctc
781 aggtgttcag tcagatgctg tcagcaagca ctagaatggt aagttccaag ggtatacagg
841 ggtgcacccc ccatacatcc ctcaactagc ctgaaattcc ataaaagcag agactgggcc
901 tatataattc accattctat tctcaacaac ttgtccggga gttggcacac ggtaggtgcc
961 caatacaggt tagagggtaa gggataaata ataaaataaa gttgagacia acctaaaaga
1021 gactgtttca tcttcaaaga cttagcctaa aggggggtgg agaggagtaa acatattatta
1081 gaagcttcca cagaattcat gcttctgcaa acctgattgc ccctaaattt agcacatcag
1141 tagttccaaa gaagagactt taagacatgc tttaagacia caacagggtg gaactctgct
1201 cccagatgac aattggcaat gtctggagac atgttatatt ggttgtcaca actggaaaag
1261 ggggtgctgct ggtagctaat gggtagggac cagagatgct attaaatata ctacaatgca
1321 caggacagct ccccacagca agaatgatt tggcccaaa tgtcaataat gccgaaatta
1381 agaaacctta ggttagaggg tggatacgtg aaccaattgg agtcaataga aaaatcaacc
1441 acaaggtcga gtgcatgac tcatgcctgt aatcccagca ctttgggagg ccaaggtgga
1501 tggatcacct gaggtccgga gttcagagacc agcctggcca acatggtgaa atcctgtctc
1561 tactgaaaat acaaaaatta gccaggcatg gaggtggaca cctgtaatcc cagctacttg
1621 ggagactgag gcaggagaat tgcttgaacc cggaggcggg ggatgcagtg agccgagatt
1681 gtgccattgc actccagcct gggtgacaag aacaagactc tatcttaaaa aaaaaaaaaa
1741 aaaaaggaag aaagtaaaag aaaaatcaat cacaactaata ggagtttctc tcttccctgg
1801 aaggagcaat atggcaccta ttttgataac acagagagaa ggctggagct gtgtgatgtc
1861 gggaaagagag gtttactagg gtgccatgtg gagcatgtag atgaaactga tgcccagaga
1921 aagctgtaga gatgaagaga aactgagcca ttgttgacca catttgagcc actagatcaa
1981 ttctccctg aagtaggtcc ttttaattta tatgagctaa taaaccctt tcttgtttaa
2041 gccagtagac ttttgtgtca tttgcaacca ggagcactcc aactgacaga atcaacgaac
2101 attaaatcct atcctccttt cctgaaagca ccagtgtaaa cagaaccttg tcttggagca
2161 cgactggttg cgagttaacag aaaccaccg aagctggcac aagtaaaaga gtatttatta
2221 catgggtcca aggttatctc caaaacctat aggagcatga agtaaaagt aaattcatgc
2281 agaaccatga tcagatactt agaaagcaat taataaatac ctggtgatct caatggaatg
2341 aacactcaat taactggaat tgataactgc actgggagcc gagataaacc cccagccct
2401 caaggagcca tttgagctct catttctgat tctccctgaa tattagcttc attctctgctc
2461 cctgtttgct tgtctctttg cccaccacct cgaattggcc actgatattt ctaaatctga
2521 ttgtgacctt aagagctcca ccaactgaca actttcttag cctctccatt aaagaaagt
2581 ctgagcaaaa agaactctgat tggctcagct taagtcagggt gacaaccact aatccaatcg
2641 gtagtcatct gccaaaagag gaagcataag tagactggca agtcaaagt atttaccca
2701 agatgggagc tcaagttagc cctgggtggg ggagatagaa gtcccactgc caccactccc
2761 caccttacct agacacagac caacaacact gagtgtaac tattccttta ttagtggtg
2821 aggggagtgg ggggctgtac aatgcagggc ttgggaattc agctacacag ggattttgac
  
```

2881 cagaagccat attacaaaa tagcctaaaa atacctctat tctgccatga ggctgggcca  
2941 gaggattacc aggaagagaa agtaccactc aggcagctcc tcttcttgta gtcccacatt  
3001 tacagggaca caaaggaatg catacatctg ccctggctcc ctttcccac cagcagcccc  
3061 ctagtcccca aagggagatg gccaaagggtg gggctggggc agagggagtt ggcagtgggc  
3121 ccttccccag caaggctatg gaacatgtcc catctggaat gcagtgaagt gagggaggtg  
3181 gggctgggag gaggctccca acactccccc gaccctggc cccagtgggt gcgccaagg  
3241 gtgcctactt ggtacatacc tgggtcgtgt actcggtaaca cgggatccca gtcatgcctg  
3301 tgtactaggt acacagaatc cagcttggcc gaatgggccc acagactcag ccccagggg  
3361 gcaactgggg gttcacaggg ggaggggagca tctccatata tatattttgt aaaaaataat  
3421 aataataagt ttgttttaaat gaatttgttt ctatacaaaa tgtaaaaaaa aaaagaaaaa  
3481 taaaaataat aataattcat tcatgtcaga gtgtaaggaa atgagaggcc agaaggagag  
3541 agaaaggaat gagggagag agaagtgggg aagaaggagg tcccagactc ttcacaagac  
3601 tgagaggggt ggtggtagcc cccaccatt ataaataaag gaagactgaa gccctgggct  
3661 gatgacctat ctcggtcacg ctgcaaggca aggatgcagg agggttggtg gaaagagaag  
3721 aggatgaaaa agactaaaag gaaggaagag acagagaaaag aaaggaggca ggacgggggt  
3781 ctcttttaggc cccaagatc taagaactga cgagcgtctc aggccaacac ttgcctctg  
3841 ggggtaactt cgggctgggg cccagctgg gtctgggct gagcaaggga gctgggaaca  
3901 gacaagggaa gtggctgggc aggcagtgcc ccagtgggtg tggccatggc cccagtggca  
3961 gtggctgggc taggcgtgga ttcggaccag tctgagaggg agggaggtga gggctggcc  
4021 cagtgtcag gggattcggg ggatggggtc aggtaagggt gctcactggg aaccgcagg  
4081 aagcgggctt ttggggggct gctgtgtgcc ccagccgccc ggtactcctc gccatgtcct  
4141 gggactgcca ggtaaggcgg gggccgctcc tgcggggaga cgggggtccc tgggtgagc  
4201 agctggggtc ccggcgccag tggcagcagg aacgaggggc ctggaggggc aggtgggggc  
4261 agccgggccc aatcgagggg cacagccaca gggttcagca ggcccaggct gactacacat  
4321 cctccagggg gctggcgccc tagaccgccc cggcctgggc caccaagctg tgccagagac  
4381 actgcagtgg cagtggcagc tgcatagggc ccctcaaggg ggaagccacc aggggaagca  
4441 gggggcccac cgaaaggccg cggggagtcc agcagatcca cgggcgacag cgtgaccgag  
4501 ctgtcagcca gggggcccgg gcaggccagc gtcagcttct tgccccgcc cggggcccc  
4561 tgcggcccca gccccgcctt cccggggggc ctctgctct tcttggacc cgactgtgcc  
4621 gctttgaggg cagggagga ggcccctgga ggacagagca gaggccccag gccgtgggga  
4681 ccgggggggc tgcggggccc actgggttga tccagcaagc gcacgatgtc ctggtgcagt  
4741 ctctcctggg ctacgtcccg cggcagcctg tccaggtggt cgggatctc acgttggca  
4801 aagtgttcca acagcagctt ggcagcctca tagctgccct cgcggggggc caggaatagg  
4861 ggggtctcct cctggggggc cagaaccac agaggtcagc gaaaccocag ctaacagtag  
4921 caaagtcttc cagcctcttg ggggagccc agagctatga gtttgtacac tggagcttga  
4981 cagacctggg ttaaateccg gctgtgctgg gggagcctgg acaagttctt tcctctaac  
5041 cacattcatc catcacctac taagtccag aggtcaaggc caggactagg gtgagttaag  
5101 cgaggtactc accttaggca taactttttt tttcttgtct tttttttttt ttttttttga  
5161 cagagtctca ctgtgtcgcc caggctggag tgcagtggcg ccatttcagt tcaactgcaac  
5221 ctacagcttc cgggttcaag cgattctcct gcctcagcct cccgagcagc tgggactaca  
5281 ggtgtgcacc actacgccc gctaattttt gtatttttaa tagagacagg gttttactac  
5341 gttggccagg ctggtctcaa actcctgaac tcaagtaat cacccatctc gacctccaa  
5401 agtgcaggga ttatagcat gaaccaccac acccagccat taggcacaaa atttaagggg  
5461 ccccaaaaaa ctttagtcac caaagtaaat tatattttaa tccaatggtt ggggttttat  
5521 tttgttttgt tttttgagag agtctcactc tgttgcccag gctggattgc aatggcactg  
5581 tgccactgct gacaccaggt ggaccaaggg ctggggctca ccttgctatc ctgcatgtcc  
5641 ttattggctc cttttttgag cagggccaaa gtggcttcca cgttgttcc agccgcagcc  
5701 cagtgtgagg ctgatttccc tggaggatga aggggagaga ggtcaggacc aggcacctgt  
5761 cctggggccc ccatgtcagg gaggcagaaa ttgtctagaa tgtggcctag gggtagagga  
5821 ggcaggcctg ccgtgaatat ccaagtactt gtgtctggat ccatgtgtgt gcaaatctgc  
5881 aaaaatgactc tgtgcaatgg ggaagcactt tgtgtctgtg tgtgtgtgtt tgtgtgtgtg  
5941 tgtgtgtgtg cgtgcatgtg catgactgtt tatacacagg ggggtgctgtg tggacctcct  
6001 tcagttgaac ctagggttgt gacatccaat acagtagtca gtcacatgtg gtttttattt  
6061 atccatttat ttatttttga ggtagtgtct tgttctgtca cccaggctgg agtccattgg  
6121 cacaatctca gttcactgca acctctacct gctgggctca ggcgattttc ctgctcagc  
6181 tcccctagta gctgggacta caggcatgta ccaccatgcc cagctgattt tttgtatttt  
6241 tgatagagat agagttccac tatgttgccc aaactggtct caaactccta agctcaagtg  
6301 atccaccagc cttggcctcc caaagtgttg ggattacagg catgagccac cgtgcccaac  
6361 ctacatgtgg ttcttttagat ttaaatttaa taaaataggt gccaggacca ttcaatgggg  
6421 aaagaatggt cttctcaact aatggtgctg gaacaactga atattcacat tcaagaatg  
6481 aagttggacc cctacctcac actttataca aacactaact caaaatggat cagaaacctc  
6541 aatatatgag ctccaactat agaactaac atctagaaga aaacataaga gtaaatcttc  
6601 atgaccttga attaggccaa gattcttaga tgtgacacca aaaccacaaa taacaataac  
6661 aactacaaaa cagataaatt ggacatgatc aaaattgaaa cttttggcca ggcacagtgg  
6721 ctacgcctta taatctcaac aatttgggag gccaaagggt gagtattgct taaggccagg  
6781 agtttgagat cagcctgggc aacaaagtga gaccctattt ccacaaaaaa ttttaagaatt  
6841 acccaggtat gctggcacac acctgtaata ccagctattc aggaggggtg gatgggagga

6901	tcacttgagc	ctgggagttt	gaggctgcag	agagccgtaa	tcatgccact	gcactccagc
6961	ctgagtgaca	gagtgatacc	ctatgtctaa	aaataaaaaga	aaaaaaaaagt	ggtccatcga
7021	cacaatggat	tattattcag	ctgaactatg	ctacaacatg	ggtgaacctt	gagaacataa
7081	tgctaactga	aaaaaaggca	gacaaaaaag	accacgtatc	atatgattac	atztatatga
7141	aatgtccaga	acaggcaaat	ccatgaagac	aaaaagcaaa	tcagtggttg	cacagggctg
7201	gacagatgag	gggcttggca	gatgatggcc	aagtggtgag	aatttatttt	tcaggtgatg
7261	aaatgttcta	aaattgattg	tggtaatagt	tgtatgactc	tgattatact	aagagccatt
7321	gaattgtaca	ctttaaattg	gtgaattgta	tggcatgtta	attatatctc	aataaagcag
7381	ttttttgttt	tttgtttttt	gttttttaaaa	aaaaggatgg	actcagtggc	tcatgcctat
7441	aaccccagca	ctttgggagg	ccaaggcagg	tggatcacct	gaggtcagga	gtttgaggcc
7501	agcctggcca	acattgcgaa	accccgtctc	tactaaaaat	acaaaaaaat	agccaggcat
