Exon sequencing of the gene encoding UCMA/GRP in healthy and clinical subjects

Ebba Frånland

Thesis performed at the Division of Clinical Chemistry, Department of Clinical and Experimental Medicine, Faculty of Health Sciences at Linköping University

2011-04-14

LITH-IFM-A-EX--11/2459--SE
Exon sequencing of the gene encoding UCMA/GRP in healthy and clinical subjects

Ebba Frånlund

Thesis performed at the Division of Clinical Chemistry, Department of Clinical and Experimental Medicine, Faculty of Health Sciences at Linköping University

2011-04-14

Supervisors
Per Magnusson, Associate professor
Majid Osman, PhD

Examiner
Uno Carlsson, Professor
Exon sequencing of the gene encoding UCMA/GRP in healthy and clinical subjects

Ebba Frånlund

Mineralization of soft tissues can cause significantly increased morbidity and mortality. The mechanism for this process is still unknown; however, patients with chronic kidney disease (CKD) are at high risk of developing vascular calcifications. Coronary artery calcification occurs faster in CKD patients undergoing dialysis in comparison with the general population. The pathological process of vascular calcification is the leading cause of death in patients with CKD.

Upper zone of growth plate and cartilage matrix associated protein (UCMA) is a novel vitamin K dependent (VKD) protein expressed in bone and the vascular system. The UCMA protein contains 15 γ-carboxyglutamic acid (Gla) residues in its 138 residue sequence which is the highest ratio between the number of Gla-residues and the size of the mature protein found in any protein so far. These Gla-residues form a domain that gives unique calcium binding properties for UCMA with high affinity for calcium phosphate crystals (i.e., hydroxyapatite). Even though the function of UCMA remains to be elucidated, it has been speculated that UCMA inhibits calcification of soft tissues and could therefore have a protective function against vascular calcification. Any mutations in the gene coding for UCMA might lead to a diminished function or defective protein.

The aim of this study was to determine whether the gene encoding UCMA in patients with the most progressed stage of CKD (stage 5 CKD) contained any mutations. This was accomplished by performing a full re-sequencing of all five exons with dideoxy sequencing in 16 patients with stage 5 CKD on hemodialysis. If any mutations were discovered, pyrosequencing would be performed on 98 healthy control individuals. This would help to determine if the mutation was exclusive for the patients or existed in the general population as well.

Genomic DNA was extracted from whole blood originating from 16 patients with CKD on hemodialysis. Each of UCMA’s five exons were amplified with PCR and the results were visualized using gel electrophoresis. Each exon was re-sequenced and pyrosequencing was performed on 98 healthy control samples. The acquired results were compared with the sequence of the UCMA gene identified at NCBI-GenBank (NCBI, build 37.2, NM_145314.1, Gene ID: 221044) and the Ensemble genome browser (ENSG00000165623). In addition, the frequencies of each SNP were calculated and compared with a study at the Ensemble database originating from the 1000 genomes project (1000GENOMES:low_coverage: CEU).

Because the population of our study group was too small to yield appropriate power for statistical calculations, no definite conclusions could be drawn from the acquired results. Nevertheless, this is the first patient group with CKD ever studied and should thus be regarded as a pilot study due to the limited size. However, no indication was found that UCMA had major defects in the investigated patients. Instead, a heterozygous transversion mutation was found in SNP rs4750328, indicating that the site of this SNP is subject to other modifications. Furthermore, a novel SNP was discovered which has not been described in other populations to our knowledge. The novel SNP is non-synonymous (i.e., causes an amino acid exchange) and located at the carboxyl-terminal of the protein. A serine is incorporated instead of threonine giving a 138Thr>Ser change since the last ACC codon in exon 5 (adjacent to the stop codon) is altered to an AGC codon. The UCMA 138Thr>Ser polymorphism was submitted to the dbSNP database and has been assigned the accession number ss283927876, which will be publicly available upon the release of the next dbSNP Build, B134. In order to determine the physiological significance of the discovered SNP, functional studies are required on both the wild-type and mutated UCMA variants.

UCMA, GRP, PCR, DNA-sequencing, dideoxy-sequencing, pyrosequencing
Copyright

The publishers will keep this document online on the Internet – or its possible replacement – from the date of publication barring exceptional circumstances. The online availability of the document implies permanent permission for anyone to read, to download, or to print out single copies for his/her own use and to use it unchanged for non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional upon the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its www home page: http://www.ep.liu.se/.

Upphovsrätt

Detta dokument hålls tillgängligt på Internet – eller dess framtida ersättare – från publiceringsdatum under förutsättning att inga extraordinära omständigheter uppstår.

Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för icke-kommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnd som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se förlagets hemsida http://www.ep.liu.se/
Preface

This master thesis was performed at the Division of Clinical Chemistry, Department of Clinical and Experimental Medicine, Faculty of Health Sciences at Linköping University under the supervision of Per Magnusson and Majid Osman.

The Division of Clinical Chemistry is a clinical laboratory organisation comprising all hospitals and district health care centres within the County Council of Östergötland where patient samples are sent for analysis in the areas of allergy, haematology, proteins, hormones, routine chemistry, immunochemistry, molecular biology, blood gases and coagulation among others. The Department of Clinical and Experimental Medicine at Linköping University conducts research and development work in cooperation with the County Council of Östergötland. Research and development is performed to maintain and improve the present competence and offer the health care a modern service in laboratory medicine of a high standard.
Abstract

Mineralization of soft tissues can cause significantly increased morbidity and mortality. The mechanism for this process is still unknown; however, patients with chronic kidney disease (CKD) are at high risk of developing vascular calcifications. Coronary artery calcification occurs faster in CKD patients undergoing dialysis in comparison with the general population. The pathological process of vascular calcification is the leading cause of death in patients with CKD.

Upper zone of growth plate and cartilage matrix associated protein (UCMA) is a novel vitamin-K dependent (VKD) protein expressed in bone and the vascular system. The UCMA protein contains 15 γ-carboxyglutamic acid (Gla) residues in its 138 residue sequence which is the highest ratio between the number of Gla-residues and the size of the mature protein found in any protein so far. These Gla-residues form a domain that gives unique calcium binding properties for UCMA with high affinity for calcium phosphate crystals (i.e., hydroxyapatite). Even though the function of UCMA remains to be elucidated, it has been speculated that UCMA inhibits calcification of soft tissues and could therefore have a protective function against vascular calcification. Any mutations in the gene coding for UCMA might lead to a diminished function or defective protein.

The aim of this study was to determine whether the gene encoding UCMA in patients with the most progressed stage of CKD (stage 5 CKD) contained any mutations. This was accomplished by performing a full re-sequencing of all five exons with dideoxy sequencing in 16 patients with stage 5 CKD on haemodialysis. If any mutations were discovered, pyrosequencing would be performed on 98 healthy control individuals. This would help to determine if the mutation was exclusive for the patients or existed in the general population as well.

Genomic DNA was extracted from whole blood originating from 16 patients with CKD on haemodialysis. Each of UCMA’s five exons were amplified with PCR and the results were visualized using gel electrophoresis. Each exon was re-sequenced and pyrosequencing was performed on 98 healthy control samples. The acquired results were compared with the sequence of the UCMA gene identified at NCBI-GenBank (NCBI, build 37.2, NM_145314.1, Gene ID: 221044) and the Ensemble genome browser (ENSG00000165623). In addition, the frequencies of each SNP were calculated and compared with a study at the Ensemble database originating from the 1000 genomes project (1000GENOMES:low_coverage:CEU).

Because the population of our study group was too small to yield appropriate power for statistical calculations, no definite conclusions could be drawn from the acquired results. Nevertheless, this is the first patient group with CKD ever studied and should thus be regarded as a pilot study due to the limited size. However, no indication was found that UCMA had major defects in the investigated patients. Instead, a heterozygous transversion mutation was found in SNP rs4750328, indicating that the site of this SNP is subject to other modifications. Furthermore, a novel SNP was discovered which has not been described in other populations to our knowledge. The novel SNP is non-synonymous (i.e., causes an amino acid exchange) and located at the carboxyl-terminal of the protein. A serine is incorporated instead of threonine giving a 138Thr>Ser change since the last ACC codon in exon 5 (adjacent to the stop codon) is altered to an AGC codon. The UCMA 138Thr>Ser polymorphism was submitted to the dbSNP database and has been assigned the accession number ss283927876, which will be publicly available upon the release of the next dbSNP Build, B134. In order to determine the physiological significance of the discovered SNP, functional studies are required on both the wild-type and mutated UCMA variants.
Sammanfattning

Mineralisering av mjuk vävnad med hydroxiapatit kan signifikant öka sjuklighet och dödlighet. Mekanismen för denna process är fortfarande okänd, men patienter med kronisk njursjukdom som behandlas med dialys har en speciellt hög risk att utveckla vaskulär kalcifiering vilket också är den ledande dödsorsaken för denna patientgrupp. Dessutom sker kalcifiering av kransarter snabbare i dessa patienter jämfört med den generella populationen.


