A molecular genetic study investigating the role of maternal and placental *laeverin* gene mutation and fetal whole genome copy number variations in the pathophysiology of preeclampsia

A Master of Science Thesis

By Hanne Listau Olsen

Department of Medical Biology, University of Tromsø and Women`s Health and Perinatology Research Group, Department of Clinical Medicine, University of Tromsø

Norway

2011

Acknowledgements

The work for this thesis was performed at the laboratory of the Women`s Health and Perinatology Research Group, Department of Clinical Medicine, University of Tromsø. The study utilized the biobank (maternal blood, umbilical cord blood and placental tissue samples) collected by Vasilis Sitras, MD, PhD and Åse Vårtun, MSc.

I would like to thank my principal supervisor Mona Nystad, Senior Scientist at the Division of Child and Adolescent Health, Department of Medical Genetics, University Hospital of North Norway, for her outstanding guidance throughout the whole project. She has provided the motivation and support during all phases of the project including planning, protocol development, practical laboratory work, data analysis and the writing process. I would also like to thank my co-supervisors Ganesh Acharya, Professor of Obstetrics & Gynaecology at the Department of Clinical Medicine and Tore Gutteberg, Professor of Medical Microbiology at the Department of Medical Biology, University of Tromsø for their excellent guidance during project planning and manuscript writing.

My appreciations go to all the supportive staff at the Division of Child and Adolescent Health, Department of Medical Genetics, University Hospital of North Norway, Tromsø, Norway for their excellent help during my study. I would also like to thank the employees Hagar Taman and Christopher Fenton at the Microarray Resource Centre, University of Tromsø for the magnificent help during the performance of the aCGH analysis.

I would like to thank my family and friends for their enormous support during my University studies.

Tromsø, November 2011

Summary

Preeclampsia is a clinical syndrome affecting only pregnant women and is defined as new onset of hypertension and proteinuria after 20 weeks of gestation. Approximately, 2-7 % of all pregnant women in the developed world are affected by this condition and it is a major cause of maternal and fetal morbidity and mortality. Despite intensive research the pathophysiology of preeclampsia is not fully understood.

Laeverin is a gene encoding for a membrane bound-cell surface metallopeptidase (MMP) expressed on extravillous trophoblast (EVT). The EVT cells are responsible for remodelling of maternal spiral arteries during placental development. It has been suggested that *laeverin* may be involved in the regulation of invasive EVTs during early human placentation. Gene expression studies using microarrays have documented that *laeverin* is ten fold-up regulated in preeclamptic placentas compared to normal placentas, indicating its possible involvement in the pathophysiology of preeclampsia. Immune fluorescence studies of placental tissue sections from normal and preeclamptic patients documented that in preeclamptic placentas *laeverin* is expressed in cytoplasma, while in normal placentas it is expressed in the cell membrane. The *laeverin* gene was therefore sequenced to detect possible mutations which could be linked to the pathogenesis of preeclampsia.

We found one variant documented exclusively in blood from one preeclamptic patient in exon 7 position g.1459G>A which replaces one Glutamic acid with Lysine. Since it was not present in any normal controls this may be a pathogenic mutation and should be further investigated.

Several studies have indicated a possible fetal contribution to development of preeclampsia. By performing array comparative genomic hybridization (aCGH) analysis which detects deletions and duplications in the whole genome in one single test, we investigated if umbilical cord blood from fetuses of preeclamptic women contained mutations which may be associated with preeclampsia. These results were compared to the results from an aCGH analysis on maternal blood.

Numerous copy number variations were detected and some of them contained genes involved in vesicle transportation within the Golgi apparatus, protein folding, protein trafficking,

immunological processes, maintenance of cell membranes, pregnancy-specific glycoproteins, complement system, tissue invasion and angiogenesis. These genes should be further investigated to examine their potential role in the pathophysiology of preeclampsia.