7561	ggtgatgggc	acctttagtc	ccagacactt	gggaggctga	gacaggagaa	tactggaac
7621	ctgggtagtg	gaggttgcag	tgagccaaga	tcacaccatt	gcactccagc	ctgggtgaca
7681	agagtgcagc	tctgtctcaa	aaaaaaaaaat	taaaaaaagg	gcaacataag	gagaaaaaaa
7741	ttaaataaac	taaaacattt	acttcttcat	ttgcctcagt	cacatttcaa	gtcactatat
7801	gtggccagtg	gctactatac	tggacagtgc	agagaataat	tccatcatca	cagaaagtcc
7861	tattggccag	ccctgcctca	gaactgtggc	tctcaactag	gagaaactcc	ctgccccag
7921	ccccagagc	atgtggcaac	gtctggagac	tgtttttttt	ttttgagatg	gagtctgtcc
7981	cccaggctgg	agtgcagggg	cacaatcacg	gctcacagac	tcaacctcct	gggctcaaac
8041	gaccctccca	cctcagtttc	tggagtagct	agggccacag	gtatgcacca	ccgcacatgg
8101	ctaactcttc	aaaaaacttt	tgtagaacct	actatgttgc	ccagtctggt	ctcaaactcc
8161	tggcctcaag	tcatectcct	gcctcagcct	cccaaagtgt	tgggatcaca	ggcatgagcc
8221	accacacatg	gcctgatgtg	tacgtttatt	ataataactt	ccagcagggtg	gaagtacagt
8281	gtgtccgtct	aacaaagcag	gtggaaatga	agtatctagg	ggaaaggcac	ctttttctaa
8341	atccctgagg	caccctaagg	actagtgggtg	accctgcatg	accctgcagg	cttcatgaag
8401	cctggttatg	tgtgcgtgag	cttcagtgat	tgggtctcaa	catgggtatg	aatttgcacc
8461	aacatgacct	tcagggccca	ggtgacacca	accagctta	gacttgattc	ctctgccaac
8521	ctaccaagct	catccacagc	attgacatca	gcatggctgg	cgatgagctc	ttccaccatg
8581	ccctctactg	ccaggcgggc	cgccaggatc	agtgccgttg	agccatctgc	catgctggga
8641	tccaagtctg	tagagcggtt	tcggatgaga	atctaggaca	gagagtggat	gcagcaggag
8701	gggtcatggc	aggaacagag	gaatcaggag	caccctaaag	tcccatgaag	tcccaacct
8761	ccatacccca	accctgaaa	ttggtgagg	tcagggagca	gctgcttgac	agatactgta
8821	ttcccatata	tccccatttt	ccaaatgaga	aaaaatgagt	ctgaaaggca	gaactggggc
8881	tcaaaccag	gtaagtctaa	tgccctgccc	agctctgagg	tccaaagtgt	gtgcctatct
8941	cacctggaag	acaccctggg	catcggtgtg	gacagctgtg	tgcaggggag	tgcggcctga
9001	gtggtcctgg	gcattggtgt	ctgccccagc	atccagcagc	cgcttggtg	catcagcagc
9061	ggcataacgg	gcagccaggt	gcaaagcagt	ctcgccagta	cggtcagtc	gtgccccag
9121	ctgagcccc	tggcagatca	ggctggagat	gatgctagct	gatgtgtcat	ctgcctcatc
9181	ctcttcagtt	ggcattggct	ccagagcccc	cccacagaag	gaagccagca	ttagcggggt
9241	gaagccatct	gcagggacag	gagtggttca	gcaggaaggt	aaacacaggg	ccttttccag
9301	gaaatgttat	caagggatcc	ccattaggag	ctaattgggac	acatggaaaa	atgacaatta
9361	gcagatcctc	ctgtctgcac	acccacacc	agaatgagga	ttcgtgggca	cgtgatttgt
9421	cacagaacta	ctcctagggc	acagggaaaca	caggacaagg	gactctggga	gtcaagcaag
9481	tgtgcaacct	ccagtaaatt	cccaaagggt	tattccacct	tgatcctcca	aagggactct
9541	ggaacatgtg	ttgtaccca	gagttcta	ccaactcaag	gcaaagaaac	catccccttc
9601	atactcctgc	acccactggt	taggagctgc	ccctttctgc	caattcattt	gagtgtttta
9661	aaaggggggg	ggaaaaaaag	agctgcccc	aggacataa	aatcccaggc	acttcaggcc
9721	ccagcagcca	gagggacttc	aaaagagagc	cataagtaca	ggccatttgg	agcataagca
9781	aatgaagct	aaggcaggaa	tgccctgaaa	cataacaaag	atggatctaa	gaggatcaa
9841	caacatagtc	cgctacactg	cacacactgt	acacaagcac	ttgtcctcat	ttttcttagg
9901	gagccacttg	catccggttc	tccattctca	agccctgggt	agagactgac	ctccccgctc
9961	caccagcaga	cccgaccccc	accagctcaa	agacaggtat	gtgaatcagg	cttgggtcaat
10021	ttgcatattc	cagccccctg	gtcatactga	ttggctcaag	gatgagcacg	tgaccctagt
10081	ctgtccaatc	agaatgcac	ctgagacttt	tctggagcat	taagaaggtt	ttgtaagatt
10141	ggagtcaaag	ggtagagagg	ggaaaactga	gtcctggatc	caaagtgtgc	acttttagtt
10201	aaatgagttg	ataaatccct	tgttttactt	aagctagtct	gagttggacc	tttttactt
10261	ccaaccaaga	gtgttgctga	atacacacac	ccttcacaga	agcttagaga	ctccccagt
10321	tctcccaaaa	acacagagtc	agaaataacc	tctcacaggt	aggatgggtg	aggggcatcc
10381	ctttgggagg	ggcactgtca	ctaacctggg	ccacgcacat	tgacatccat	gccatcagca
10441	tctgcgtcgc	cctgtggtgg	tgctcagtgc	atggctgggtg	ccacgcggat	gtcagcagca
10501	accagatggt	gttgagtcca	ctgacggcaa	tccacagcct	cctcagcccc	catgcctggc
10561	tcctctacct	ggaggggcag	gcaccctgga	tctcagatct	caccctttgg	tgtgggggag
10621	agccactgcc	agggatgggg	ctgggcattc	ctacaggttt	ctcttttagga	gatacacaca
10681	gagccctggg	aaatcattta	ccatctgttt	ggtgttttat	ttgtttttga	gacagagtct
10741	cgctctgtcc	cccaggctgg	agtgcagtgg	catgatctca	gctcactgca	acccccgct
10801	ccgggttca	agtgattctc	ctgcctcagc	ctctggagta	gctgggatta	caggcacatg
10861	ccacggtgtc	cagctaattt	tttgtatttt	tagtagagaa	ggggtttcac	catgttggcc

10921 aggctggtct cgaactcctg acctcaggtg atcagcccct ctcagcctcc caaagtgctg  
10981 agattaacag gaaatcaatg tgtgccatag cccaaggagg taataaaagg aagacacaag  
11041 gtttcttttt tttttttttt ttggtctttt tgagtaataa atggctttaa ttatttatgt  
11101 atatatTTTT tagagactga gtctcactgt gtcactcagg ctggagtgca gtgctgcaat  
11161 catacctcac tgcagcctca acctcctggg ctcaaagat cctcccacct cagcggccca  
11221 agtagctggg actataagca cacagataaa gcagattggg atagcatctg atccggttc  
11281 aagttccaac cccaccctt agtagctgga aggctctggg caaatcactt tgaagctctg  
11341 agcttctcat tggcaaaatc aggatgacgc agccattcca ccatgctaag ataccaccac  
11401 aattacattg ggccaccctt ttaatacctc taatacctac tttgggatgg gaaggcagaa  
11461 ggaaaagtct gggttgagca ggttttttta tctgtccact gtaatagatg attccatttc  
11521 acaacattaa aatatggaca gaggaagggg gtttgtaaat tgaagtatat gtcaatagct  
11581 actttacagc gtcctcagag agatgaaata agatgtttta tgtgcattag agacattcag  
11641 cgggtgcctc agtaatgcta gcaggatctg tatggactca attgtgcccc ctctaaaaat  
11701 catacaatgg cggggcacgg tggctcacgt ctgtaatccc agcactttgg gatgccgagg  
11761 cgggtggatc acttgaggtc tgcagcttga gaccagcctg gccaacatgg tgaacccca  
11821 tctctactaa aaatacaaaa attagccggg catggtagca catgcctata atcccagcta  
11881 ctcgggaggc tgaggcaggg gaatcacttg aacctgggag gtggaggctg caatgagcca  
11941 agatcacgcc actgcactcc agcttgggca acagagtgag acctggtctc aaaaaataaa  
12001 taaataaata aacagaaatc atacactaaa accctaaata atatataaaa tttccatgtg  
12061 cctgcatttg aagataggac tttattttta gacaggtct tgcctctgta cccaggctgg  
12121 agtgaagtga tgcagtcata gctcactgca ggctccaact cctggactca agtgatcctc  
12181 ctgcctcagc ctcccagta gctgaaatta ctggtgcaag ctaccaagcc tggtgagata  
12241 gggcctttaa ggaggtaat aaggttataat gaggtataa atgtgggacc ctaatccaat  
12301 aggattggtg tccttacaag aagaggatgg gatgtcaaat ctgtctctct ctctctctct  
12361 ctctctctct ctctctctct ctctctcccc tctgtgtaca cagaggaaaag gccatgtgag  
12421 gctacagtga gaaggcggcc atctgcacac caggaagaga gcccttacca ggaaccagcc  
12481 ctgacagcac cttgggcttg gacttccaac ctccagaacc gtgaaaaaat taatttctgt  
12541 tgtttaggtc acccatgctg tgggactttt tttttttttt gagacggagt ttcgctcttg  
12601 ttgccagtc tggagtgcca ggttcaaggg attctcctgc ctcagcctcc caagtagctg  
12661 ggattacagg tgcccgccac catgcccagc taatcttttg tatttttagt acagacgggg  
12721 tttcacatg ttggtcaggc tggctcgaa ctctgacct caggtgatcc acctctggtg  
12781 gattttttaa tggcagccaa agctactca tactgaacct aggggcatga cccatgagc  
12841 tcagaaccgg ctcagagcac tttcaggcta agtccatgct ggacaaaata ggaactgaca  
12901 gagataaggc agactcgctc tctgggctgg gggacacct ccaggaagag tgcaggctgg  
12961 ccaggaagca ggatcctggc tggggagggg caggatggag gggaggattt tccacctgcc  
13021 cggagggacc cttggggccc cagtggcaaa tcttctctcc tctgaggtea cagaaccaca  
13081 ctgagaagaa agggttgcaa agctgggact attctctgga gctagggagg tggggacagc  
13141 ggtgccatcc cctgatcacg cccatcatcc actggggacc ccaccagta ccctgagggg  
13201 aggggtcaga ggagggggca aaggctcagag ggggggcagt acctttagcc gcttggcctc  
13261 tgggcaactc gtgtccatcc agtctgtggc cacctcccc atcaggctct cacccttggc  
13321 catgttcctg gcgacaatg ggaagagggg atgtgagggg gacactagag gtgtccagct  
13381 gggaaggagg atttgggggt gggctcagag tctgagtcag gtcaaggcgg ggtccagggt  
13441 tcacaaggct cccgtagtca gggacagggg gcgagcaggg ttctcacttc atgccagcg  
13501 cgtcctggcc cacgggttcc cgccggccct tgtgaccaga ggccacgtcc ttgtgcagtg  
13561 agaagccctc agggaaccag aggggtgctgt gctcgcgctt gcgcccggcc accatgacac  
13621 ccaggacgag aatgaccagc agcaagacag cgcccgccac tagcagtggc agcagcggga  
13681 cgctgggttc tggaggctcc agcggctccc ctaagagcag gaagcagagg gcggggcttc  
13741 agccgagggc ggggcttttg cctcaggcgg ggcttcggcc gccccagct ccgccccgc  
13801 caccctctcc ccaaggcagg gccgcacccc gcacgtcccg cagtgggtac gggagttcca  
13861 ggcgctccac cgctgacaac gctcccaggt agtcagcggc gctctgggca tcggggaagc  
13921 agtgatcatt ctcaggcgac tgcaggcaga gccggttgtc aatctccagc attactaccg  
13981 agctgcaggg acagcagggg gggaccagag ggtcaaaaa ttgcccgatg caccctctgg  
14041 taccaagccc cactttccct ccccagcgc ttggtctgaa gcgccatctc caccacagg  
14101 tctccatggc ccacctgcat ccacctggca tccaaccca ctgcacctc aaaacgtgtc  
14161 ttgctgtggt cagcagctcc cccagagcca tgaatccaat ggatattatt ttttgatctt  
14221 gcaccgaggg tggagtgcag tgggtgcgac tgggtcact gcaacttccc ctctgggtt  