Syftet med den här studien var att undersöka om genen som kodar för UCMA i patienter med kronisk njursjukdom i steg fem innehåller några mutationer. Detta skulle ske genom att UCMAfs fem exoner sekvences med dideoxy sekvenserings hos 16 patienter. Om några mutationer hittades skulle pyrosekvensering utföras på 98 friska kontrollpersoner. Detta skulle hjälpa till att bestämma om mutationen var exklusiv för patienterna eller om den även finns i den generella populationen.

# Table of contents

1 **INTRODUCTION** ................................................................. 1  
   1.1 **AIM** ............................................................................. 1  
       1.1.1 **Limitations** ......................................................... 1  

2 **THEORETICAL BACKGROUND** ......................................................................................... 3  
   2.1 **SINGLE NUCLEOTIDE POLYMORPHISM** .................................................. 3  
   2.2 **VITAMIN-K DEPENDENT PROTEINS** ...................................................... 3  
       2.2.1 **Upper zone of growth plate and cartilage matrix associated protein** ........ 4  
       2.2.2 **UCMA and chronic kidney disease** ................................................. 5  
   2.3 **POLYMERASE CHAIN REACTION** .......................................................... 6  
   2.4 **GEL ELECTROPHORESIS** ....................................................................... 7  
   2.5 **DideoxY SEQUENCING** ......................................................................... 8  
   2.6 **Pyrosequencing** ...................................................................................... 9  

3 **MATERIALS AND METHODS** ......................................................................................... 11  
   3.1 **SPECIMEN AND ETHICAL PERMISSION** ............................................... 11  
   3.2 **PCR** ......................................................................................... 11  
   3.3 **GEL ELECTROPHORESIS** ...................................................................... 12  
   3.4 **DideoxY SEQUENCING** ....................................................................... 13  
   3.5 **Pyrosequencing** .................................................................................... 14  

4 **RESULTS** ..................................................................................................................... 15  
   4.1 **PCR AND GEL ELECTROPHORESIS** .................................................. 15  
   4.2 **DideoxY SEQUENCING** ....................................................................... 19  
   4.3 **Pyrosequencing** ...................................................................................... 21  

5 **DISCUSSION** ................................................................................................................. 23  
   5.1 **FUTURE STUDIES** ............................................................................. 25  

6 **CONCLUSIONS** ............................................................................................................ 27  

7 **ACKNOWLEDGEMENTS** .............................................................................................. 29  

8 **REFERENCES** .............................................................................................................. 31  

**APPENDIX 1** .................................................................................................................... 33  
   **PRIMERS** ................................................................................................. 33  

**APPENDIX 2** .................................................................................................................... 35  
   **PCR PROGRAM 1** .................................................................................. 35  
   **PCR PROGRAM 2** .................................................................................. 35  
   **PCR PROGRAM 3** .................................................................................. 35
<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>APS</td>
<td>adenosine 5’ phosphosulfate</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>CKD</td>
<td>chronic kidney disease</td>
</tr>
<tr>
<td>ddNTPs</td>
<td>dideoxy nucleoside triphosphates</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>deoxycytidine triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>dNTPs</td>
<td>deoxy nucleoside triphosphates</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dTTP</td>
<td>deoxythymidine triphosphate</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>Gla</td>
<td>γ-carboxyglutamic acid</td>
</tr>
<tr>
<td>Glu</td>
<td>glutamic acid</td>
</tr>
<tr>
<td>GRP</td>
<td>Gla-rich protein</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PP_i</td>
<td>pyrophosphate</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>SSB</td>
<td>single-stranded binding protein</td>
</tr>
<tr>
<td>UCMA</td>
<td>upper zone of growth plate and cartilage matrix associated protein</td>
</tr>
<tr>
<td>VKD</td>
<td>vitamin-K dependent</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Aim

Upper zone of growth plate and cartilage matrix associated protein (UCMA) is a novel vitamin-K dependent (VKD) protein expressed in bone and the vascular system. The UCMA protein contains 15 γ-carboxyglutamic acid (Gla) residues which is the highest ratio between the number of Gla-residues and the size of the mature protein found in any protein so far. These Gla-residues form a domain that gives unique calcium binding properties for UCMA with high affinity for calcium phosphate crystals (i.e., hydroxyapatite). It has, therefore, been speculated that UCMA inhibits calcification of soft tissues and thereby has a protective function against vascular calcification. [1-3] Although the mechanism of this process is largely unknown, patients suffering from chronic kidney disease (CKD) are at high risk of developing vascular calcifications. [1] Coronary artery calcification occurs faster in CKD patients undergoing dialysis in comparison with the general population and vascular calcifications are the leading cause of death for CKD patients. [4-5]

It is not known why certain individuals have an increased risk of developing vascular calcifications. It is, therefore, of importance to investigate if the sequence of the coding regions (exons) in the UCMA gene contains any mutations in patients with CKD which could cause a mutated protein. An amino acid exchange in the sequence of UCMA does not definitely lead to a reduced function, but could affect the functional properties of UCMA which might lead to a defective protein.

The aim of this study was to determine whether the gene encoding UCMA in patients with the most progressed stage of CKD (stage 5 CKD) contained any mutations. This was accomplished by performing a full re-sequencing of all five exons with dideoxy sequencing in 16 patients with stage 5 CKD on hemodialysis. If any mutations were discovered, pyrosequencing would be performed on 98 healthy control individuals. This would help to determine if the mutation was exclusive for the patients or existed in the general population as well.

1.1.1 Limitations

The entire exons were only re-sequenced for the patient samples whereas the sequence of the UCMA gene identified at NCBI-GenBank (NCBI, build 37.2, NM_145314.1, Gene ID: 221044) and the Ensemble genome browser (ENSG00000165623) was used as reference material in the dideoxy sequencing experiments. Pyrosequencing was only performed on 98 healthy control samples for the novel SNP position.
2 Theoretical background

2.1 Single nucleotide polymorphism

Deoxyribonucleic acid (DNA) is a linear polymer composed of four different nucleotides, A, T, G and C, which each consists of a sugar, a phosphate and a base. The four nucleotides are bound together in a specific sequence making a strand and it is this sequence that holds our genetic information. [6-7]

DNA exists as a double helix consisting of two complementary nucleic acid strands. [6-7] This means that for each given location of the DNA-sequence, called locus, there are two nucleotides, called alleles. The helix forms chromosomes and each person carries 22 pairs of homologous chromosomes and one pair of sex chromosomes. One of the chromosomes in a pair is inherited from the mother whereas the other originates from the father. At fertilization, a recombination process known as crossover takes place and the two homologous chromosomes exchange genetic material between each other, implying that the offspring receives a unique genetic composition. [8]

A single nucleotide polymorphism (SNP) refers to a common genetic variation at any given chromosomal locus, which generally occurs when the variation is ≥ 1 percent in a population. [9] Mutations, on the other hand, are normally rare and do not affect the general population. As there are two homologous chromosomes, an individual may be homozygous for the most frequent allele (or wild-type), heterozygous (i.e., one chromosome is affected whereas the other contains the wild-type allele), or homozygous for the least frequent allele (or minor allele). SNPs can be either synonymous or non-synonymous. It is synonymous when the SNP does not result in a change of amino acid in the protein whereas it is non-synonymous if it does. [8] SNPs tend to be relatively stable genetically and occur frequently in the genome, with an average of 1 SNP per 1000 base pairs. [9-10]

Most SNPs exist in the non-coding areas of the genome where they may or may not be responsible for diseases. SNPs found in coding areas may directly affect the gene product (protein) and could therefore be associated with, and occasionally cause, a specific disease making them a target for research. SNPs may also act as biological markers to find a disease since they usually are located near genes associated with various diseases. An SNP may also determine the progression of the disease or confer susceptibility or resistance to the disease and can affect a person’s response to drug therapy. [9-10]

2.2 Vitamin-K dependent proteins

VKD proteins play an essential role in many biological functions such as skeletal mineralization, growth control, apoptosis, signal transduction and in the coagulation process. The mechanistic function is mainly attributable to a Gla-domain, which is characteristic for all VKD proteins. The protein matures when a number of repetitive glutamic acid (Glu) residues become carboxylated into Gla-residues by γ-glutamyl carboxylase. This domain is responsible for calcium binding properties; however VKD proteins can also bind coagulation factors by means of a calcium bridge. [2]
2.2.1 Upper zone of growth plate and cartilage matrix associated protein

UCMA was first discovered as a novel secreted cartilage-specific protein. [11-12] It was suggested that UCMA was a highly specific marker for distal chondrocytes unique to cartilage. [12] Later on it was discovered, by another independent group, that UCMA is a VKD protein with numerous Gla-residues and was consequently named Gla-rich protein (GRP). [1-3] Here, the protein will be called UCMA since this is the official name according to the HUGO Gene Nomenclature Committee.