Table of contents

Abbreviations

aCGH	Array comparative genomic hybridization
BAC	Bacterial artificial chromosomes
bp	Base pair
cDNA	complementary DNA
CGH	Comparative genomic hybridization
CNV	Copy number variation
DGV	Database of Genomic Variants
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
dNK	Decidual natural killer cells
dNTP	Deoxynucleotides
ddNTP	Dideoxynucleotides
EVT	Extra-villous trophoblasts
FISH	Fluorescense *in situ* hybridization
GWLS	Genome-wide linkage screens
HELLP	Hemolysis elevated liver enzymes low platelet count
HLA	Human leukocyte antigen
IUGR	Intrauterine growth restriction
LVRN	Laeverin
Mb	Mega bases
MLPA	Multiplex ligation-dependent probe amplification
MMP	Matrix metalloproteinases
mRNA	Messenger RNA
PCR	Polymerase chain reaction
SNP	Singe nucleotide polymorphism
STC	Sample tracking control
qPCR	quantitative PCR

1.0 Introduction

1.1 Genetics

Genetics involves the study of inheritance in organisms. The first genetic study was published by Gregor Mendel in 1865. He crossed strains of garden peas with different traits and studied their offspring`s phenotype distribution pattern. The distribution of the traits followed a certain inheritance pattern, now known as the Laws of Mendel. At the beginning of the 19th century evidences was found indicating that deoxyribonucleic acid (DNA) is the genetic material in organisms, and that genes are the basic elements of inheritance (1). In humans the genetic material is distributed on two genomes, the nuclear and the mitochondrial.

The DNA is made up of long linear stretches of nucleotides. A nucleotide consists of a phosphate group, a nitrogen base and a sugar called deoxyribose. There are four different types of nucleotides; Guanine, Cytosine, Adenine and Thymine. The only difference between these four nucleotides is the structure of the nitrogen base. The four nucleotides are connected to each other by sugar-phosphate bonds which connects the hydroxyl group on the 3`carbon with the phosphate group on the 5`carbon of another nucleotide, making a long linear chain of DNA. Human DNA is double stranded, consisting of two linear chains of DNA paired together, forming a double helix with hydrogen bonds between the nitrogen bases in the nucleotides; Guanine pairs with Cytosine, and Thymine with Adenine (2).

The DNA helixes are packed with a type of proteins called histones, and two DNA helixes surrounded by histones constitute a chromosome. There are five different types of histones, H1, H2A, H2B, H3, and H4 (2). Positively charged amino acids in the histones form electrostatic interactions with negatively charged phosphate groups in the DNA molecules and are essential for the molecule`s structure. The human genome consists of twenty-three pairs of chromosomes. The two chromosomes which form a chromosome pair are called homologous chromosomes. Since humans have two sets of each chromosome, our genome is said to be diploid. Twenty-three chromosomes are inherited from an individual’s mother and twenty-three from the father. Chromosomes 1-22 are called autosomes and are similar for both sexes. X and Y are the sex determining chromosomes. Men have one X- chromosome and one Y-chromosome, whereas women have two X- chromosomes. The different chromosomes vary in

size and amount of genes, but they all consist of one p arm (petite) and one q arm (queue), separated by a centromere (2).

DNA is the basis for production of proteins, and is divided into coding sequences, repetitive elements and none- coding elements. The coding part of the double helix is made up of genes that code for amino acids and is essential in protein production. The genes are divided into exons and introns. Exons are nucleotide sequences which code for amino acids. Three nucleotides together form one codon which encodes for one specific amino acid. When genes are expressed, RNA polymerase synthesizes a complementary strand from a template strand in the gene in a process called transcription and forms pre-mRNA. Pre-mRNA consists of both introns and exons and is immature. Mature mRNA is formed when introns are cut out of the pre-mRNA. The mature mRNA is translated into polypeptide chains in a translation process where the amino acids encoded for are bound together by peptide bonds and forms long linear polypeptide chains which are further processed into proteins (2).