14281 caagcagttc tcctgcctca gcctcccaag tagctgggat tacaggcgtg tgccaccatg  
14341 cccagctaatt ttttttgtat ttttagtaga gacggggtt caccacgttg gccaggctgg  
14401 tcttgaactc ctgacctcaa gtgatcctcc ctgccttggc ctcccaaagt gctgggatta  
14461 caggcgtgag ccgccactgt gcctggcctg tttttttgt tgtttgtttg ttctgtttt  
14521 ttttttgaga cggtttctact ctgtcggcca ggctagagca cagtggcgaa atcagagctc  
14581 actgtagcct tgacctcccg ccttcaagca atctttccac ctcatctcc caaagatctg  
14641 ggaccacaga catgtgccac cacacctggc taatttttgg ggggggttgg ggggtgaggc  
14701 atagagatta ggtgtcacta tgttgcccag gctagtcttg aactccgggg ctcaagcaat  
14761 cctcccacct cagcctccca atgtgctggg ttacaggcc tgggccaacta cgcccagcca  
14821 agggacattg tttgtttttt gtttttgaga tgaagtcttg ctctgtcgac aggctggagt  
14881 gcagtggtgc aatctcagct cactgcaacc tccgactccc tggttcaagt gattctctgc

14941 ctcagcctcc tgagtagctg ggattacagc cacgcgctgc cacgcccagg taatttttgt  
15001 attttttagta gagacggggt ttcaccgtgt tggccaggac attgtatttg tttgttttgt  
15061 tgcttgtttt gagacagagt ttcactcttg ttgcccaggc tggagtgcaa tggtgctatc  
15121 tcagctcgcc acaacctcca cgtccagggt tcaggaatt ctctgctc agcctctga  
15181 gtagctggga ttacaggcat gcgccaccac gccccagtaa ttttgtatt ttagtataca  
15241 cagggtttct ccatgttggt caggctggtc ttgaactccc gacctcaggt gatccaccg  
15301 cctcaacctc ccaaagtgct gggatcacag ttgtgagcca tcacacctgg cctctcctc  
15361 ttttttgagc taattccatt catccaatcc ttaggtctcc gtgtccccag attaattctt  
15421 ctgttccttc tctccccact cctttgagtc cctctcatgg cttcaatgat catcaacact  
15481 tttatatctt tcttgtttgt gtctttagcc cccaggatc tccttgagg ccttcccaat  
15541 acttccaggt ccctcagact ctgcaagctc caaatgacc tgtttttcac tttgtggccc  
15601 tatcactcag cttagccagg ggctggtaca ccaaccactc tccttcaccc ctctaccct  
15661 ccccttccac caccacatc ccaatatctg caaaacccaa ccctcctctc tgtatcagag  
15721 acccaccctc ctcaggtctc tgccttctcc tcaacaatag ctacacctcc cttctggtac  
15781 ctctaccagg ccactagcag gatacttcta tttttttct ttttctttt ttttttttt  
15841 gagacagagt ctggctctgt tgccccggct ggagtgcagt agcacaatct tggctcactg  
15901 caacctcac ttctgagtt caagtgattc tcatgcctca gcctcctgag tagctgggac  
15961 taaaggcata caccaccacg cctggctaatt ttttgtatt ttttagtagag acagggtttt  
16021 gccatgttgc ccaggctggt ctcaaactcc tgacgtcaaa tgaactgcct tccaaagtgc  
16081 tgggattaca ggcgtgagcc accgagcccc tcctatttct ttcttcttct tcctctcct  
16141 cctcctcctc cccctcccc tcttcttct tcttctctt cctccccctc cccctcccc  
16201 tccctcctcc ctcttcttct tcttcttct cttcttctt ttcttcttct tcttcttct  
16261 cttttttttt ttttttttt tttttaagag atcggaggcc gggcatggtg gctcacgcct  
16321 gtaatccag cactttggga ggccgaggca ggtgatcac gaagtcagga gatccagacc  
16381 atcctggcta acacactgaa acccctctc ttactaaaa tacaaaaaat tagccaggca  
16441 tgggtggcggc gcctctgagt cccagctact cgggaggctg aggcaggaga atggcatgaa  
16501 cccgggaggc ggagcttgca gaaagccgag atcgcgctc tgcactctag cctgggcgac  
16561 agagcgagac tccgtctcaa aaaaaaaaaa aaaaaaaaaa agagagagag atgggagct  
16621 tgctatgttg cccaggctag atttgaattc ctgggctcaa gtgatcctcc aagtagctg  
16681 gactacaaga atatatgtg cccggcttta tcttctatgt ctttaagaag aagaaaatta  
16741 gcccggtgtg gtggtggatg cctgtaatcc cagctactct ggaggctgag gcagagaatt  
16801 gcttgaacct aggaggcaga ggttgcagtg agctgagatc acgccactgc actccagcct  
16861 gggcgacaga gactccgtct caaaaaaaaaa aaaaaaaaaa aaagaagaag gagaaggaga  
16921 aggagaagga gaaggagaag gagaagaaga agaagaagaa gaaaaaaacc ctacacatgt  
16981 atttaaccac tctgccaagc cacttagcaa cctgcctccc cacaaagccag cactagacct  
17041 gggcaattct accccagaag cactagctcc atctggaggc attttttgct tgtcacagcc  
17101 aggaggcggg gctcctgaca tctggtgggt aaaggcatta agtgaaatga aacacacaag  
17161 acctggatca acagcatctc tctctcccc agccaccacg gcttttccag gtggggtcac  
17221 tcacccgatc acctcggggg ccagctcccg acgggcccgg ggttcggagc caggactagg  
17281 ccggtggtaa ggaagacca tggcctggcc gtgcgcgtcc aggcggaagc gcagcgaggt  
17341 gcgcaggatg gcgctgagcc gctgcagaaa gtcggcgtg gaacgcagta gctcctctgg  
17401 cggcagcagc actgtgagca ccagcacgcc gcgggccagc agggccggca cctcgtggc  
17461 acaatccagc ccatcccagc cgcactcctc cgtgttgag ccctggctgc agcggccgtc  
17521 ggcaaagtgg tcggcgcagt acttctcgta caccgggctg gtggggaagg gtgaggcaga  
17581 gaggggcgga gtcaggggtc agaggagaca atccccttc atgcattcac accgtagagt  
17641 acagagactc cagacaaggc ttgcacaggg gactggaata tacggacca tgcaggctt  
17701 tctgcaaadc ttgccattca ctggataatt ttttttttt ttttttgaga cggagtcttg  
17761 ctctgtcggc caggccggaa tgcagtagca tgatctctca gctcactgca gcctccacct  
17821 ccccggttcc agcgattctc ccgcctcagc ctcccaagta gctgggatta caggcatgca  
17881 ccaccagcc cgtctaattt ttgtattttt agcagagacg gggttttact atgttgcca  
17941 ggctgctggt gaactcctga cctcaggtga tccacctgcc tcggcctccc aaagtctgg  
18001 gatcacaggc ataagtcacc gcacgcagcc cttgataata tttttttaa ttgattgatt  
18061 gaaatggggt cttgctctgt tgcccacgcc ggagtgcagt ggtgcaatca tagctctctg  
18121 caatctcaa ctectggact caaattatcc tccctcttca gcctcccaag tagctaggac  
18181 tacaaccca caccacactc ccggcctcac tggctaacta ttaattaatc agtaataact  
18241 aactgactac catgtgttgg gcactctgc aggcacaagg ataaggccg gccacaattc  
18301 ctggtccacg tgcagctggc attccactgg gtgaagacag atagtaaggg agatacatga  
18361 gccaaatata cagcatctcc attactcatc aataaaataa aaatacaggc tgggtcgggt  
18421 ggctcatgcc tctaattcca gcactttggg aggccaaagg gtacggataa cctgaggtca  
18481 ggagttctag gccagcctgg ccaaaatggt gaaaccccc tctctactaa aaatacaaaa  
18541 attagtcagg catggtggcg ggcgcctgta atcccagcta ctcggaagc tgaggcagga  
18601 gaatagcttg aaccaggag gtggagggtg cagtgagcca agatcacacc actgcactcc  
18661 agcctgggca atgggagaga aatcctgtct aaaaaaaaaa aaaaaaaaaa tacatatata  
18721 tatatatata tgtatatata tgtatatata cacacacaca cacacacaca cgcacgcaca  
18781 cacatagtgc tggatggaat taaatgctaa gcagaaaaat aaagcagagg aagggtggg  
18841 ggtgaccagg gaggtgctca ctgagaagg accaccattt cagcaaagac ttgaaggagg  
18901 tgaggaggga gccatgtggg ggaagggaac agccaggcag taggaacagc ttgtgcaag

18961 gccctgatat gggatcaagc ctggggatc tgggaaacag cagagaggcc aggggtggctg  
19021 gaacagagtg agtgagggga aagcagagaa gatgaggatg tggagattaa gggggtggag  
19081 ggttccaagc agacagagag gtacgcagag gacagaaggg gcaggggtgta gggatggaca  
19141 gatggactgc ctaggacaga cagagagagg ggagaagag gggcagagat agtcaggggc  
19201 acccacagat agccatatag aggagcaaaa agagtgcatt gatggacaga ttccggcccaa  
19261 ggatggacaa acagagatgg aggaagacag ttgggtacac agaggaacaa aagcaatddd  
19321 tttttttttt tttttttttt gtgagacaga gatttactct tgctgcccag gctgagatgc  
19381 aatggcacga tctcagctca cctcatcctc cacctcccag gttcaagcaa ttctcctgcc  
19441 tcagcctcct gagtagctgg gattacagcc atgtgccacc acaccgggct aatdddgtat  
19501 ttttagtaga gacggagttt caccatgttg cccaagctca aactcctgac ctcagggtgat  
19561 ctgccctcct ctgcctccca aagtactggg attacaggcg tgagcactgt gcccgcccaa  
19621 taaaaggaat ttttaggcta gatgtggtgg ctcaggccca caatcccagc actttgagag  
19681 gctaagctgg gtggattgct ttgagctcag gtgttcaaga ctagcctgga caacatggcg  
19741 aaaccccatc tctaagaaaa agcacaaaag ttaccggggc gcgggtggctc acacctgtgg  
19801 tcccagctcc tagggaggct gaggtgtgag ggttacttga gcccaggagg tagagattgc  
19861 actgagccaa gatcacacca ctgcaactca gcctgggcca cagagcaaga ccctgtccta  
19921 aaaaaaaaaa aaaaaaaaaa aaaaaggagc caggtgcggt ggctcacacc tgtaatcca  
19981 gcactttggg aggccaaaggc gggcggatca caaggtcagg agttcgtgag cagcctggcc  
20041 aatatggtga aaccctgtct ctactaaaaa taaaaaaaaa aagccaggcg tgggtggcagg  
20101 tacctgtagt cccagctact cgggaaggctg aggcaggaga atcgctttaa cccgggaggc  
20161 ggaggttgca gtgagccaag atggagccat tgcactccag tctgggcaag agagcgagac  
20221 tccgtcttaa aaaaaaaaaa aaaaaaaaaa aggaattcgg ctgggcacgg tggctcacac  
20281 ctgtaatccc agctactcca gaggtgagg caggagtact gcttaaacc gggaggcaga  
20341 ggttgcaagt agccgagatc gcgccactgc actctagcct gggccacaga gcaagactcc  
20401 atctcgggaa aaaaaaaaaa aagtaattat tggatggaga aaagggcaga caaacggggc  
20461 agagaagtgc ataggcagat agaaacaaga acaagataga cagagcacat ggatgaacag  
20521 acacacggac agacacgtgg gacacaccga tgggcagata gagtgcacag agaaacagac  
20581 ggggcaagtg gatgggcagg tggaaagaga ggcagggcc acggacaaac agactgggat  
20641 ggatgcatag acagacggat cgatcgggtg gatgggtca cttgcaagtg cgctcgcggc  
20701 caccggcgtg gcagtcgaag ttgtcgtaga ggcaggcgg cgagctgcag gcggggtcgc  
20761 accggtggtt gttgaagagg cgcagcact gcagcgcctc gcattgccc caggggtcgc  
20821 ccacgctcag cgacagctcg ccgccctccc agccgcagcc tgggctggtg cactcgcggt  