UCMA is a small protein with a total molecular weight of 17-kDa and a net negative charge density of 20, but is nevertheless insoluble at natural pH. At first UCMA was believed to be a cartilage-specific protein, but it has now been found to accumulate in bone, skin and the vascular system as well. [1-2]

![Figure 1: Predicted 3D model of UCMA where the green and yellow dots represent the N- and C-terminal, respectively.](image)

The UCMA gene is according to NCBI-GeneBank and Ensemble genome browser composed of five exons, each 131, 66, 96, 99 and 434 base pairs (bp), respectively. [2-3] The transcript of UCMA encodes 138 amino acids and consists of a signal peptide, a coiled-coil domain and a cleavage site. UCMA is secreted as an uncleaved precursor into the extracellular matrix where it is cleaved by a furin-like protease into its amino- and carboxyl-terminal fragments of 37 and 74 amino acid residues, respectively. [11-12] The secondary structure of UCMA has yet to be determined by NMR or crystallography. However, fig. 1 shows a predicted 3D model generated by I-TASSER (http://zhang.bioinformatics.ku.edu/).

As mentioned, all VKD proteins are characterized by a Gla-domain. However, UCMA has a total of 15 Gla-residues and the highest ratio between the number of Gla-residues and the size of the mature protein of any protein known to date. The large number of Gla-residues makes UCMA a very efficient binder of calcium ions or calcium crystals with high affinity for hydroxyapatite. The Gla-domain in UCMA differs significantly in comparison with other VKD proteins within the protein sequence and gene organization as the Gla-domain spreads throughout the entire protein in UCMA. However, the core Gla-domain is encoded by exon 4. This unique domain has a high degree of conservation since UCMA has orthologs in all taxonomic groups of vertebrates, with 78 % identity between sturgeon and human UCMA, and a paralog in bony fish. However, no orthologs have been identified in chickens. [1-3, 11-12]
The high evolutionary conservation as well as the high number of Gla-residues suggest that UCMA has a pivotal function. [2] The function is yet to be determined, nevertheless several hypotheses exist and some progress has been made.

The importance of Gla-residues for the coagulation process was accentuated when insufficient \( \gamma \)-glutamyl carboxylation was found to prevent Gla-mediated calcium binding and thereby inhibiting clot formation. [2] The high affinity for calcium binding suggests that UCMA has a function in the regulation of calcium in extracellular matrix where UCMA may play a role as a physiological calcium modulator. [1-2] When high accumulation of UCMA was found at sites of medial calcification co-localized with mineral deposits whereas this was almost absent in noncalcified areas in the vascular system, it was concluded that UCMA definitely is associated with ectopic mineralization of connective tissues. [1] Another study indicated that UCMA may participate in early chondrogenesis. [3] It is furthermore proposed that mammalian UCMA can bind anionic proteoglycans, which is highly abundant in cartilage, in the same way VDK proteins bind coagulation factors. [2]

### 2.2.2 UCMA and chronic kidney disease

Mineralization of soft tissues with calcium phosphate crystals (i.e., hydroxyapatite) can significantly increase morbidity and mortality. The mineralization process normally progress with age and occurs naturally in any body tissue, but particularly in skin, kidney, tendons and the cardiovascular system. [2] Therefore, the process must be actively inhibited and it is essential that inhibitors such as matrix Gla protein are expressed accurately. The mechanism for this process is still unknown; however, patients with CKD are at high risk of developing vascular calcifications. [1-2] Coronary artery calcification occurs faster in CKD patients undergoing dialysis in comparison with the general population and vascular calcifications are the leading cause of death for CKD patients. [4-5]

The effects of CKD on vascular calcification are the net result of multiple pathogenic mechanisms and it is likely that there are circulating inhibitors yet to be discovered. [5] It has been suggested that both extracellular calcium and phosphate increase the influx of phosphate into vascular smooth muscle cells which accelerates the mineralization. This hypothesis is supported by the progression of coronary artery calcification in patients with CKD, which seems to positively correlate with serum phosphate, the calcium \( \times \) phosphate product, and daily calcium intake. These findings support the hypothesis that a disturbance in the calcium and phosphate homeostasis in CKD patients plays a decisive role in the pathological progress of vascular calcification. [4] Since UCMA has a calcium binding domain it is interesting to investigate whether it contributes to the pathological process of vascular calcification.
2.3 Polymerase chain reaction

The polymerase chain reaction (PCR) was invented by Kary Mullis in 1983, for which he was awarded a Nobel Prize in Chemistry ten years later. The technique enables selective copying of a specific target DNA-sequence by mimicking the cellular DNA replication which occurs in the cells prior to cell division. This technique generates large amounts of DNA samples for downstream analysis such as DNA sequencing. [6-7, 13-14]

In order for PCR to work the so called flanking sequences on both sides of the target sequence on both DNA-strands must be known. Two different short oligonucleotides called primers are required, each complementary to a stretch of DNA on these flanking sequences. The primers anneal to the DNA-strands where they act as a starting point for the additions of new nucleotides and thus enable copying of the target sequence. [6-7, 13-15] For best results, a primer that is selective only to the specific stretch of DNA should be used to avoid unspecific priming that might otherwise create undesired background signals. Primers are in general 14-40 nucleotides long and should preferably contain approximately the same amount of each nucleotide. It is also important that primers do not form secondary structures or base pair with each other or with itself. [7, 13, 15]

In addition to the primers, deoxy nucleoside triphosphates (dNTPs), a Taq-polymerase and a buffer are needed. The four dNTPs, deoxyadenosine-, deoxycytidine-, deoxyguanosine- and deoxythymidine triphosphate (dATP, dCTP, dGTP and dTTP, respectively), act as building blocks for making of the new template. [6] Taq-polymerase catalyzes the synthesis of DNA by selecting the correct dNTP to incorporate. It is very heat stable and remains active after exposure to 94 °C. [7, 13, 15]

The buffer is used to create a positive environment for the reaction and contains KCl, \((\text{NH}_4)_2\text{SO}_4\) and magnesium. KCl assists the primer annealing to the template by binding to the phosphates on the backbone of the DNA-strand. \(\text{NH}_4^+\) helps to destabilize the hydrogen bonds and thus the DNA to remain in single stranded form. Taq-polymerase is dependent on magnesium to function properly and it is therefore crucial that the concentration of magnesium is optimal for polymerase activity. [13]

PCR is carried out in a thermal cycler, a programmable instrument that can alter temperature rapidly. The cycle consists of three steps (denaturation, annealing and extension), which are repeated 20-40 times. The temperatures and length of the steps depend, among other things, on which primers are used, the length of the target DNA sequence, what polymerase is used and the concentration of the dNTPs. [7, 13, 15]

An incubation time of 10-15 min, at 94-95°C, is required prior to start of the PCR cycling program to give the Taq-polymerase the hot start necessary to its activation. For DNA amplification, each cycle starts with a short denaturation step. In the denaturation step, the two DNA strands unwind into single strands at a temperature that is normally 94-95°C. The high temperature breaks the hydrogen bonds of the double helix leading to the formation of single-stranded DNA. [7, 13]

In the second step, the sample is cooled to the annealing temperature which is specific to the primers. At the annealing temperature the primer hybridizes with the DNA-strands on their specific sequences respectively. [7, 13] The primers melting temperature (\(T_m\)) is a good indicator of which annealing temperature should be used. \(T_m\) is the temperature where half of the primers are annealed to the DNA-strands. The easiest way of calculating the \(T_m\) is by simply adding 4°C for every C or G in the primer and 2°C for every A or T. This will give a rough estimation of the \(T_m\) and the annealing temperature is often 5°C below the calculated \(T_m\). [13, 15]
The third step of the cycle allows the target DNA sequence to be synthesized. The sample is allowed to reach the extension temperature, where elongation of both forward and reverse primers occurs and both strands of the target sequence are replicated. The extension temperature is chosen at the optimal extension temperature for Taq-polymerase, which is 72°C. [13, 15] Since more and more target DNA strands are made, these can be used as templates to create more DNA strands in the following cycle. This allows the newly formed DNA strands of the target DNA sequence to increase exponentially. [6-7, 14] The last cycle is then followed by a hold at the elongation temperature for final product extension. At the end of the program, the temperature is switched to 4°C for an indefinite time in order to store the samples correctly if they are not taken out immediately.