1.2 Mutations

A mutation is a permanent alteration in the DNA sequence and may be pathogenic or benign depending on the mutation`s genome position and type of mutation. Large mutations may involve whole chromosomes or parts of chromosomes resulting in a decreased or increased production of several proteins and may cause severe disease. However, pathogenic mutations could also originate from one single mutated nucleotide. Most of the pathogenic mutations are detected in exons, an exception from this is mutations localised in splice sites where introns are cut out of mRNA. There are several types of mutations and they are grouped into gene mutations, chromosome mutations and genome mutations (3).

1.2.1 Gene mutation

A gene mutation involves substitution, deletion or addition of one or a few base pairs (bp) in a gene and may result in alterations of codons. Mutations causing replacement of one codon with another coding for the same amino acid has no effect and are called silent mutations. If the codon is replaced by a codon coding for a different amino acid the mutation is a missense

mutation. Gene mutations could also alter the reading frame and form new codons which change the amino acid composite in the protein. This type of mutation is called frameshift mutation. In most cases the outcome of an altered reading frame is a truncated (non-functional) protein. Gene mutations may also result in the formation of a premature stop codon and lead to production of a truncated protein. This type of mutation is called a nonsense mutation (3).

1.2.2 Genome mutation

A genome mutation is a mutation altering the number of chromosomes. The human genome normally constitutes of forty six chromosomes. Aberrations in the number of chromosomes could involve the whole chromosome set, for instance sixty nine chromosomes instead of forty six (triploid), or the number of individual chromosomes. Some of these syndromes are viable eg. Klinefelter syndrome (47, XXY), Triple X syndrome, (47, XXX), Turner syndrome (45, X), and Down syndrome (trisomy 21). Edward syndrome (trisomy 18) and Patau syndrome (trisomy 13) are lethal genetic syndromes and individuals with these syndromes die in utero or within a short period after birth (3).

1.2.3 Chromosome mutation

A chromosome mutation is a mutation affecting a part of the chromosome and is large enough to be detected using a light microscope. These are classified as duplications, deletions, inversions, and translocations. Deletions involve loss of a chromosome segment while duplications results in gains of an extra part of a chromosome (3). In inversions one segment is cut out of a chromosome due to two chromosomal breaks on the same chromosome and is inverted before it is re-inserted between the two breaks (3). A translocation is a mutation where two chromosomes exchange segments, and usually involves non-homologous chromosomes (3). Several syndromes are caused by chromosome mutations, for example Charcot Marie tooth syndrome (dup 17p12) and Cri du Chat syndrome (del 5p15).

1.3 Inheritance patterns

A genetic disease is caused by mutations in the DNA and may be inherited or acquired. Mutations in germ cells are inherited and are passed through generations, while mutations in somatic cells arise *de novo* and are not inherited. Genetic diseases caused by a mutation in one single gene are called single gene disorders. Some genetic diseases are dependent on gene-gene interactions and gene-environmental interactions to cause diseases and are complex genetic disorders (4). Many different inheritance patterns exist and some of these are presented below.

1.3.1 Autosomal dominant inheritance

An autosomal dominant disease originates from a mutation in one single gene on an autosome, and one mutated copy is enough to cause disease (3). Each offspring of a parent who has the mutation has a 50 % chance for inheriting the mutated gene and developing the disease. In most cases the disease is represented in every generation, and every affected individual has an affected parent (3).

1.3.2 Autosomal recessive inheritance

An autosomal recessive disorder is caused by a mutation in one single gene on both of the autosomes. Two copies of the mutated gene are necessary for manifestation of the disorder, and therefore both parents of an affected individual are carriers of the disease (3). In most cases the parents are unaffected by the disease and may not be aware of their carrier status. Each offspring has a 25 % chance for inheriting the disease (3).