20881 cgcagcgtg tgcccgcgcg ttggcctggc aggcggcgcg cgggcaccgc ggtcctccg  
20941 agacctcggg tgccgcggcg ggcgcctcgc agcgcggccc ggtccagccc tgcgcgcaag  
21001 cgcagcggaa gaagggcgcg agcggcgcg ggcggcagga gccccctgg agacagggg  
21061 cggccgcgca gctggcgttg ctggcccccg gcggcgacc cgggaagctg cggcaggagg  
21121 gtcccgacaa ccctgggggg caggcgcagc gcggcccgcg gggcgtctgc tggcatggga  
21181 cgcccaccgg gcaactgcagc tcccggcagg agcgcgccac ccgctcga caagcaccac  
21241 agaacggctg ggggtcgggt ttggggaggg agggataaaa atgaggggtg ggagtggagg  
21301 gaaggaggag aagaaatgga ggagtgggga aaagaagagt caaggggaga gagggggaag  
21361 agaagaggtc aagactaagg aaggttgaga cccctcaac ctccccatgt cccaccaac  
21421 cgctcccctt tccacactgg agaccccacc tggcagataa gccctgcat cccaggaggc  
21481 cacgcccccc aatcaccatc ctcacccttt tccctcaaa cctccaggaa atgtcggggc  
21541 agggggtgtt ttctttggtt ttgagacgg agtgtcactc agtcgccagg ctggagtgca  
21601 atggcgtgat ggctcactgc aacctccgc tcccgggttc aagtgattct cctgcctcag  
21661 cctcccagat agctgggatt acaggcagct gccactacgc cgggtaatt tttatatttt  
21721 tagtagagac ggggtttcac catggtgccc aggtggtct ctaactcctg gcctcaagtg  
21781 atcctcccgc ctcggcctcc caatgtgctg ggattacagg tgtgagctac cacaccggc  
21841 ggaaatgtca gcattttcca gaaactccct tccctcga gtctccccta aagccacatc  
21901 ctctccttag acgccagcc cctactactc cagaggccac gccctactc ctctccaaa  
21961 ggccgccacc cacacctacc tgggcacagt gacaggtgaa ggtcagccca ccccaggac  
22021 ccgggctagg acggcactgg cctccatgct ggcattgctg ggactcgcag ggagacagga  
22081 cagtctgaca gcgaggacct gagcagcgg gagcatgtag atcagccaca atgggggaat  
22141 gacaggcggg gtcccaggcc agccccttcg ccaacgctta cctgagaagc cagcatgaca  
22201 aaggcaacgg aaacctccgc ctgggtcctg caggcagtc cgggtgtgtg ccgctgggga  
22261 ggcactgag cgacactcat tgatctctgc ctcgcagcgc aaaccagtgt atcctggggg  
22321 acaggtgca gggaaaccac ccaccagctc cacgcaggtg ccattgtgta ggcaccggg  
22381 ccctgagtcc agcgggtggc ctgggcccga gtcactcctca ttaatctcgc agagcaccac  
22441 tgggggaaga aacgaggggt ggtcaagagg gaatgaagac agcctctcat cctgtcccc  
22501 caaccctggc cctggcatac ccagcgttcc tggggacag gagcagagat agcgggccac  
22561 gaggtcaatg catgaacccc cgtgctggca gggctgggag gcacactcgt ccacgtcgtc  
22621 ctcacagtta tcaccattgt agccaggag acacttcagt ggggtaagag agggaccac  
22681 tcagcttagt gggacacaga cccacctgga cataccatg aagacgcaa gaaccccagc  
22741 acaaaaggac acacacagat gttcagcga ctaccagtag gtaatcccag cactttggga  
22801 ggccgagggt ggcagatcac ctgagatcag gagttcgaga ccagcctggc caacatggtg  
22861 aaaccctgtc actatgaaaa atacaaaaat tagctgggcg tgggtggggg tgctgtact  
22921 cccagctata atgagaggct gaagcagaag aatcacctga acccaggagg caaaggttg

22981 agtgagctga gatcgcgcca ctgcactcgt ctgggtgaca gagtgagaac ctgtctcaaa  
23041 aaaacaaaca aataaaaaaa caccacagat ctaccagag tgacaccac agatacacca  
23101 agagtcacaa aagcaaaaac tcagacatat ctagagacat acccatacca agccacacag  
23161 aaatgtgtgc ccagacgcac ccaagcatgc ccacctctc tccctctcc tggggagcgc  
23221 ccccttacct cacacatgta gcccccata tagccacggc aggtccccc atgtctggcag  
23281 ggctgggcca agcaggggtc cacctcctgc tcacagtggc taccagtacg gccctctggg  
23341 cacagcagtg agtgggagct gtcttcatcc acacactgcc caccgctcg acacagctgc  
23401 tccagccgca cccctgcaaa gaggagagtg gcacaggaac agaggtaacc ccatacatcc  
23461 cccttctgcc tccatcacac agatcattca acatccttgg tggaatgagt ccccaggatc  
23521 gcaccacac tccatcaggt tgtcacacat cccccagca ggcaaggctc catcactgac  
23581 atacacacac atgccttcaa cgctcacagc agctcggta caccctgtag cacacacacc  
23641 atccctgctg tctcacagt acagacacac accatcccgg ccagttgtca catgctcaca  
23701 caccacacac actcagccac acaccccac atgcccata ggctctgtgg cacccccatt  
23761 cggctcacac tagcaggagg tacgtgcatg agccccttc caaggcccca cacgcccgcc  
23821 cacatgctcc cactcaccga tctgggctgc ggctccctg cagggcaagc ttcggatgct  
23881 acagaggcgt ccgctccatc cagggggaca aaggcaatag gcccagctc ggacgcagcg  
23941 acccccgttt tgacaaggct ggcggctgca ccaatccacc agcgtctgga ggggaagcac  
24001 tcagagtcag tactgtgggg tgggggtag tctgggagaa actgtgcccc gacttggtct  
24061 ctccctccta ggatcccagg caggctcctc cccagggtcc ccagtaactc caccacctg  
24121 gcaactgccc cccgtgaagc tctcagggca ggtgcagcgg aagccagggt gggcggcgt  
24181 gcagacgccc ccgtgtaggc agggccgcga gaggcaggg tctgcctcat gttggcagtg  
24241 ggctcctgtg tagccgggac ggcacaggca gctgaacgag ttcacgccgt ccacacaggt  
24301 cccgccattg aagcaggagc tggagcggaa ggagtgggag ggaggatcag gctccgcccc  
24361 ctcacaggcc ctgccctcct tctgagctc cgctgttagc gctggggacg tcccactccc  
24421 tgacccttat ctcggaaga agctgcagcc ccagcagaag ccaatcaggt cccttagctc  
24481 caattcacca gggacttcag ccctaagggt agagccacct ggaagccatg ccctctcagg  
24541 caccagagac tccaagggtga gctccaaaca gacgcctgc tccttccctg gccctgcccc  
24601 atcagtcac agagtccgca gtgggtacca agctgaggac tcccccaagt cgttgctgct  
24661 cccacttcgt cccagggtccc agggcccac ccaagccaga gtcctgctc tccaagcaga  
24721 ggccccgccc acctggggct gcagtcgggc aggtcctgtt cgcagtggaa gcctccgtag  
24781 cctggcgggc aggtgcaggt gaaggagcc acgtggctcg tacaggtgcc cgggcccag  
24841 gggttgctca ggcactcatc cacatcgcgg cgcctcgtg gcccggcga accagggagg  
24901 caggagcagg aaaaggagcc cacgccgtct tggcacgagc caccgttcag gcatgggtct  
24961 gcggacagga ggaaggcgg ctggctaccc accttgcctc cattagccca gtctgactgg  
25021 gatatgcctt ggattgagaa gaattcagga ggaaccagg ttcaactggg atgggggtgca  
25081 tttggctcaga ggaatctaaa tccatcccct tctgtggcct cgtccccctc tttatctgga  
25141 aacctcacct atgtggcaag ggcaggatgt tgggaaaacc aaaaggctcc ttaagctgc  
25201 caagcaaacg ctgctaattg tgctgatct aattcctgtg cttggaacag gacaaggaca  
25261 gtggccatta gccacatgtg gctactgagt ccttgaatt gtgactaaat ggagatgtgc  
25321 tataaatgta aaatacatac tgggtttcag attttgaaga cttgggatga aaaaaaaaaat  
25381 aggtggggcg tgggtggcaca cacttgtaat cccagcactt tgggaggccg aggtgggcag  
25441 atcacttgag ctccaggagt cgagaccaac ctgggcaaca tggcaagacc ccatttctat  
25501 aaaaaataca aaaattagct gggcatgatg gtgcacacct gtagtcccag ctagtgggga  
25561 ggctgagacg ggaggatcac ttgagcctgg gaggtaagg ctgcagtgag tcatgatcgc  
25621 accactgcac tccagtctgg gcgacagagt gagaacttgt ctcaaaaaa aaaaaggaaa  
25681 caatgtaaaa tatctcatta acaattttat actggttaca tgttgaaatg atatgttaat  
25741 atattggggtt aaatagatta tatacattta tttatttatt tttttattta cttttatfff  
25801 tttttttatt ttttttgaga tggggtttcg ctcttgttgc ccaggctgga gtgcaatggc  
25861 acgatctcag ctccaccaca cctccgcctc ccgggttcaa gtgattctcc tacctcagcc  
25921 tcctgagtag ctgagattac aggcattgtc caccagccc agctaatttt gtgtttttag  
25981 tagagacggg gcttctccat gttggtcagg ctggtctcaa acttctgacc tcaggtgatc  
26041 cacctgcctc agcctcccaa agtgcctggga ttaccgggtg gagccactgc gccagccga  
26101 ttatatacat tcaataaatt aaatttaaat atattaaatt tggtcaggca cagtggcagc  
26161 agtctgagag tccaaggcag gaggatggct ctacagaatt aaaataaaaa tattagctag  
26221 aacgcagtga gaccctatct ctacagaatt taagacagga ggaatgcctc agcccaggag  
26281 agtgcctgta atcctagcta ctccaggagt taagacagga ggaatgcctc agcccaggag  
26341 ttcgagactg cactgagcca agatcgcacc actgcactcc agcctgggag acagagcctc  
26401 actctgaata cataataaat aaataaataat aaaatattaa atttaatatc tcccatttct  
26461 gttttatfff taatgtggct actagaaaat ttaaaattgc atatgtggct tgtgttattt  
26521 atctgtttgg cagcaccctc ttttgtttgt ttgtttgttt gtttgttctg ttttgagacg  
26581 gagttatgcc ttgttgccca ggctgcagta caatgacaca atctcagctc accgtaatct  
26641 ccacctccaa ggttcaagtg attctcctgc ctcagcctcc caagtagctg ggattacagg  
26701 catgcgccac catgccctgc taattttgta tttttagtag agatgggggt tctccatggt  
26761 ggccaggctg gtctcaaaact cctgacctca ggtgatccgc ccaccttggc ctcccaaagt  
26821 gctgagatta caggtgtgag ccacctcgc cggcctctgg ggggcctttt atctcatact  
26881 ttgcacatga tattaactcg attgctctca aagacagtgc acttaggtcg ggtgtgggtg  
26941 ctcacgcctg taatcccagc actttgggag gtggagacag gtagattgct tgagcccagg

27001 aattggagac caccgtgggc aacatggcga aactccgtct ctacaaaaaa cacaaaatat  
27061 tagctgggcyg ctgtgggtgtg cgcctatagt cgcagccact ggggaggctg aggtgcttga  
27121 gccagaagc cagaggttgc agtgagcgat gattgcaaca ctgcactcca gcgacagagc  
27181 taagactctg tctcaaaaaa aaaaaaaaaa aaaaaaaaaa aagaggcagc acacagactc  
27241 caaggttcta agcccagctc attccctaac tcgttccatg ggcttgcaca agaagtactc  
27301 caacctcctc gggcctcagt ttccatataa ataaaacggg aaaagtaaga taactgtgaa  
27361 gatgaaatga cagcacagtg ccaggcacac agttcaagct taatgactgt gttccccaga  
27421 gcagcaccce caagagctcc cctgcactca ctggggtcac agtcattgat gtcctgatcy  
27481 caggaagggc cagtgtaccc tccatggcag gtgcagctga aactccctgc caggttggtg  