2.4 Gel electrophoresis

Gel electrophoresis is used to separate proteins or other macromolecules (such as RNA or DNA) after size when an electric current is applied over the gel. Agarose gels are frequently used in techniques which allow the DNA to move unhindered through the gel matrix according to their net charge. Since DNA has a negative net charge, the sample will migrate towards the positive pole when an electric current is applied over the gel. Gel electrophoresis is thus very sensitive to changes in pH, since it alters the overall net charge of a molecule. DNA-gels are always run horizontal and are submerged in buffer. [6, 14]

A horizontal 2 or 3 % agarose gel was used in this study. Agarose is a linear polysaccharide made up of the basic repeat unit agarobiose. When a gel is composed, agarose forms both intra- and inter-molecular hydrogen bonds within and between the agarose chains. In this way, a network of pores are fabricated which small molecules easily move thorough whereas larger molecules are almost immobile. Molecules of medium size migrate through the gel with various degree of facility. The pore size varies depending on how much agarose is used; the higher concentration of agarose, the smaller pore size will be obtained. The size of the studied DNA can be determined by using a base pair ladder of known size. [14]

The samples are mixed with a gel loading solution containing a buffer that facilitates the DNA sample to enter into the wells of the gel. The most common dye used for the loading solution is bromphenol blue. It is a small molecule that migrates without retardation, therefore giving a good indication about when to terminate the electrophoresis. Another component of the gel loading solution is often sucrose or glucose, which provides the sample with high density to prevent it from leaving the wells after it is applied onto the gel. [14]

In this study, a gel loading solution from Sigma-Aldrich containing bromphenol blue and sucrose was used. The solution also contained SDS, to help dissociate DNA-protein complexes which could interfere with the electrophoresis, and EDTA, to terminate the action of enzymes that require divalent cations.

In order to visualize the results a gel stain solution is added to the gel. Different fluoresceins, such as ethydium bromide and SYBR®, which intercalate into DNA are commonly used. [14] In this study, SYBR® Safe DNA gel stain was used which allows the DNA-bands on the gel to be visualized on a PhotoDoc-IT™ Imaging System equipped with a benchtop UV transilluminator (UVP, Upland, CA, USA).
2.5 Dideoxy sequencing

The most popular method for sequencing DNA is called dideoxy sequencing. It was developed in 1970s by Frederick Sanger and colleagues and awarded Sanger his second Nobel Prize in chemistry in 1980. [6, 15] Back in the 1970s, Sanger could only determine the sequence of 15-200 bases long fragments with reasonable accuracy [16] but the instruments used today are capable of determine sequences up to 1000 bases. [17]

The principle of the method is to mimic the cellular replication of DNA but also to incorporate dideoxy nucleoside triphosphates (ddNTPs). The PCR product undergoes a cleanup procedure and a primer is annealed to the single-stranded DNA template which is then extended by DNA polymerase producing a complementary strand. Since all four ddNTPs are present in the reaction at low concentrations, they will occasionally become incorporated instead of the dNTPs at random. The extension process is terminated every time a ddNTP is incorporated instead of a dNTP causing fragments of different sizes to be produced. ddNTPs are identical with its corresponding dNTP apart from the absence of a 3’-hydroxyl group, inhibiting a new nucleotide to bind, as shown by fig. 2. [6, 15]

![Figure 2: Overview over the difference between dNTP and ddNTP which obstructs ddNTP to bind new nucleotides. The illustrated base in each analog is either adenosine, cytidine, guanosine or thymidine.](image)

The DNA fragments can later be separated after size using electrophoresis. When dideoxy sequencing first was developed, four different samples with different ddNTPs had to be produced. The samples were loaded into separate wells on a polyacrylamide gel and electrophoresis was performed. Thereafter, autoradiography was used to visualize the radioactive DNA bands on the gel and the sequence was determined manually. [14-16] Since then, various instruments have been developed along with computer software which determines the sequence and the reliability of the readings. Today, the ddNTPs are labeled with different fluorescent tags causing the sequence to be determined in a single reaction. The sequence can thus easily be determined by electrophoresis followed by detection of the fluorescence of the fragments. Another advantage with the new procedure is that it has eliminated the use of radioactive reagents. [6] The Genetic Analyzer 3500 (Applied Biosystems, Carlsbad, California, USA) used in these experiments uses capillary electrophoresis to separate the fragments prior to sequence determination. [17]
2.6 Pyrosequencing

Pyrosequencing is a sequencing-by-synthesis method based on real-time detection of incorporated bases onto a primed template, assembling a new complementary DNA strand. During the DNA-polymerase reaction, pyrophosphate (PP\(_i\)) is released upon nucleotide incorporation displayed in Eq. 1. PP\(_i\) is then converted into adenosine triphosphate (ATP) by sulfurylase in the presence of adenosine 5’ phosphosulfate (APS) demonstrated by Eq. 2. ATP is then used in another reaction which results in the production of visible light by luciferase, as shown in Eq. 3. [10, 18-22] The emitted light is detected and measured by a pyrosequencing instrument. [19] The amount of light is proportional to the amount of released PP\(_i\), which, in turn, is proportional to the amount of incorporated nucleotide. Hence, the amount of released PP\(_i\) is equimolar to the quantity of incorporated nucleotides and the number of incorporated bases can therefore be established by comparing the intensities of the signals. [19, 22-23]

\[
(DNA)_n + dNTP \xrightarrow{\text{DNA-polymerase}} (DNA)_{n+1} + PP_i \tag{Eq. 1}
\]

\[
PP_i + APS \xrightarrow{\text{ATP sulfurylase}} ATP + SO_4^{2-} \tag{Eq. 2}
\]

\[
\text{Luciferin} + ATP \xrightarrow{\text{Luciferase}} \text{Oxyluciferin} + AMP + CO_2 + \text{light} \tag{Eq. 3}
\]

Pyrosequencing is always preceded by PCR in order to produce a sufficient amount of template in which one primer is biotinylated enabling separation of the DNA strands with streptavidin-sepharose. [10, 20-22] Therefore, there is only one direction for the reaction to progress when the sequencing primer is added to the sample. [21]

The instrument adds small volumes of dNTPs separately to the sample in a repeated cycle. If the dNTP is complementary to the template strand, it will be incorporated into the growing strand and release PP\(_i\). Since the process requires the detected light signal to exclusively originate from the last incorporated nucleotide, excess ATP and unincorporated nucleotides are therefore degraded by the enzyme apyrase between nucleotide additions to allow processivity, as shown in Eq. 4 and 5. It is, therefore, necessary to add new enzymes to each cycle of dNTP. Furthermore, single-stranded binding protein (SSB) is optionally added to the reaction in order to disrupt secondary structures in the template. [19-21]

\[
\text{ATP} \xrightarrow{\text{Apyrase}} AMP + 2P_i \tag{Eq. 4}
\]

\[
dNTP \xrightarrow{\text{Apyrase}} dNMP + 2P_i \tag{Eq. 5}
\]

Two types of strategies for pyrosequencing are available; cyclic and sequential. In the cyclic protocol the four bases are added according to a specific iterative cycle. This protocol is used when the target sequence or the SNP is not known. The sequential approach is applied when the target sequence is known and is typical for SNP scanning studies. The order in which the nucleotides are added is determined by the target sequence. If the sequence contains a SNP, both possible nucleotides are added separately. [10, 19] This protocol is preferred since it requires fewer nucleotide additions and leads to a more rapid procedure. Additionally, it allows flexibility in primer position and enables identification of closely located SNPs. [19]
3 Materials and methods

3.1 Specimen and ethical permission

A full exon re-sequencing study on DNA samples from 16 patients with CKD on haemodialysis was performed. The annotated reference sequence for the UCMA gene was identified at NCBI-GenBank (NCBI, build 37.2, NM_145314.1, Gene ID: 221044) and at Ensemble genome browser (ENSG00000165623). In an effort to investigate possible UCMA gene mutations, or variations, not found in the reference sequence, the results were compared with 98 healthy control individuals. The local research ethics committee of Linköping University, Sweden, approved this study.

3.2 PCR

One hundred μL of genomic DNA was extracted from whole blood using a BioRobot EZ1 instrument (Qiagen, Hilden, Germany). The instrument was equipped with an EZ1 DNA Blood Card containing the protocol for nucleic acid purification.

The whole blood was thawed at room temperature and mixed thoroughly to ensure homogeneity. Thereafter, 200 μL was transferred into 2 ml screw-capped sample tubes. The instrument was then loaded with the components of a EZ1 DNA Blood 200 μL Kit (Qiagen, Hilden, Germany); 2 ml screw-capped sample tubes, disposable tip holders, disposable filter-tips, 1.5 ml elution tubes and presealed reagent cartridges. A negative control was performed by loading a tube without any blood which should result in a tube with only water and no DNA. The whole blood was first lysed and magnetic particles were added to the sample which bound to DNA. This enabled magnetic separation of DNA from the rest of the blood. A program with extra wash and buffer instead of 80 % ethanol was selected in which the separation step was performed twice with a washing step in between. After the separation 100 μL of pure DNA was eluted into the elution tubes.

Primers were designed by Primer3 program (http://frodo.wi.mit.edu/primer3/). All primers passed the In silico PCR test provided by University of California Santa Cruz, which signifies that they are specific only to one part of the genome. For the dideoxy sequencing experiments the primers were designed to include the entire exon as well as a part of the intron on either side. As a result, the template sizes were 358, 296, 343, 288 and 595 bases for exons 1, 2, 3, 4 and 5, respectively. For the pyrosequencing experiments, shorter templates are better and therefore a 75 bases long template was used. All primers are shown in Appendix 1.

All samples were prepared as described in table 1. Since the UCMA gene contains five exons, five different mastermix solutions had to be made. In these solutions, only the selection of primers varied. Negative controls were prepared for each exon by adding 5 μL of water instead of DNA. The solutions were mixed in clear 96-well Multiplate® PCR Plates™ (Bio-Rad Laboratories, Hercules, USA) or MicroAmp™ 8-Tube Strip (Applied Biosystems, Carlsbad, CA, USA).

The optimal concentration of magnesium had earlier been investigated, at concentrations 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 and 3.5 mM, respectively. The two lowest concentrations did not yield any products on the gel. A magnesium concentration at 1.5 mM gave only a weak band on the gel whereas 2.0 mM yielded a stronger DNA band. For concentrations of 2.5 mM or higher, more than one band was detected which indicates unspecific binding of the primers. Therefore, additional magnesium was added to the samples in excess of the magnesium in the PCR buffer making the final concentration 2.0 mM.
Table 1: Sample preparation for PCR.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Concentration</th>
<th>Volume in well</th>
<th>Final amounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 X PCR buffer (Qiagen, Hilden, Germany)</td>
<td>25.0 mM</td>
<td>2.0 μL</td>
<td>1.0 X 1.5 mM Mg(^{2+})</td>
</tr>
<tr>
<td>Magnesium (Qiagen, Hilden, Germany)</td>
<td>2.5 mM</td>
<td>0.4 μL</td>
<td>0.5 mM</td>
</tr>
<tr>
<td>dNTP (VWR, Stockholm, Sverige)</td>
<td>5.0 μM</td>
<td>1.0 μL</td>
<td>125.0 μM</td>
</tr>
<tr>
<td>Forward primer (Biomers.net, Ulm, Germany)</td>
<td>5.0 μM</td>
<td>1.0 μL</td>
<td>5.0 pmol</td>
</tr>
<tr>
<td>Reverse primer (Biomers.net, Ulm, Germany)</td>
<td>5.0 μM</td>
<td>1.0 μL</td>
<td>5.0 pmol</td>
</tr>
<tr>
<td>HotStar Taq-polymerase (Qiagen, Hilden, Germany)</td>
<td>5.0 U/μL</td>
<td>0.1 μL</td>
<td>0.5 U</td>
</tr>
<tr>
<td>Nuclease free water (Sigma-Aldrich, St Louis, MN, USA)</td>
<td>9.5 μL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA-sample</td>
<td>5.0 μL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final volume</td>
<td>20.0 μL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to obtain the optimal annealing temperature with no unspecific binding of primers to DNA, a PCR program with a temperature gradient was carried out in a Mastercycler®PCR instrument (Eppendorf, Hamburg, Germany). The temperature gradient ranged from 55 to 65°C and the used PCR program was already optimized for the used Taq-polymerase. This procedure was employed to determine the annealing temperature for both dideoxy sequencing and pyrosequencing. The temperature gradient is shown in Appendix 2 where the PCR program used for optimization of annealing temperature is listed as “program 1”. The program used to obtain templates for the dideoxy sequencing and pyrosequencing experiments is listed as “program 2” in Appendix 2.

3.3 Gel electrophoresis

To visualize the results from the PCR performed for optimization of annealing temperature and dideoxy sequencing, gel electrophoresis was performed on a 2 % agarose gel made of 1.000 g NuSieve® 3:1 Agarose (Lonza, Rockland, ME, USA) dissolved in 50 ml 1xTBE-buffer (70mM TrisBase, 90 mM Boric acid, 2.8 mM EDTA, pH 8.3, Substrate unit, Clinical Microbiology, Linköping University Hospital). The gel solution was boiled for 2 minutes and 4 μL SYBR® Safe DNA gel stain (Invitrogen Molecular Probes, Eugene, OR, USA) was then added into the melted gel after it was being cooled to approximately 50°C.

5 μL of each sample were mixed with 5 μL of Gel Loading Solution (Sigma-Aldrich, St Louis, MO, USA) and applied on the gel. To conclude if the right exon was expressed, a base pair ladder (2.5 μL of 100 bp DNA ladder (Invitrogen Life Technologies, Carlsbad, CA, USA), 2.5 μL of ultrapure water (Millipore, Billerica, MA, USA) and 5 μL of Gel Loading Solution (Sigma-Aldrich, St Louis, MO, USA)) was added into one of the wells on each gel. The electrophoresis was set on 134 V until the sample had migrated about 5 cm on the gel. The DNA-bands on the gel was visualized using a PhotoDoc-IT™ Imaging System equipped with a benchtop UV transilluminator (UVP, Upland, CA, USA).

The DNA-templates were smaller (only 75 bases) in the pyrosequencing experiments than in the experiments with dideoxy sequencing. Therefore, a 3 % agarose gel was used as it provides smaller pores. In addition to the change of gel, a 50 bp DNA ladder (Invitrogen Life Technologies, Carlsbad, CA, USA) was used. Gel electrophoresis was carried out as previously described.
3.4 Dideoxy sequencing

The samples first underwent a PCR cleanup procedure to remove primers and dNTP from previously PCR. 3 μL of the sample was transferred to a new strip to which 1.2 μL of ExoSAP-IT® (USB Corporation, Cleveland, OH, USA) was added which contains exonuclease I and shrimp alkaline phosphatase. Exonuclease I degrades remaining single-stranded primers and extraneous single-stranded DNA produced in the previous PCR whereas the shrimp alkaline phosphatase hydrolyses redundant dNTP molecules. The sample was then put in a Mastercycler® PCR instrument (Eppendorf, Hamburg, Germany) which ran at 37°C for 15 minutes followed by 80°C for 15 minutes.

Two solutions with only single-stranded forward and reverse templates, respectively, were mixed according to table 2. In order to anneal the primers to the templates another PCR program was run, described as “program 3” in Appendix 2.

Table 2: Sample preparation for dideoxy sequencing.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Forward reaction</th>
<th>Reverse reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template</td>
<td>2.0 μL</td>
<td>2.0 μL</td>
</tr>
<tr>
<td>Forward primer 2,5 μM (Biomers.net, Ulm, Germany)</td>
<td>1.3 μL</td>
<td></td>
</tr>
<tr>
<td>Reverse primer 2,5 μM (Biomers.net, Ulm, Germany)</td>
<td></td>
<td>1.3 μL</td>
</tr>
<tr>
<td>Big Dye Mix (Applied Biosystems, Carlsbad, CA, USA)</td>
<td>1.0 μL</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Big Dye Terminator Sequencing Buffer (5X) (Applied Biosystems, Carlsbad, CA, USA)</td>
<td>1.5 μL</td>
<td>1.5 μL</td>
</tr>
<tr>
<td>MILLI-Q water (Substrate unit, Clinical Microbiology, Linköping University Hospital)</td>
<td>4.2 μL</td>
<td>4.2 μL</td>
</tr>
<tr>
<td>Final volume</td>
<td>10.0 μL</td>
<td>10.0 μL</td>
</tr>
</tbody>
</table>

A post-reaction cleanup was required to degrade and remove unused dNTP, DNA polymerase and buffer before the sample could be sequenced. The desired templates were first precipitated with EDTA (Sigma-Aldrich, St Louis, MN, USA) and sodium acetate (Merck, Darmstadt, Germany) and the resulting pellet was washed using ethanol. Next, the pellet was resuspended in Hi-Di™ Formamide (Applied Biosystems, Carlsbad, CA, USA) and transferred to a 96-well optical reaction plate (Applied Biosystems, Carlsbad, CA, USA). The plate was fitted into the genetic analyzer 3500 instrument from Qiagen for analysis programmed according to manufacturer’s recommendations. The process is described in the flow chart in fig. 3.