27541 cagataccat gagggccaca ggggtgcgggg ccagcacact cgtccacatc ctgctggcat  
27601 cgtgggcctg ggggtaggga gcaaggttac acctagggtt acagggtaaca gagcagggtc  
27661 tcagggacct ggggctcagg aagaacctgc aggccagat ggggtgaaggc aaaataagcc  
27721 gtgctgtcat tcgtgtcttg gcattttgca tatgcagttt cagaatggcc gtctccctgt  
27781 ctcagtaggt ggcaaatacc tactatgagg tctgtctctc caggaagctg taccctaacc  
27841 cccgccctaa gtcccagcat ctctagcccc ataactcttc ctcttccaag ggctgacca  
27901 cacaggctgc atatcagtgt ccagctctgt ccttccaaga aggaagcccc tctcaaaggc  
27961 agggctgggg atgggtccct ctctggatcc ccagcatcat ccctgatagg gtagggggca  
28021 gaggagatgg agaggaggag ggaagagaag caggtggcat accttgccag ccctgggggc  
28081 aggagcagac aggcagctgg ccaggggcag actcgcagcy gccccatgc tcacaggggt  
28141 tcggggtgca gggggagagg agttcacact gacgtcctgt tgggggtgga agagagggaa  
28201 gcagagatag ccttgaggga ctccctgatc ccatctcccc ccacctcccc caaccccagg  
28261 gtccccacct tgatacccac ctcccagct cctggaggga aatgattgaa agcaaaaaag  
28321 aagctaaccat agcgggagga gagagtagag gagaagagag atgagaaggc ccatggtgtt  
28381 ggtggggctg cagaggggag gtgaggtaca caccctggac accaggcggg caggtgcagt  
28441 ggaacccat tccatcgtg ctgcatgtcc caccggccct gcacggctgg gactcacagg  
28501 cgtctcgggc caggctctgg ctgcagcggg gccactcca gccaggctca cacacacagc  
28561 ggaacctggc aggggaaggt agtcaggcca gggagggtgg ccagggagag ggggcagtgt  
28621 ctgaggctga gaagggccct caccgcag gtgcatcata gcagatgccg tgactgcagg  
28681 gctcatgggc acagggatgg ctcgggggga ggacagatgg gggcaaggag ccaggcgggc  
28741 agaggcagcy gaagccatct tccccatcca cacaggaacc tcctcgcgc catgggtgg  
28801 aagcacactc attgatctcc acgttacaac ggggccctgg ggagtacaca agcaatctca  
28861 tctcagaaca aagtcagcag ggacaaccag ggaggacga cgtgaccca cttagcacac  
28921 ccacaccccc gagcaatgac ctctttttca taacgcatca gctcttgtgc agattaggac  
28981 acccgcctcc tcaaccaaga gctgacttgc ccagatcat tctgggtccc aaacttccat  
29041 gaagtccctt tctttttctt ttcttcttct tcttattttt tgagatgggg tctcactctg  
29101 tcgcccgggc tggactgcag tggcatgatc atagctcacc acagcctcca ctctctgggc  
29161 tcaagcgatt ctcccacctc agcctcccga gtggctggga ctacagttgc gcaccaccac  
29221 gtctggctaa ttgtgtgtgc atgtgtgtgc gtgtttatct acgtattgac taatctggct  
29281 ttgggagtgt gtttattttt tgtagagaca gggtttcatt atgttgocca ggctagtctc  
29341 gaactcctgg gctcaagtga tctcatgtc tcagcctccc aaagtgtctg gggattacag  
29401 acatgagcca acacacctag caaagtcaca ttcatttcca ctaaggatta gtttagacct  
29461 attcagaaat aacctagtca ctgagtttgg acactttcat ctatgttcca gaataatctt  
29521 tttgtttgtt tttgtatgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtttg  
29581 aatgaagtc tactctgtc gccagactg gactacagtg gcacaatctt ggctcactgc  
29641 cgctccgccc tcctgaggtc aagcaattct cctgcctcag cctcccagat agctaggact  
29701 acaggcgcce accaccacac ccgactaatt tttgtatttt tagtagagat ggggctcgc  
29761 aacgttgccc acgctggtct cgaactcctg acctcaagt atccacttgc ctggcctcc  
29821 cagagcgtg ggattacagg cgtgagccac cgcgccagc ctgcagaata atctcgaaac  
29881 aaccttatgc caatgagaca gcacagactc agggcaaagc acggacaacc tcgttggaaca  
29941 agagtctgca aagatacggg caaaacaggc ccacagagac gagtgaactc accctcgatc  
30001 taaggacccc ctctcatggc agccacttgc ccacctgtga agccagggtg gcagacacag  
30061 tcgtagcggc tgatgccatc acggcagact ccaaaggctg aggggttgc ggcacagtcg  
30121 tcaatgttca cttcgcagtt cacacctagg gccaggaac atggcatgga gtggccacca  
30181 ctgtgcccc aatgatgcac cattccccaa cctctgtgc cctccaggt gtgctgtttc  
30241 tgccccagcc cccgggtccc cctgtggtcc cagaaggca gcggcagag taccgttcca  
30301 ccaggtctag ccatttgccc ccatggggc agggctggct gcggcattcg tccacctggc  
30361 tctcgcagcy tgtgcccgtg tagccaggag cacaggcaca tgagaagctg gcgatccat  
30421 ccacgcagcy accatggtgg catgggtcag gggagcagtc gtccacgttg cgatcacaca  
30481 gcgtgccctc aaagcctgtg gggccaagag ggtcaggctc cggccacttg ccagggcct  
30541 gccacaagt gaggcctcgg accaatcctg gttcagctct atctgaggcc ctgcccata  
30601 agctgtcagg aggcagctc aatacaggcc ccacccccca cccccactt acttccactc  
30661 cccccacta caagctccag aatttgtccc tagccaagac tctcccacag ttaaattccc  
30721 cccccagctg tgccccacc tccagccctg tttctgcccc accctccact ggagatggga  
30781 gcacatggca ggtgccaggg gaaggtattg ggccctcccc attcagttac aggcacctg  
30841 tgcccagccc aacactttgg ctccacacgt agccttatgt cagtctatat gcattattgg  
30901 ctccaccccc caactctgtc actgggtcct gccttcttac aaccccgcc ccaagctctc  
30961 cccaagtctg ttattggccc tgtcggccc aagccccgc tcctgattct tgtcggactg



31021 tcattggccc gcctcaccct cggcacagcg gcactcgtag ccatcgggct ggtccacgca  
31081 tttggcgcca ttcttgcagg gcgtgctggc gcattcgtcc acgtccagct gacacgtgga  
31141 gccgctgaag cctgggggtgg ggagtgggat gagcagaggc ccagaaaggg tgagagcagt  
31201 acaccctct ctagccccgc cttgtttggg tggaaaaacc ctttacttg gatttaaatt  
31261 aacttcaatg ttccataaca tatggtttct attgtaagtt atccgctaac gtggttttac  
31321 aggttcccac cctggagttt ttgccccttc ccagacatgt cttttcgggc tccctctcct  
31381 gagtcctca cccgaggggc aggtgcagct gaagccattg actcggctct tgcagacccc  
31441 accgttgaca caggggctac tctgacactc gtcaatgtcc acctcgcaat aggttcctgt  
31501 gaagcctggg gcaggggaata gggcttagga aagcgggggc taccctatgg tgtgaacggg  
31561 gtgcaaggaa ggaggtactt cccttctctac tcatcgacaa atccccgag ccttcgtccc  
31621 cacctctaca ctcccagccc aaccttaatt ttacgccaca ggctggactg tcaccagcc  
31681 gtggtcccaa ctggcccaag gctcatgggt gtccgcagag gccttgcctt caaacccctc  
31741 ctctctctcc cagtctcag actaagccct ttccccagtc cccagcccc agccccaggg  
31801 cccccatag cttttgccc ggttactgc aagcttgtct gcaggcttca ccttgacctc  
31861 acagcagccc cgactctgcc tctctatccc caacccttc ccgagtgggt cctaaggcca  
31921 agaatactcc cagcccacct gaggttctac tgactccatt tgaccaagcc cccagaatca  
31981 cacagtattc ctttctggtc ctgacccac ctcttccag gcttcagtct ctaagggtcc  
32041 cactccaaac ccacttacac ccattctgc tcagctcccc atcccgggc ttctctgtcc  
32101 ccgccccctg cctcaggacc cgcccaggcc acgcccacca cccacctgcc atacagatac  
32161 aggtgaactg gcctatgagg tccagggcac tggcctggtt tccgagggc cccgacagac  
32221 actcgttgac atcgggtctca cagcagggtc cagtgtagcc acgaccgcac tggcacagga  
32281 aggagccctg cgtgttcacg cacctgccc aagtgtcgc ggggttggcg cctgcccgat  
32341 ggagtgcgat cgggtgtggc gtggctggcc gggacccct cctctcccct ctttcccagc  
32401 ccattcacag acgatggagc tcccctcacc gatagagcac tegtccacat cctggtcaca  
32461 tgccccacc gtgaagccgg gaggacaggt gcaaatggcc cggccgttca ccggatttgt  
32521 gtcacagata gcatcctcgt ggcaggggt gctgacacag gcgtcatcca ggtgacacag  
32581 gaggcctggg aagtggtaag cagaagtcac aggcagatct tctgtctctg cccaaaaggc  
32641 ccacaccct tgagtattgt ttgtctgttg ttgttgttt gttttgttt ttgagacgaa  
32701 gtttcgctct tgttgcccaa gctggagtac aatagcgcga tcttggctca ctgtaacctc  
32761 cacctcccag gttcaagcga ttctctgtc cagcctccc agtagctggg attacaggca  
32821 tgcgccacca cgcccggcta atttttgta ttttagtag aaacggggt tccaaagt  
32881 agccaggtg gtctcaact cctgacctca ggtgatctgc ctgcctcagc ctcccaaagt  
32941 gctgggatta cagcggtgag ccaccgcgcc cggcctggtg tttttattaa acctaaacag  
33001 tatttagtag ctgctttttt aatccattta ttatgagaac tatcaaacat acagaaaagt  
33061 tgaccattca tatacccacc acctagattc aactattggt agcttaatct accaatctc  
33121 ttttctaaac catttaaaaa taaattgcag accgggcaca gtggctcagc tctataatcc  
33181 tagcactttg ggaggccaag gtgagtgcac cgcttgagtt caagagttca agaccagcct  
33241 gggcaatata gtgagaccgc tctctacaaa aaatacaag cttagccggg catggtggcg  
33301 cacgcctgta atcccagcta cttgggggac tgaggtggga ggatcacttg agcccaggaa  
33361 gtcgaggetg cagtgagctg agatcatgcc actgtactcc agcctgggtg acaaagtgag  
33421 actctgtctc aaaaataaaa taaaataaat tgcaaacatc cacatattgt tttaaaacct  
33481 tttgactct ctgattacaa aaggtccta cactccagcc tttgctctg taataacctc  
33541 tgccatgat gccctctttc tatatccacg ttcccctct aaacatgttc tctgatctt  
33601 gccacttcca aaggttaact ccaggagttc ccattccaa tggcttctgc cctgagccat  
33661 gcatcagtg aatgtcagcc tctcccag cctctcccag gcttacact cccactaaa  
33721 gctgcctcca gcatcatctt ttttctctc tttttttgt tatacagggt cttacactgt  
33781 ctcccaggct ggaacgcagt ggctcagctc cagctcactg caacctccac cttctgggct  
33841 caagccatcc ttcaacctta gcatccagag tagctgggac tacaggcatg cgctaccatg  
33901 cccagctaat tttttgtatt tttggtagag atagggtttt gccatgttgc ccaggctggt  
33961 ctcaaactgc tgagctcaag caatctacc gcctgggct cccaatgtgt tggggttaca  
34021 ggctgagcc accacacca gcctaccaag tttttgtaga gatgacaaa actcactatg  
34081 ttgtccaggc tggctctaaa gtccctgggc caagcagatc tccatcttg gcctcccaa  