![Figure 3: Flow chart describing the post-reaction cleanup procedure which removed unused dNTP, DNA polymerase and buffer.](image-url)
3.5 Pyrosequencing

Twenty μL of PCR-product were mixed with 2 μL of Streptavidin-Sepharose™ High Performance (GE Healthcare AB, Uppsala, Sweden), 20 μL of 2xBW buffer (10 mM Tris-HCL, pH 7.6, 2 M NaCl, 1 mM EDTA, 0.1 % Tween 20, Substrate unit, Clinical Microbiology, Linköping University Hospital) and 38 μL of ultrapure water (Millipore, Billerica, MA, USA). The PCR-plate was then agitated in a Thermomixer comfort (Eppendorf AG, Hamburg, Germany) at 1400 rpm for 5 minutes. Thereafter, single-stranded DNA was prepared using a PyroMark™ Q24 vacuum prep workstation (Biotage, Uppsala, Sweden). The biotinylated strands were first immobilized on streptavidin-coated supermagnetic beads. The sample was incubated in 70% ethanol (Substrate unit, Clinical Microbiology, Linköping University Hospital) for 5 seconds and 0.2 M NaOH (Substrate unit, Clinical Microbiology, Linköping University Hospital) for an additional 5 seconds before it was washed in 0.001 M Tris, pH 7.6 for 10 seconds.

The sample was released in 7.5 nmol sequencing primer (Biomers.net, Ulm, Germany) and 24.5 μL of annealing buffer (20 mM Tris-acetate, pH 7.6, 5 mM MgAc₂, Substrate unit, Clinical Microbiology, Linköping University Hospital) on a PyroMark™ Q24 Plate (Qiagen, Hilden, Germany). The samples were annealed to the primer by heating the plate to 80˚C for 2 minutes on a Stuart® digital hotplate SD160 (Bibby Scientific Limited, Staffordshire, UK) before they were allowed to cool to room temperature for 5 minutes. The plate was then fitted into the PyroMark™ Q24 instrument from Qiagen. In each run, the instrument has the capacity to analyze 24 samples out of which one was a negative control sample without DNA.

A sequence-specific dispensing protocol was designed using the PyroMark Assay Design 2.0 program in which eight downstream bases were sequenced in addition to the studied position for acquisition of reference signals, according to fig. 4. The studied sequence was therefore TSCTGATCCC, where the first T acts as a negative control and S is the variable position which is either C or G. The protocol gives rise to a unique sequence to each of the two allelic alternatives possible for the SNP. The wild-type allele gives double CC at the variable position whereas the G peak is absent. Heterozygosity shows one and a half C peak following a half G peak. For homozygosity, one full peak of G and one full peak of C are shown on the pyrogram. The similarity in the predicted pattern and the raw data allows the SNP to be determined.

Figure 4: The theoretical pyrosequencing output for the analyzed SNP. The bases under the horizontal axis show the dispensation order. An incorporation of a single base gives a signal magnitude corresponding to a value of 1, whereas a signal value of 2 is obtained for two nucleotides, and so on. The wild-type allele gives double CC at the variable position whereas the G peak is absent. Heterozygosity shows one and a half C peak following a half G peak. For homozygosity, one full peak of G and one full peak of C are shown on the pyrogram.
4 Results

4.1 PCR and gel electrophoresis

The PCR and gel electrophoresis results in the experiments made to optimize the annealing temperature showed that the amplified target sequence gave a pure band on the gel and was free from unspecific background amplifications. Further, all the negative controls showed no contamination, suggesting successful PCR setup. The annealing temperature was set to 60.5°C for dideoxy sequencing since this temperature gave a satisfactory result for all exons. For the same reason, the annealing temperature was set to 65.0°C for pyrosequencing. In the dideoxy sequencing experiments a 100 bp DNA ladder was used whereas a 50 bp DNA ladder was used in the pyrosequencing experiments. Complete PCR programs can be found in Appendix 2 and fig. 5-9 show the results for dideoxy sequencing for exon 1-5, respectively. Results from the experiment conducted to optimize the annealing temperature for pyrosequencing are viewed in fig. 10.

![Figure 5: Results of temperature program for exon 1. The gel revealed amplification of the 358 bp fragment and that no unspecific background amplification occurred. The weak result for 63.9°C indicates that the sample has not been properly transferred to the well.](image1)

![Figure 6: Results of temperature program for exon 2. The gel revealed amplification of the 296 bp fragment and that no unspecific background amplification occurred. The result for 63.9°C indicates that the sample has not been properly transferred to the well.](image2)
Figure 7: Results of temperature program for exon 3. The gel revealed amplification of the 343 bp fragment and that unspecific background amplification occurred for 55.0, 55.2, 55.8 and 63.9 °C.

Figure 8: Results of temperature program for exon 4. The gel revealed amplification of the 288 bp fragment and that no unspecific background amplification occurred. The result for 57.8°C indicates that the sample has been contaminated by the DNA ladder.

Figure 9: Results of temperature program for exon 5. The gel revealed amplification of the 595 bp fragment and that unspecific background amplification occurred for 55.0 – 59.1 °C.
Figure 10: Results of temperature program prior to pyrosequencing. The gel revealed amplification of the 75 bp fragment and that no unspecific background amplification has occurred.

All gels performed on the patient samples prior to dideoxy sequencing also showed pure bands and show none or very little unspecific background amplifications. All the negative controls showed no contamination, suggesting a successful PCR setup for these experiments as well. A 100 bp DNA ladder was used and the results are visualized in fig. 11-15 for exons 1-5, respectively. No gel electrophoresis was performed prior to pyrosequencing experiments why no results can be shown.

Figure 11: Results from amplification of exon 1 in patient samples.
Figure 12: Results from amplification of exon 2 in patient samples.

Figure 13: Results from amplification of exon 3 in patient samples. The result from patient 3 indicate that the sample has not been properly transferred to the well whereas the result for patients 4 and 5 indicates that the samples have been contaminated by the DNA ladder.

Figure 14: Results from amplification of exon 4 in patient samples. The results for patients 1, 2 and 5 indicate that the sample has not been properly transferred to the well.
4.2 Dideoxy sequencing

Heterozygosity was discovered at several SNPs and in various patients. The acquired results are shown in table 3 where R is the nucleotide code describing heterozygosity between A and G; Y describes heterozygosity between C and T; S describes heterozygosity between C and G; K describes heterozygosity between G and T; whereas M describes heterozygosity between A and T.

All five exons in the patients’ DNA were sequenced and aligned with the sequence of the UCMA gene identified at NCBI-GenBank (NCBI, build 37.2, NM_145314.1, Gene ID: 221044) and the Ensemble genome browser (ENSG00000165623) using NCBI’s BLAST function. By comparing the calculated frequency acquired in the experiments with the frequency of each allele obtained from a reference study at the Ensemble database (1000GENOMES:low_coverage:CEU) any variation could be established, as shown in table 4. Unfortunately, the chosen reference study did not contain the frequency of all the studied SNPs.

At SNP rs4750328 (A/G) in intron 2, a heterozygous transversion mutation was found in patient 4 which involved an exchange of the ancestral A allele to a T base. Instead of A/G, this patient had a T/C heterozygosity at this position. In patient 7, a novel non-synonymous coding SNP was found. This SNP involves an alteration of the last ACC codon for threonine in exon 5 (adjacent to the stop codon) to an AGC serine codon (138Thr>Ser).
Table 3: Acquired results from diodeoxy sequencing experiments.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Ancestral allele</th>
<th>Alleles</th>
<th>Location in gene</th>
<th>Patient 1</th>
<th>Patient 2</th>
<th>Patient 3</th>
<th>Patient 4</th>
<th>Patient 5</th>
<th>Patient 6</th>
<th>Patient 7</th>
<th>Patient 8</th>
<th>Patient 9</th>
<th>Patient 10</th>
<th>Patient 11</th>
<th>Patient 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2399954</td>
<td>G</td>
<td>A/G</td>
<td>Intron 5-1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>rs942418</td>
<td>G</td>
<td>A/G</td>
<td>Intron 1-2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>rs41291317</td>
<td>A</td>
<td>A/G</td>
<td>Exon 2</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs10796042</td>
<td>G</td>
<td>G/T</td>
<td>Intron 2-3</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>T</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>rs3829927</td>
<td>C</td>
<td>C/T</td>
<td>Intron 2-3</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>rs3829926</td>
<td>C</td>
<td>C/T</td>
<td>Intron 2-3</td>
<td>Y</td>
<td>C</td>
<td>Y</td>
<td>C</td>
<td>C</td>
<td>Y</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>Y</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>rs4750328</td>
<td>A</td>
<td>A/G</td>
<td>Intron 2-3</td>
<td>G</td>
<td>R</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>R</td>
<td>R</td>
<td>G</td>
</tr>
<tr>
<td>rs3829925</td>
<td>G</td>
<td>G/T</td>
<td>Exon 3</td>
<td>K</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>K</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>K</td>
<td>K</td>
<td></td>
</tr>
<tr>
<td>rs4750327</td>
<td>A</td>
<td>A/G/T</td>
<td>Intron 3-4</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs7893239</td>
<td>G</td>
<td>C/G</td>
<td>Intron 3-4</td>
<td>C</td>
<td>S</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>S</td>
<td>S</td>
<td>C</td>
<td>S</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>rs4750320</td>
<td>A</td>
<td>A/G</td>
<td>Intron 3-4</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs74123515</td>
<td>C</td>
<td>C/T</td>
<td>Exon 4</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>rs80070283</td>
<td>A</td>
<td>A/G</td>
<td>Exon 4</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs2281797</td>
<td>C</td>
<td>A/C</td>
<td>Intron 4-5</td>
<td>A</td>
<td>M</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>C</td>
</tr>
<tr>
<td>rs2281796</td>
<td>T</td>
<td>C/T</td>
<td>Exon 5</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11547943</td>
<td>G</td>
<td>C/G</td>
<td>Exon 5</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs7894436</td>
<td>G</td>
<td>A/G</td>
<td>Exon 5</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>rs17851610</td>
<td>A</td>
<td>A/T</td>
<td>Exon 5</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>rs76274665</td>
<td>G</td>
<td>A/G</td>
<td>Exon 5</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

Table 4: Comparison of allele frequency between reference study and the acquired results from patients participating in this study.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Alleles</th>
<th>Frequency* (%)</th>
<th>Calculated A (%)</th>
<th>Calculated C (%)</th>
<th>Calculated G (%)</th>
<th>Calculated T (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2399954</td>
<td>A/G</td>
<td>10.0 / 90.0</td>
<td>12.5</td>
<td>0</td>
<td>87.5</td>
<td>0</td>
</tr>
<tr>
<td>rs942418</td>
<td>A/G</td>
<td>0.8 / 99.2</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>rs41291317</td>
<td>A/G</td>
<td>99.2 / 0.8</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs10796042</td>
<td>G/T</td>
<td>50.8 / 49.2</td>
<td>0</td>
<td>0</td>
<td>40,625</td>
<td>59,375</td>
</tr>
<tr>
<td>rs3829927</td>
<td>C/T</td>
<td>Not available</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs3829926</td>
<td>C/T</td>
<td>65.0 / 35.0</td>
<td>0</td>
<td>75</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>rs4750328</td>
<td>A/G</td>
<td>21.7 / 78.3</td>
<td>12.5</td>
<td>3,125</td>
<td>81,25</td>
<td>3,125</td>
</tr>
<tr>
<td>rs3829925</td>
<td>G/T</td>
<td>90.8 / 9.2</td>
<td>0</td>
<td>0</td>
<td>84,375</td>
<td>15,625</td>
</tr>
<tr>
<td>rs4750327</td>
<td>A/G/T</td>
<td>0.0 / 100.0 / 0.0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>rs7893239</td>
<td>C/G</td>
<td>74.2 /25.8</td>
<td>0</td>
<td>75</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>rs4750320</td>
<td>A/G</td>
<td>38.3 / 61.7</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>rs74123515</td>
<td>C/T</td>
<td>Not available</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs80070283</td>
<td>A/G</td>
<td>Not available</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs2281797</td>
<td>A/C</td>
<td>25.0 / 75.0</td>
<td>21,875</td>
<td>78,125</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs2281796</td>
<td>C/T</td>
<td>Not available</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>rs11547943</td>
<td>C/G</td>
<td>Not available</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>rs7894436</td>
<td>A/G</td>
<td>Not available</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>rs17851610</td>
<td>A/T</td>
<td>Not available</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs76274665</td>
<td>A/G</td>
<td>Not available</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

*Reference frequency originating from the Ensemble database (1000GENOMES:low_coverage:CEU), a study containing 60 participants from central Europe.
4.3 Pyrosequencing

The pyrogram data was evaluated by comparison of peak height, which is proportional to the number of incorporated bases, as described earlier. The C/C genotype represents homozygosity for the wild-type allele, whereas C/G stand for heterozygosity. None of the analyzed subjects displayed homozygosity for the minor G-allele (G/G). Representative examples of the genotypes found in the study are shown in fig. 16 and 17.

![Figure 16: A pyrogram representing C/C homozygosity at the analyzed SNP. This genotype was carried by 93 of the 98 healthy control individuals. E and S indicate the dispensation of enzyme and substrate mixtures, respectively.](image1)

![Figure 17: A pyrogram representing C/G heterozygosity at the analyzed SNP. This genotype was displayed by 5 of the 98 healthy control individuals. E and S indicate the dispensation of enzyme and substrate mixtures, respectively.](image2)
Discussion

Patients with CKD are at high risk of developing vascular calcifications. Even though the mechanism for this process is largely unknown, the Gla-domains high affinity for calcium binding suggests that UCMA inhibits calcification in soft tissues and has therefore a protective function against vascular calcification. The aim of this study was to investigate if the UCMA gene in patients with stage 5 CKD contains any mutations which could explain why these patients suffer from high calcium deposits. However, no correlation between a defect gene coding for UCMA and the disease could be found, but other findings were made which warrant further experimental and clinical investigations.

The target sequence was annealed as intended in the PCR and was free from unspecific background amplification and other contaminations. The annealing temperature was optimized and set at a temperature which gave satisfactory results for all exons in the dideoxy sequencing experiments as well as for the pyrosequencing experiments, asserting a successful PCR setup. Some samples from the PCR did not present any result in the gel electrophoresis experiments while others showed to be contaminated by the DNA ladder. However, it was established that this was due to human errors, eliminating the suspicion of an unsuccessful PCR.

The method used in the dideoxy sequencing experiments was developed previously and used routinely in the laboratory. Thus, no optimization experiments had to be performed on this method. All five exons in the patients’ DNA were sequenced and aligned with the sequence of the UCMA gene identified at NCBI-GenBank (NCBI, build 37.2, NM_145314.1, Gene ID: 221044) and the Ensemble genome browser (ENSG00000165623) using NCBI’s BLAST function. The sequence for the UCMA gene in NCBI-GenBank and Ensemble genome browser was chosen as a reference since it includes a diverse population. Also, sequencing the entire gene or using pyrosequencing for every SNP in the control population would be very expensive as well as time consuming.

The calculated frequencies of each allele in the SNPs acquired in our experiments were compared with those obtained from a reference study originating from the 1000 genomes project. The reference study was published in the Ensemble database (1000GENOMES:low_coverage: CEU) and consisted of 60 participants of Central European descent, but did not, however, contain frequencies of all the studied SNPs. One could argue that these SNPs are not published in the database because those variations were not found in the 1000 genomes project and, thus, are not common for a Central European population. This would be consistent with our study as no variation for these SNPs was found in patient or control population which indicates that they are not common for a Swedish population either. However, no definite conclusions could be drawn from the acquired results in this study since the population in our patient study group was too small to yield appropriate power for statistical calculations. Nevertheless, this is the first patient group with CKD ever studied and should thus only be regarded as a pilot study due to the limited size.

The frequencies of the SNPs in our patient study and the pilot study from the 1000 genomes project are quite similar, suggesting that UCMA may not be defect in patients with stage 5 CKD. If so, UCMA would not be the cause of the patients’ high risk of developing vascular calcifications, though this does not exclude that some other substance which is involved in the expression of UCMA is affected. Therefore, it would be interesting to compare the concentration of UCMA in patients with stage 5 CKD with a healthy control population.
As stated earlier, some new findings were made. One patient had a heterozygous transversion mutation in SNP rs4750328. The patient had a T/C heterozygosity in this position instead of A/G which suggests that this position is subject to other modifications.

A new mutation was also discovered in one of the 16 patients which has not been described in other populations to our knowledge. In order to investigate this mutation further, pyrosequencing was performed on 98 healthy control samples. Pyrosequencing was preferred since it enables a shorter DNA-strand being sequenced than in dideoxy sequencing. The experiments revealed that five out of the 98 control samples were heterozygous for the novel mutation. This gave the major C-allele a frequency of 97.4 % in the control population and the minor G-allele 2.6 %, thus concluding that the novel mutation is a SNP.