34141 gtgctgggat tacaggcatg agccactgtg cccggcctag cataatctt tttttatctt  
34201 ttaaatttga gaaggagtct cactctgtca cccaggctag agtgcagtg tgcaatcttg  
34261 gctcactgca atctccacct cctgggttca agctactctc gtgccttagc ctctgagag  
34321 acttgatta caggtgtgca ccaccagcg cagctaatt ttgtgtttt agtagagatg  
34381 gggtttacc atattgacca ggctggctc aaactcctga cccagtgat ccgcccgcct  
34441 gggcctccca aagtgtatg attacaggca tgagccactg tgcccagcct agcataatct  
34501 ttctaact caagtcagac ttcttatttg ccctcactaa aaacctoca tggctcctg  
34561 cagagaaaac ggccactcac cagtcttgcc catggggcag gcacagtaga aagaagccac  
34621 gcggtcatgg caggtggccc catggaagca cacggctgtg gcacagtcac cgatattctg  
34681 actgcagctc tcgcctgtcc agccattgac acacacgcag ctgtggccac ccagcgtgtt  
34741 gaagcaggta cccccattgt ggcaggcgtt gggctgcagc tgacactcat ccacgtctc  
34801 cgtgcagaac tggcctgtgg cacacagatg cagcagtcga gccacctggc gcagtccac  
34861 ccgaggcctg cctccccgct cctctggcc gcagtccca cctgtccact caggagggca  
34921 ctggcagtta taggtgttga cgccatccac gcatgtcccc ccattgagac atcgggtctc  
34981 tggacagtcg tccacgttca cttcacaatt ctgacctca aacctagca gggaggggg

35041 caaggatggt caccgcccggg ctggcctgct gtccccacgc ccacccctct gactctcctg  
35101 agtagggctc actcaccagg aagacaggca cagtcgtaag tgaggtcgcc actctgcctg  
35161 caggtgcccc cgttacggca tggtaggggt gcacagggca ccgccccggt ctcacatagt  
35221 ggccctgtgt agccagctgg acactggcag cgaaggagc caggtgtgtt gaggcaggtg  
35281 ccacatggc ggcagggctc acccaccggg cactcatcca cgtcgcttcg gcagctgcgg  
35341 ccctggtagc caggtgggca ggcagcaggg aagcgtccat cgggccccac tgagcagcgg  
35401 gcaccgtggg cacaagggct gctgagggc ggatctggca gggagcagtc ggggcttggg  
35461 gggaccagga caggggtgagt ttaggactga ccacaccccc gactacctcc cctccagact  
35521 cttcccctct cacctcggaa gccacggggg caccggcatg agaatcgggc ggtgccagcc  
35581 accactgaac tctggcagac accacggcca gcacaggggc ctgagtgaca ggggtcctcc  
35641 agctgacacc gctcaccacc ccagccaggc gggcacctgt gggcagagat ggcttggtt  
35701 ggcagcacag ggcaggatgg cccagacac aaagatacac acaagacagg caagaaacac  
35761 gcagggaaac agagcagcaa acacacattt gctgggagca gtggctcagc cctgtaatcc  
35821 cagaactttg ggaggctgcg gcgggcagat cacttggggg caggagtcca agaccagcct  
35881 agccaacatg gtgtaacccc gtctctacta aaagtacgaa actttgctgg gcatagtggg  
35941 gcgtgcctgt agtcccagct actcaagagg ctgaggcagg agaattgctt gaactcggga  
36001 ggtggagggt gcagtgagcc aagatcgcgc cactgcactg cagcctgggt gacagagtaa  
36061 gactcggctc caaagaaaca aaacaaaaca aaacacacat tcattcatgc aacgaatatt  
36121 tactgagtac ctactacggt ctgagcaata caaatacagc aacaaacaca atgaaagggg  
36181 tccctatccc tctggatcct tgcagccagg tccaggcaag gagaacatat acgaataaca  
36241 aattaatata caacatgctg gcagggggcg gtgaaagcca tcagaacat gaaggatgta  
36301 gtcagttggg tcacagcccc tgaaaagatg acatccatgt cccagaacct gtgactagga  
36361 ccttactttt tttttaaatg gctctttgca gatgtaattt tttttttta agagacggag  
36421 tcttgctctg tcaccaggc tgcagtgca tggcgcgatc ttggctcact gcaacctcca  
36481 cctccccggg ttcaagcaat tctttgtctc agcctcctgc gttagctggga gtacaggcac  
36541 acgccaccag gcctggctat gcagatgtaa ttaaatgaga gatgtcaaga tgagattatt  
36601 ctgaattatc cacctgggcc ctaaatcag tgacaagtgt tcatacaaga gaacggctca  
36661 caactgtaat cccaacactt tgggaggtca agacaggata atcacttcaa gccagaagtt  
36721 caagatcagc ttgagcaaca tagggagact tttttccac aaaaaaaaaa taaaaattat  
36781 aaccgggatc ggtggcagga gcctgtaatc ccagctactc tggaggctga ggcaggagaa  
36841 ctgcttgaac ccaggaggcg gaggttcag tgagccgaaa ttgcaccact gcactccagc  
36901 tctggtgaca gagggagact ccatctcaaa aaaaaacaaa aacaaaaaaa aacaaacagg  
36961 ccaggcagc tggctcatgc ctgtaatccc agcactctgg gagactgagg tgggcgcatc  
37021 acgaggtcag gagttcaaga ccagcctggc caatatggtg aaacctcgtc gctactaaaa  
37081 atacaaaaat tagctgggcg tgggtggtgca ctctgttagt ccccgtact caggaggctg  
37141 agacagaaga attgcttgaa cctgggaggc ggaggttgca gtaagccgag atcgtgccac  
37201 tgcactccag cctgggtgac agagagagac tccacctaaa aaataaaaaat aaaagattag  
37261 ctggccgtgg tggctcacac ctgtagtccc agcttcttgg gaggtgagg tgggaggatc  
37321 gcttcagcct gggagttgaa ggctgcagtg aactaagatt gcaccactgc actccagcct  
37381 gggtagacaga gcaagactct gtctcaaaaa aagagagagg cagccaaaag gagatttggg  
37441 tgggagaagg aagagaagaa gacagagaca gggactggag tgagacagcc actagctgag  
37501 aaacgtctgg agccaccaga agctgcaaga gtcagttctc ccctgggtcc ccagatgagag  
37561 gctcagccct gccgatatct tgatttcaga ctccaacct ccagaacttc aagagaataa  
37621 attactattg ttcaagctac ccagtttatg gtaatttggt agagcagccg taggaagcta  
37681 aaccacaggg ttaaggggga tagaatgatg ggtcaagga ggtggaattc tattttcaca  
37741 gacatatgaa ggtaaggagg gggccatggg gctttctggg gaacggcatt ccaggcagag  
37801 ggaacagcaa atgaaaagtc tcagacccat gtccaagcgc ggtggctcac gcctgtaatc  
37861 ccaacacttt gggagaccga ggcaggtaga tcacctgaca tcaagagttc gagaccagcc  
37921 tgaccaacat gaagaaactc cgtctctact aaaaacacaa aaaatcagcc aagtgtggtg  
37981 gcacatgcct gtaatcccag ctactccgga ggctgaggca ggaatcgc ttgaaccag  
38041 gagatagagg ctgtagtgag ccaagatcga gccactgcac tccagcctgg gcaacaagag  
38101 cgatactcga tctcaaaaaa aaaaaaaaaa aggaaaagga aaaaaagaag tctcagacc  
38161 caggacacaa ccacgcaggc acacacagac aggccagaa accacagcta cacacacaga  
38221 gggacacacc aattggacgc catggcctgg ccagactca agacaccaac ccacagcatc  
38281 ccacgcggaa atgagcaact ctcatcaca gggacagaca gcagggtcc cacaaaacct  
38341 caagtccctt ctctgggact tatatagcct cttatgcagg gatgggaggg tggagggtt  
38401 ggtactgac accactgttg gtaccctcat cttggggggg tggctctca acctgcagcc  
38461 ccctcccaa gccaccaag gagtgtgagg aattggaact gggaggggtt ggtcttaaga  
38521 ctggaaatct ccaaggggca ctagcggggg cagagggagg ggcttagaga agccagagag  
38581 tccaagctcc gcccccgac gctgccaacc ctaccccc tctatcccc caccctccgc  
38641 ccctggcgcg gttgtcctcc ctccggcttg tcccagacgc ctgagactga actcaacaag  
38701 cggctgactc agtccttagc gtcaccgcga agatccccc caccocggcc cccagcgaaa  
38761 caggtcgggg cacggccgag caagcaaggg tgcgggtggc gttgggaggg gagggcacag  
38821 ctcccgtccc ctctccaag gcccgaggg cgaccactgc ggccccgacc gcgctcggg  
38881 agggggcgcg ctgcgctggg cgccgaggat agctggtccc accgccaggc agggcgcgac  
38941 ccgcgcgggc cccctcccga cgccccggg gggttcccac gctctgcgtc ccgcgcccc  
39001 gccgaccggt cgggcccgtt ccagccccag cccgcgcccc gccccacggc ctaagctttg

39061 ggggagacgg aggggggagc gcagacctag ttaggggggg cacggcgggt gaggcgagga  
39121 ctaggagatg gggacgaggt ctatggaaac gtctgctgcc ggtggggggg tcttggggga  
39181 tccatgggtg gggtgacagc ttgggggcat ccttgaaagg acttgagtat ttgagaccgt  
39241 gaaagagact cgtgggggggt ggagggcgtg tagggagtgg tccccacagc ccaactcggg  
39301 ggactgagcc cgctccactc agggcgatcc gggctccggg cgccgcgcgc gccggtgatt  
39361 cacctgtcga agggcggggc agccgcgggc tcattcacct gtcggcagta ttccaggcgc  
39421 ggagctccac tgcgcaggcg cccggcccca ccgctcgccc tcccagggtg gtggctccag  
39481 gtaaggtcac acttttcccc atcccccccc accttcgcag accccccagc cccaggaaac  
39541 cctctcagtc aggctcagcc ctccacaaaag gtgagcccga gtgggagatc cagccaaaat  
39601 aaatgtaggt gtaacttcaa agcggctctt cttggtctaa aatcacccca gaccaccca  
39661 gcttgggctt gaccagtac ttggcacaca tgatgctgaa taaatgaata cattaatgag  
39721 tcaatcaatg gctactgctg gaactgctgt tattaatgat agcatgaaag gagccaagca  
39781 tttcctcata accttgagag acttctcctt agagcacctg tttcctgcct ccctaccatt  
39841 attatttata ataatgccta tgttgagcac ataactatatt ctgagcactt aatgtatggt  
39901 gagtgcatac aggagtattt attgccccat tttccagatg caaaaactga ggctcaaaaa  
39961 gcttgaatgc tcaaggtgtc ccagacccca gaatttaagc tgagctcttt ctgtctctgc  
40021 tcagctcctt cccagcctct gctctgctcc ctgcaacctc ttgcctttcc ttgtgatatc  
40081 taacaactaa attgttaagt cccttcacag tcgccagtga agcattccct atttcacagg  
40141 tgcaatgatg atcagagagg gaaagtgact agcccaaagt cacactgcaa gtaagtggca  
40201 gtgtgggact gtattatagg cctcctatgt cccctccac taaatggtga gagatgggag  
40261 ggggtgtggag ggtgtgagcc actgcgattg gtccctgata gtttgttgat tgaataatac  
40321 ctacctcatg gggggcttta catgtgacct caggtggcca agaactgtga agctgggatt  
40381 actgggcccc ttttgacagc catgacactg agatgtaaag agtttaattc ctagccgggtc  
40441 agcaacatga tgtgggactt gaacccatgc caagaagagt gcagagcccc tgctcagttt  
40501 ccaatgggga aacacgagag gttgcccaag ccacacacat gtaggaagtg gtagagacat  
40561 ccatggtgaa cacagaggca gagggagaag acaaatcgcc cctccccccc gcccccacac  
40621 acagggcccc ctggtggctc tgagccaggc actcacaggc aggcagcctc ccgggagggc  
40681 agctgggtgc aacgacctcc atttgcacac gggcttccgt ccaggcaagg ggggctgtg  
40741 tgggggtgaa ggaaggtgga ggatcagcca ggtgccagg aaccccaggc ccaagtgaca  