The novel SNP is non-synonymous (i.e., causes an amino acid exchange) and located at the carboxyl-terminal of the protein. A serine is incorporated instead of threonine giving a 138Thr>Ser change since the last ACC codon in exon 5 (adjacent to the stop codon) is altered to an AGC codon. The UCMA 138Thr>Ser polymorphism was submitted to the dbSNP database and has been assigned the accession number ss283927876, which will be publicly available upon the release of the next dbSNP Build, B134.

In the hominiae lineage (humans, chimpanzees and gorillas), the preferred amino acid at the carboxyl-terminal is threonine whereas serine is not found in any of the 28 species aligned at the UniProt database. Instead, threonine is exchanged for isoleucine at this position in most other species. In addition, threonine and serine are structurally similar which is why the exchange might not cause any physiological significance. Studies have shown the two amino acids to be functionally interchangeable in some proteins, whereas an exchange results in a dysfunctional protein in others. [24] Therefore, any conclusion as to the physiological significance of the novel SNP is not possible to make without further studies.

The serine and threonine residues at the carboxyl-terminals are sometimes the site for phosphorylation. However, the two amino acids have different degree of phosphorylation which might cause a change in the process. This is, for example, demonstrated when a serine in the active site of serine proteases is exchanged for a threonine which results in the termination of the physiological function of the enzymes. In order to determine the physiological significance of the discovered SNP, functional studies are required on both the wild-type and mutated UCMA variants.
5.1 Future studies

Since UCMA is such a novel protein, there is yet much to be discovered. It would be very interesting to conduct functional studies on wild-type UCMA and establish its secondary structure. By examining mutated UCMA variants with the different SNPs, their impact on the protein could also be determined.

In order to acquire results with appropriate power, which yield statistical significance in testing the hypothesis of this study, a full re-sequencing study would be necessary on the UCMA gene in larger patient and control populations. Such a study would provide frequencies of all SNPs for both reference individuals and patients. By comparing the results to the 1000 genomes project (when the project is finished), the results could also be compared with a more diverse population.

This study did not investigate the circulating concentration of UCMA in the patients or in the control population. Even though the gene did not include any mutations in the patients, UCMA might not be expressed in the same amount as in healthy individuals. Thus, UCMA could still be involved in vascular calcification in these patients.

Measurements of the level of UCMA in serum and/or plasma could be established by enzyme-linked immunosorbent assay (ELISA). However, no ELISA kits are commercially available for measurements of UCMA as of today, but an ELISA assay can be set up since antibodies against UCMA are available. Though, the ELISA assays might not work if the SNPs cause structural changes resulting in no binding of the UCMA variant to the antibodies due to changed antigenic domains.

Furthermore, comparisons of UCMA with other inhibitors of vascular calcification, such as osteopontin and inorganic pyrophosphate, could be alternative investigations as to yield new knowledge concerning the mechanistic properties of UCMA.
6 Conclusions

The results of this study can neither confirm nor contradict any correlation between a defect gene coding for UCMA and patients with stage 5 CKD since the population of our patient study group was too small to yield significant power for statistical calculations. Nevertheless, this is the first patient group with CKD ever studied but should be more regarded as a pilot study due to the limited size. However, other findings were made which warrant further experimental and clinical investigations.

A heterozygous transversion mutation was found in SNP rs4750328 which was coding for a T/C heterozygosity instead of A/G for one patient. In addition, a novel non-synonymous SNP located in the coding region of the gene was discovered in both patient and control populations. The novel SNP is non-synonymous (i.e., causes an amino acid exchange) and located at the carboxyl-terminal of the protein. A serine is incorporated instead of threonine giving a 138Thr>Ser change since the last ACC codon in exon 5 (adjacent to the stop codon) is altered to an AGC codon. The UCMA 138Thr>Ser polymorphism was submitted to the dbSNP database and has been assigned the accession number ss283927876, which will be publicly available upon the release of the next dbSNP Build, B134. In order to determine the physiological significance of the discovered SNP, functional studies are required on both the wild-type and mutated UCMA variants.
7 Acknowledgements

I would like to thank all the personnel at the Division of Clinical Chemistry, Department of Clinical and Experimental Medicine, Faculty of Health Sciences at Linköping University for their kind welcome and support throughout this project.

Special thanks go out to my supervisors Per Magnusson and Majid Osman for their commitment and patience. Thank you Per for the opportunity to perform my master thesis in such an interesting area and never failing to answer any of my questions. Majid, thank you for your guidance and helping me understand the theoretical background to the methods as well as the execution of the experiments. Thank you both for sharing your tremendous knowledge and all your help in writing this report!

Ewa Lönn Karlsson and Karin Ehrlin, thank you for your company and help with the DNA extraction from whole blood.

Thank you Annette Molbaek and Åsa Schippert for all your help with the dideoxy experiments which was a new area for me, your help has been invaluable.

Peter Nord Andersson, thank you for helping me edit the figures.

Finally, I would like to thank my beloved family for all their love, support and belief in me.
8 References


### Primers

**Table 5: Primers (Biomers.net, Ulm, Germany).**

<table>
<thead>
<tr>
<th>Method</th>
<th>Exon 1 Forward primer</th>
<th>Sequence 5'-3'</th>
<th>Length</th>
<th>T&lt;sub&gt;m&lt;/sub&gt; (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MegaBase</td>
<td>Exon 1</td>
<td>aacatctcgggtgtgctccatc</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Exon 1 Reverse primer</td>
<td>tccaacctgagcagccctacct</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td>MegaBase</td>
<td>Exon 2 Forward primer</td>
<td>ctcgagctggtaacgtgaa</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Exon 2 Reverse primer</td>
<td>agagccctgtcaaccagcagt</td>
<td>20</td>
<td>54</td>
</tr>
<tr>
<td>MegaBase</td>
<td>Exon 3 Forward primer</td>
<td>tgcacagaagggagaacagt</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Exon 3 Reverse primer</td>
<td>gactgtggagagggagggag</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td>MegaBase</td>
<td>Exon 4 Forward primer</td>
<td>aatcaacaattecccccacaac</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Exon 4 Reverse primer</td>
<td>ccatgtgatttcteccccaac</td>
<td>20</td>
<td>52</td>
</tr>
<tr>
<td>MegaBase</td>
<td>Exon 5 Forward primer</td>
<td>ctctgtgtaggtcaccacag</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Exon 5 Reverse primer</td>
<td>agggaaagacagcatgacacc</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td>Pyrosequencing</td>
<td>Forward primer</td>
<td>ctctctctacaacgecaccacaca</td>
<td>23</td>
<td>56</td>
</tr>
<tr>
<td>Pyrosequencing</td>
<td>Reverse primer</td>
<td>caatgtgggctacaagcagcggc</td>
<td>23</td>
<td>55</td>
</tr>
<tr>
<td>Pyrosequencing</td>
<td>Sequence primer</td>
<td>acaaccgccaccaca</td>
<td>15</td>
<td>48</td>
</tr>
</tbody>
</table>
Appendix 2

PCR program 1
94°C for 15 minutes
45 x
1. 94°C for 30 seconds
2. temperature according to table 6 for 45 seconds
3. 72°C for 60 seconds
72°C for 7 minutes
4°C ∞

Table 6: Temperature gradient for DNA-sequencing.

<table>
<thead>
<tr>
<th>Well</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>55.0 °C</td>
</tr>
<tr>
<td>2</td>
<td>55.2 °C</td>
</tr>
<tr>
<td>3</td>
<td>55.8 °C</td>
</tr>
<tr>
<td>4</td>
<td>56.7 °C</td>
</tr>
<tr>
<td>5</td>
<td>57.8 °C</td>
</tr>
<tr>
<td>6</td>
<td>59.1 °C</td>
</tr>
<tr>
<td>7</td>
<td>60.4 °C</td>
</tr>
<tr>
<td>8</td>
<td>61.7 °C</td>
</tr>
<tr>
<td>9</td>
<td>62.9 °C</td>
</tr>
<tr>
<td>10</td>
<td>63.9 °C</td>
</tr>
<tr>
<td>11</td>
<td>64.6 °C</td>
</tr>
<tr>
<td>12</td>
<td>64.9 °C</td>
</tr>
</tbody>
</table>

PCR program 2
94°C for 15 minutes
45 x
1. 94°C for 30 seconds
2. 60.5°C for 45 seconds
3. 72°C for 60 seconds
72°C for 7 minutes
4°C ∞

PCR program 3
96°C for 1 minute
25 x
1. 96°C for 10 seconds
2. 60°C for 5 seconds
3. 60°C for 3 minutes
4°C ∞