40801 aaccccctcc ctcccctccac caggcaacag ctgcatgggg atgggggctg ctctgcgggt  
40861 gcagggagtc cccccatcca cagttccac gcccccccc cccccgccc agccccagct  
40921 gttgaggctg cagctgtccc tgccagctcg agccctggga accacaaagc aggggaaggc  
40981 aggaggagg gctgtcttat gccagatgtg gggctcggg aagctgccag agctaagatg  
41041 tgcagaatgg ggtgtgctga gacaggaagt aagggtgctg agtgttacag ggaggtctgg  
41101 ggcttcatgc ccaggttcag gggtaataga aggggaagg gcaggggatc

//

[Disclaimer](#) | [Write to the Help Desk](#)  
[NCBI](#) | [NLM](#) | [NIH](#)

Controlos (N=105)	Variações de sequências no exão 11		
	C568Y	T575T	G594G
C1	-	-	-
C2	-	-	-
C3	-	-	-
C4	-	-	-
C5	-	-	-
C6	-	-	-
C7	-	-	-
C8	-	-	-
C9	-	+	-
C10	-	-	-
C11	-	-	-
C12	-	-	-
C13	-	-	-
C14	-	-	-
C15	-	-	-
C17	-	-	-
C19	-	-	-
C20	-	-	-
C21	-	-	-
C22	-	-	-
C23	-	-	-
C24	-	-	-
C25	-	-	-
C26	-	-	-
C27	-	-	-
C28	-	-	-
C29	-	-	-
C30	-	-	-
C31	-	-	-
C32	-	-	-
C33	-	-	-
C34	-	-	-
C36	-	-	-
C37	-	-	-
C38	-	-	-
C40	-	-	-
C41	-	-	-
C42	-	-	-
C43	-	-	-
C44	-	-	-
C45	-	-	-
C47	-	-	-
C48	-	-	-
C49	-	-	-
C50	-	-	-
C51	-	-	-
C52	-	-	-
C54	-	-	-
C55	-	+	-
C56	-	+	-
C57	-	-	-
C59	-	-	-
C60	-	-	-
C61	-	-	-

Controlos	Variações de sequência encontradas no exão 11		
	C568Y	T575T	G594G
C63	-	-	-
C66	-	-	-
C68	-	-	-
C69	-	+	-
C70	-	-	-
C71	-	-	-
C72	-	-	-
C73	-	-	-
C74	-	-	-
C75	-	-	-
C76	-	-	-
C77	-	-	-
C78	-	-	-
C79	-	-	-
C80	-	-	-
C81	-	-	-
C82	-	-	-
C83	-	-	-
C84	-	-	-
C85	-	-	-
C86	-	-	-
C87	-	-	-
C90	-	-	-
C91	-	-	-
C92	-	-	-
C93	-	-	-
C94	-	-	-
C95	-	-	-
C96	-	-	-
C98	-	-	-
C99	-	-	-
C100	-	-	-
C101	-	-	-
C102	-	-	-
C103	-	-	-
C104	-	-	-
C105	-	-	-
C106	-	-	-
C107	-	-	-
C109	-	-	-
C110	-	-	-
C111	-	-	-
C113	-	-	-
C114	-	-	-
C116	-	-	-
C117	-	-	-
C118	-	-	-
C119	-	-	-
C120	-	-	-
C121	-	-	-
C122	-	-	-

Controlos (N=102)	Variações de sequências no intrão 18 IVS18+73A>G
C1	-
C2	-
C3	-
C4	-
C5	-
C6	-
C7	-
C8	-
C9	-
C10	-
C11	-
C12	-
C13	-
C15	+
C17	-
C19	-
C20	-
C21	-
C23	-
C24	-
C25	-
C26	-
C27	-
C28	-
C29	-
C30	-
C31	-
C32	-
C33	-
C34	-
C35	-
C37	-
C38	-
C39	-
C40	-
C41	-
C42	-
C43	-
C44	-
C46	-
C47	-
C49	-
C53	-
C54	-
C55	-
C56	-
C57	-
C58	-
C59	-
C60	-
C61	-
C63	-
C64	-
C65	-

Controlos	Variação de sequência no intrão 18 IVS18+73A>G
C66	+
C69	-
C70	-
C71	-
C72	-
C73	-
C74	-
C75	-
C76	-
C77	-
C78	-
C79	-
C80	+
C81	-
C82	-
C83	-
C84	-
C85	-
C87	-
C88	-
C89	-
C90	-
C91	-
C92	-
C93	-
C94	-
C95	-
C96	-
C97	-
C98	-
C99	-
C100	-
C101	-
C102	-
C103	-
C105	-
C106	+
C110	-
C111	-
C113	-
C114	-
C115	-
C116	+
C117	-
C118	-
C119	-
C120	-
C121	-

Controlos (N=102)	Variações de sequências no exão 19	
	A1020P	W1028C
C1	-	-
C2	-	-
C3	-	-
C4	-	-
C5	-	-
C6	-	-
C7	-	-
C8	-	-
C9	-	-
C10	-	-
C11	-	-
C12	-	-
C13	-	-
C15	+	-
C17	-	-
C19	-	-
C20	-	-
C21	-	-
C23	-	-
C24	-	-
C25	-	-
C26	-	-
C27	-	-
C28	-	-
C29	-	-
C30	-	-
C31	-	-
C32	-	-
C33	-	-
C34	-	-
C35	-	-
C37	-	-
C38	-	-
C39	-	-
C40	-	-
C41	-	-
C42	-	-
C43	-	-
C44	-	-
C46	-	-
C47	-	-
C49	-	-
C53	-	-
C54	-	-
C55	-	-
C56	-	-
C57	-	-
C58	-	-
C59	-	-
C60	-	-
C61	-	-
C63	-	-
C64	-	-
C65	-	-

Controlos	Variações de sequências no exão 19	
	A1020P	W1028C
C66	+	-
C69	-	-
C70	-	-
C71	-	-
C72	-	-
C73	-	-
C74	-	-
C75	-	-
C76	-	-
C77	-	-
C78	-	-
C79	-	-
C80	+	-
C81	-	-
C82	-	-
C83	-	-
C84	-	-
C85	-	-
C87	-	-
C88	-	-
C89	-	-
C90	-	-
C91	-	-
C92	-	-
C93	-	-
C94	-	-
C95	-	-
C96	-	-
C97	-	-
C98	-	-
C99	-	-
C100	-	-
C101	-	-
C102	-	-
C103	-	-
C105	-	-
C106	+	-
C110	-	-
C111	-	-
C113	-	-
C114	-	-
C115	-	-
C116	+	-
C117	-	-
C118	-	-
C119	-	-
C120	-	-
C121	-	-

Controlos (N=102)	Variações de sequências no exão 20 C1099Y
C1	-
C2	? -> Sequenciada
C3	-
C4	-
C5	-
C6	-
C7	-
C8	-
C9	-
C10	-
C11	-
C12	? -> Sequenciada
C13	-
C14	-
C15	-
C17	-
C19	-
C20	-
C21	-
C22	-
C23	-
C24	-
C25	-
C26	-
C27	? -> Sequenciada
C28	-
C29	-
C30	-
C31	-
C32	-
C33	-
C34	-
C35	-
C36	-
C37	-
C38	-
C39	-
C40	-
C41	-
C42	-
C43	-
C44	-
C45	-
C46	-
C47	-
C48	-
C49	-
C50	-
C51	-
C63	-
C64	-
C65	-
C66	Não digerido -> Sequenciada
C67	-

Controlos	Varição de sequência no exão 20 C1099Y
C68	-
C70	-
C71	-
C72	-
C74	Não digerido -> Sequenciada
C75	-
C76	-
C77	-
C78	-
C80	-
C81	-
C82	-
C83	-
C84	-
C86	-
C87	-
C88	-
C89	-
C90	-
C91	-
C92	-
C93	-
C94	-
C95	-
C96	-
C97	-
C98	-
C99	-
C100	-
C102	-
C103	-
C104	-
C105	-
C106	-
C107	-
C108	-
C109	-
C110	-
C111	-
C113	-
C114	-
C115	-
C116	-
C117	-
C118	-
C119	-
C120	-
C121	-
C122	-

## **7. Referências bibliográficas**

---



1. Sourander, P. and J. Walinder, *Hereditary multi-infarct dementia*. Lancet, 1977. **1**(8019): p. 1015.
2. Stevens, D.L., R.H. Hewlett, and B. Brownell, *Chronic familial vascular encephalopathy*. Lancet, 1977. **1**(8026): p. 1364-1365.
3. Davous, P. and C. Fallet-Bianco, [*Familial subcortical dementia with arteriopathic leukoencephalopathy. A clinico-pathological case*]. Rev Neurol (Paris), 1991. **147**(5): p. 376-384.
4. Sourander, P. and J. Walinder, *Hereditary multi-infarct dementia. Morphological and clinical studies of a new disease*. Acta Neuropathol (Berl), 1977. **39**(3): p. 247-254.
5. Sonninen, V. and M.L. Savontaus, *Hereditary multi-infarct dementia*. Eur Neurol, 1987. **27**(4): p. 209-215.
6. Tournier-Lasserre, E., et al., *Autosomal dominant syndrome with strokelike episodes and leukoencephalopathy*. Stroke, 1991. **22**(10): p. 1297-1302.
7. Joutel, A., et al., *A gene for familial hemiplegic migraine maps to chromosome 19*. Nat Genet, 1993. **5**(1): p. 40-45.
8. Joutel, A., et al., *Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia*. Nature, 1996. **383**(6602): p. 707-710.
9. Opherk, C., et al., *Long-term prognosis and causes of death in CADASIL: a retrospective study in 411 patients*. Brain, 2004. **127**(Pt 11): p. 2533-2539.
10. Razvi, S.S., et al., *The prevalence of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) in the west of Scotland*. J Neurol Neurosurg Psychiatry, 2005. **76**(5): p. 739-741.
11. Chabriat, H., et al., *Patterns of MRI lesions in CADASIL*. Neurology, 1998. **51**(2): p. 452-457.
12. Davous, P., *CADASIL: a review with proposed diagnostic criteria*. Eur J Neurol, 1998. **5**(3): p. 219-233.
13. Auer, D.P., et al., *Differential lesion patterns in CADASIL and sporadic subcortical arteriosclerotic encephalopathy: MR imaging study with statistical parametric group comparison*. Radiology, 2001. **218**(2): p. 443-451.
14. O'Sullivan, M., et al., *MRI hyperintensities of the temporal lobe and external capsule in patients with CADASIL*. Neurology, 2001. **56**(5): p. 628-634.

15. Joutel, A., et al., *De novo mutation in the Notch3 gene causing CADASIL*. Ann Neurol, 2000. **47**(3): p. 388-391.
16. Coto, E., et al., *A new de novo Notch3 mutation causing CADASIL*. Eur J Neurol, 2006. **13**(6): p. 628-631.
17. Chabriat, H., et al., *Clinical spectrum of CADASIL: a study of 7 families. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy*. Lancet, 1995. **346**(8980): p. 934-939.
18. Dichgans, M., et al., *The phenotypic spectrum of CADASIL: clinical findings in 102 cases*. Ann Neurol, 1998. **44**(5): p. 731-739.
19. Kalimo, H., et al., *CADASIL: a common form of hereditary arteriopathy causing brain infarcts and dementia*. Brain Pathol, 2002. **12**(3): p. 371-384.
20. Oberstein, S.A.L., Elles MJ Boon, and M. Dichgans, *CADASIL [Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy]*. 2006, GeneReviews at GeneTests: Medical Genetics Information Resource (database online). Copyright, University of Washington, Seattle. 1997-2007.
21. O'Sullivan, M., et al., *Diffusion tensor imaging of thalamus correlates with cognition in CADASIL without dementia*. Neurology, 2004. **62**(5): p. 702-707.
22. Vikelis, M., M. Xifaras, and D.D. Mitsikostas, *CADASIL: a short review of the literature and a description of the first family from Greece*. Funct Neurol, 2006. **21**(2): p. 77-82.
23. Peters, N., et al., *Effects of short term atorvastatin treatment on cerebral hemodynamics in CADASIL*. J Neurol Sci, 2007.
24. Buffon, F., et al., *Cognitive profile in CADASIL*. J Neurol Neurosurg Psychiatry, 2006. **77**(2): p. 175-180.
25. Taillia, H., et al., *Cognitive alterations in non-demented CADASIL patients*. Cerebrovasc Dis, 1998. **8**(2): p. 97-101.
26. Caeiro, L. and J.M. Ferro, *Cognitive profile in CADASIL patients*. J Neurol Neurosurg Psychiatry, 2006. **77**(2): p. 144-145.
27. Viswanathan, A., et al., *Lacunar lesions are independently associated with disability and cognitive impairment in CADASIL*. Neurology, 2007. **69**(2): p. 172-179.
28. Liem, M.K., et al., *Lacunar infarcts are the main correlate with cognitive dysfunction in CADASIL*. Stroke, 2007. **38**(3): p. 923-928.

29. Leyhe, T., et al., *CADASIL: underdiagnosed in psychiatric patients?* Acta Psychiatr Scand, 2005. **111**(5): p. 392-6; discussion 396-397.
30. Nakamura, T., et al., *CADASIL with NOTCH3 S180C presenting anticipation of onset age and hallucinations.* J Neurol Sci, 2005. **238**(1-2): p. 87-91.
31. Lagas, P.A. and V. Juvonen, *Schizophrenia in a patient with cerebral autosomally dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL disease).* Nord J Psychiatry, 2001. **55**(1): p. 41-42.
32. Desmond, D.W., et al., *The natural history of CADASIL: a pooled analysis of previously published cases.* Stroke, 1999. **30**(6): p. 1230-1233.
33. Chabriat, H., et al., *Autosomal dominant migraine with MRI white-matter abnormalities mapping to the CADASIL locus.* Neurology, 1995. **45**(6): p. 1086-1091.
34. Hutchinson, M., et al., *Familial hemiplegic migraine and autosomal dominant arteriopathy with leucoencephalopathy (CADASIL).* Ann Neurol, 1995. **38**(5): p. 817-824.
35. Feuerhake, F., et al., *Reversible coma with raised intracranial pressure: an unusual clinical manifestation of CADASIL.* Acta Neuropathol (Berl), 2002. **103**(2): p. 188-192.
36. Schon, F., et al., *"CADASIL coma": an underdiagnosed acute encephalopathy.* J Neurol Neurosurg Psychiatry, 2003. **74**(2): p. 249-252.
37. Lesnik Oberstein, S.A., et al., *Myocardial infarction in cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL).* Medicine (Baltimore), 2003. **82**(4): p. 251-256.
38. Cumurciuc, R., et al., *Electrocardiogram in cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy patients without any clinical evidence of coronary artery disease: a case-control study.* Stroke, 2006. **37**(4): p. 1100-1102.
39. Schroder, J.M., et al., *Peripheral nerve and skeletal muscle involvement in CADASIL.* Acta Neuropathol (Berl), 2005. **110**(6): p. 587-599.
40. Sicurelli, F., et al., *Peripheral neuropathy in CADASIL.* J Neurol, 2005. **252**(10): p. 1206-1209.

41. Roine, S., et al., *Ophthalmologic findings in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy: a cross-sectional study*. *Ophthalmology*, 2006. **113**(8): p. 1411-1417.
42. Roine, S., et al., *Neurologic symptoms are common during gestation and puerperium in CADASIL*. *Neurology*, 2005. **64**(8): p. 1441-1443.
43. Dichgans, M., et al., *Quantitative MRI in CADASIL: correlation with disability and cognitive performance*. *Neurology*, 1999. **52**(7): p. 1361-1367.
44. van den Boom, R., et al., *Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy: MR imaging findings at different ages--3rd-6th decades*. *Radiology*, 2003. **229**(3): p. 683-690.
45. Singhal, S., P. Rich, and H.S. Markus, *The spatial distribution of MR imaging abnormalities in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy and their relationship to age and clinical features*. *AJNR Am J Neuroradiol*, 2005. **26**(10): p. 2481-2487.
46. Coulthard, A., et al., *Distribution of cranial MRI abnormalities in patients with symptomatic and subclinical CADASIL*. *Br J Radiol*, 2000. **73**(867): p. 256-265.
47. Dichgans, M., et al., *Cerebral microbleeds in CADASIL: a gradient-echo magnetic resonance imaging and autopsy study*. *Stroke*, 2002. **33**(1): p. 67-71.
48. Lesnik Oberstein, S.A., et al., *Cerebral microbleeds in CADASIL*. *Neurology*, 2001. **57**(6): p. 1066-1070.
49. Viswanathan, A., et al., *Blood pressure and haemoglobin A1c are associated with microhaemorrhage in CADASIL: a two-centre cohort study*. *Brain*, 2006. **129**(Pt 9): p. 2375-2383.
50. Peters, N., et al., *Brain volume changes in CADASIL: a serial MRI study in pure subcortical ischemic vascular disease*. *Neurology*, 2006. **66**(10): p. 1517-1522.
51. Jouvent, E., et al., *Brain atrophy is related to lacunar lesions and tissue microstructural changes in CADASIL*. *Stroke*, 2007. **38**(6): p. 1786-1790.
52. Ruchoux, M.M., et al., *Systemic vascular smooth muscle cell impairment in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy*. *Acta Neuropathol (Berl)*, 1995. **89**(6): p. 500-512.
53. Malandrini, A., et al., *Diagnostic value of ultrastructural skin biopsy studies in CADASIL*. *Neurology*, 2007. **68**(17): p. 1430-1432.

54. Ishiko, A., et al., *Notch3 ectodomain is a major component of granular osmiophilic material (GOM) in CADASIL*. *Acta Neuropathol (Berl)*, 2006. **112(3)**: p. 333-339.
55. Markus, H.S., et al., *Diagnostic strategies in CADASIL*. *Neurology*, 2002. **59(8)**: p. 1134-1138.
56. Lesnik Oberstein, S.A., et al., *Evaluation of diagnostic NOTCH3 immunostaining in CADASIL*. *Acta Neuropathol (Berl)*, 2003. **106(2)**: p. 107-111.
57. Ruchoux, M.M. and C.A. Maurage, *CADASIL: Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy*. *J Neuropathol Exp Neurol*, 1997. **56(9)**: p. 947-964.
58. Joutel, A., et al., *Notch3 mutations in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), a mendelian condition causing stroke and vascular dementia*. *Ann N Y Acad Sci*, 1997. **826**: p. 213-217.
59. Louvi, A., J.F. Arboleda-Velasquez, and S. Artavanis-Tsakonas, *CADASIL: a critical look at a Notch disease*. *Dev Neurosci*, 2006. **28(1-2)**: p. 5-12.
60. Hassan, A. and H.S. Markus, *Genetics and ischaemic stroke*. *Brain*, 2000. **123 ( Pt 9)**: p. 1784-1812.
61. Dichgans, M., *Genetics of ischaemic stroke*. *Lancet Neurol*, 2007. **6(2)**: p. 149-161.
62. Meschia, J.F., T.G. Brott, and R.D. Brown, Jr., *Genetics of cerebrovascular disorders*. *Mayo Clin Proc*, 2005. **80(1)**: p. 122-132.
63. Yanagawa, S., et al., *Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy*. *Neurology*, 2002. **58(5)**: p. 817-820.
64. Razvi, S.S. and I. Bone, *Single gene disorders causing ischaemic stroke*. *J Neurol*, 2006. **253(6)**: p. 685-700.
65. Tournier-Lasserre, E., et al., *Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12*. *Nat Genet*, 1993. **3(3)**: p. 256-259.
66. Artavanis-Tsakonas, S., M.D. Rand, and R.J. Lake, *Notch signaling: cell fate control and signal integration in development*. *Science*, 1999. **284(5415)**: p. 770-776.

67. Lai, E.C., *Notch signaling: control of cell communication and cell fate*. Development, 2004. **131**(5): p. 965-973.
68. Weng, A.P., et al., *Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia*. Science, 2004. **306**(5694): p. 269-271.
69. Bolos, V., J. Grego-Bessa, and J.L. de la Pompa, *Notch signaling in development and cancer*. Endocr Rev, 2007. **28**(3): p. 339-363.
70. Oda, T., et al., *Mutations in the human Jagged1 gene are responsible for Alagille syndrome*. Nat Genet, 1997. **16**(3): p. 235-242.
71. Bulman, M.P., et al., *Mutations in the human delta homologue, DLL3, cause axial skeletal defects in spondylocostal dysostosis*. Nat Genet, 2000. **24**(4): p. 438-441.
72. Li, L., et al., *Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1*. Nat Genet, 1997. **16**(3): p. 243-251.
73. Haines, N. and K.D. Irvine, *Glycosylation regulates Notch signalling*. Nat Rev Mol Cell Biol, 2003. **4**(10): p. 786-797.
74. Karlstrom, H., et al., *A CADASIL-mutated Notch 3 receptor exhibits impaired intracellular trafficking and maturation but normal ligand-induced signaling*. Proc Natl Acad Sci U S A, 2002. **99**(26): p. 17119-17124.
75. Federico, A., S. Bianchi, and M.T. Dotti, *The spectrum of mutations for CADASIL diagnosis*. Neurol Sci, 2005. **26**(2): p. 117-124.
76. Joutel, A., et al., *Strong clustering and stereotyped nature of Notch3 mutations in CADASIL patients*. Lancet, 1997. **350**(9090): p. 1511-1515.
77. Peters, N., et al., *Spectrum of mutations in biopsy-proven CADASIL: implications for diagnostic strategies*. Arch Neurol, 2005. **62**(7): p. 1091-1094.
78. Ferreira, S., et al., *Mutações no gene NOTCH3 em doentes portugueses com CADASIL*. (comunicação oral) 10ª Reunião da Sociedade Portuguesa de Genética Humana, Coimbra, 9 a 11 de Novembro de 2006.
79. Arboleda-Velasquez, J.F., et al., *C455R notch3 mutation in a Colombian CADASIL kindred with early onset of stroke*. Neurology, 2002. **59**(2): p. 277-279.
80. Tuominen, S., et al., *Phenotype of a homozygous CADASIL patient in comparison to 9 age-matched heterozygous patients with the same R133C Notch3 mutation*. Stroke, 2001. **32**(8): p. 1767-1774.

81. Singhal, S., et al., *The influence of genetic and cardiovascular risk factors on the CADASIL phenotype*. Brain, 2004. **127**(Pt 9): p. 2031-2038.
82. Harper, P.S., *Practical Genetic Counselling*. 6th edition ed. 2004, London: Arnold Publishers.
83. *Guidelines for the molecular genetics predictive test in Huntington's disease*. International Huntington Association (IHA) and the World Federation of Neurology (WFN) Research Group on Huntington's Chorea. Neurology, 1994. **44**(8): p. 1533-1536.
84. Gardner, R.J.M. and G.R. Sutherland, *Chromosome abnormalities and genetic counselling*. 3rd edition ed. Oxford Monograph on Medical Genetics. 2004, New York: Oxford University Press.
85. Milunsky, A., et al., *The prenatal diagnosis of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) by mutation analysis*. Prenat Diagn, 2005. **25**(11): p. 1057-1058.
86. Konialis, C., et al., *Pregnancy following preimplantation genetic diagnosis of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL)*. Prenat Diagn, 2007. **27**(11): p. 1079-1083.
87. Escary, J.L., et al., *Evaluation of DHPLC analysis in mutational scanning of Notch3, a gene with a high G-C content*. Hum Mutat, 2000. **16**(6): p. 518-526.
88. Jiang, R., et al., *Sequence-based prioritization of nonsynonymous single-nucleotide polymorphisms for the study of disease mutations*. Am J Hum Genet, 2007. **81**(2): p. 346-360.
89. Mullenbach, R., P.J. Lagoda, and C. Welter, *An efficient salt-chloroform extraction of DNA from blood and tissues*. Trends Genet, 1989. **5**(12): p. 391.
90. Cotton, R.G. and C.R. Scriver, *Proof of "disease causing" mutation*. Hum Mutat, 1998. **12**(1): p. 1-3.
91. National Human Genome Research Institute: <http://www.genome.gov/> (data de acesso: 18/10/2007).
92. Cotton, R.G., *Communicating "mutation:" Modern meanings and connotations*. Hum Mutat, 2002. **19**(1): p. 2-3.
93. Baralle, D. and M. Baralle, *Splicing in action: assessing disease causing sequence changes*. J Med Genet, 2005. **42**(10): p. 737-748.

94. den Dunnen, J.T. and M.H. Paalman, *Standardizing mutation nomenclature: why bother?* Hum Mutat, 2003. **22**(3): p. 181-182.
95. The Human Gene Mutation Database: <http://www.hgmd.cf.ac.uk/> (data de acesso: 18/10/2007).