1.3.3 X linked – dominant inheritance

An X-linked dominant disease is developed from a mutation in one gene on the X chromosome. Both genders are affected, but characteristic for this inheritance pattern is that the boys are more severely affected than the girls (3). Girls have two copies of the X-chromosome and in each cell only one of them is expressed. In some cells the mutated X-chromosome is inactivated and in others the normal, they therefore produce some of the

proteins encoded for by the normal X. Boys have only one copy of the X- chromosome and if they inherit a mutated X, it is expressed in all cells causing more severe disease. Sons of a man with an X-linked mutation do not inherit the mutated gene while all his daughters do. Both sons and daughters of an affected female have a 50% risk for inheriting the mutated X (3).

1.3.4 X- linked recessive inheritance

In X- linked recessive disorders boys are more often affected than girls. Boys have only one copy of the gene and will therefore develop disease if they inherit the mutated copy from their mother. Girls are only affected if they inherit one mutated copy from each parent. In cases where the father has the mutated gene, his sons are unaffected and his daughters are carriers. If the mother carries the mutated gene her sons have a 50 % risk of being affected and her daughters have a 50 % risk of being carriers (3).

1.4 Detection of mutations

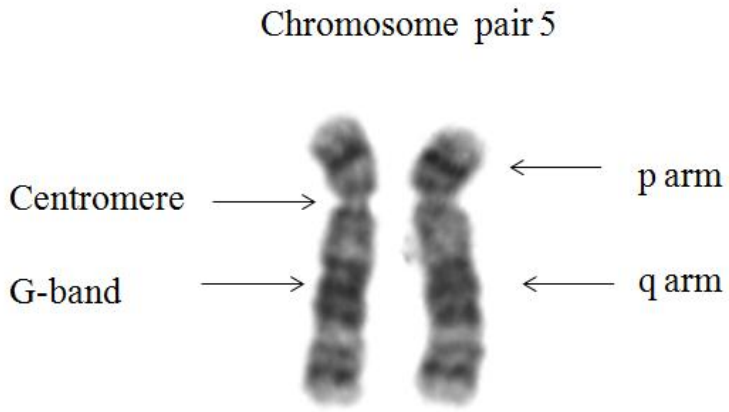
During the last decades there has been a significant expansion in the number of methods used in medical genetics to detect mutations.

Banding techniques

Cytogenetic is the study of chromosomes. Chromosomal banding techniques developed in the 1970s were the first methods for detection of chromosomal aberrations (5). There are several types of banding techniques. Common for them is the colouring of chromosomes in different patterns to detect aberrations (3). The methods identify chromosomes, detect aneuploidy and large aberrations such as deletions, duplications and translocations (5). Chromosomal banding techniques can detect aberrations larger than 3-5 megabases (Mb) (6).

One banding technique is G (Giemsa)-banding where the chromosomes are stained into dark and light bands which divide the chromosome into different regions (figure 1) (3). The dark and light bands are labelled with different letters and numbers. Bands on the long (q) arm are

labelled q followed by a number indicating their chromosomal position. On the short (p) arm the bands are labelled p followed by their number position. The numbers following the letter p or q are counted outwards from the centromere, meaning that both ends of the chromosome have the highest number and the centromere the lowest (7). The specific chromosome bands are used to pinpoint the localisation of the mutation in the genome (7).



Figur 1:. Illustration of chromosome pair 5, after G-banding. Staining with Giemsa creates a pattern of light and dark bands which divides the chromosome into different regions and pinpoints the chromosome location for detected mutations. The centromere represents the thinnest area on the chromosomes. The p arm is positioned above the centromere and the q arm below. The picture is modified from a karyogram made by Mona Nystad, Divison of Child and Adolescent Health, Department of Medical Genetics, University Hospital of North Norway, Tromsø, Norway.

The whole set of one individual`s chromosomes is referred to as a karyotype. Normal males have 46, XY as karyotype, while normal females have 46, XX. Chromosomes are often displayed in a picture with pairs of homologous chromosomes beside each other, called a karyogram (figure 2). Each chromosome pair has distinct features that make it possible to distinguish them from each other (7).

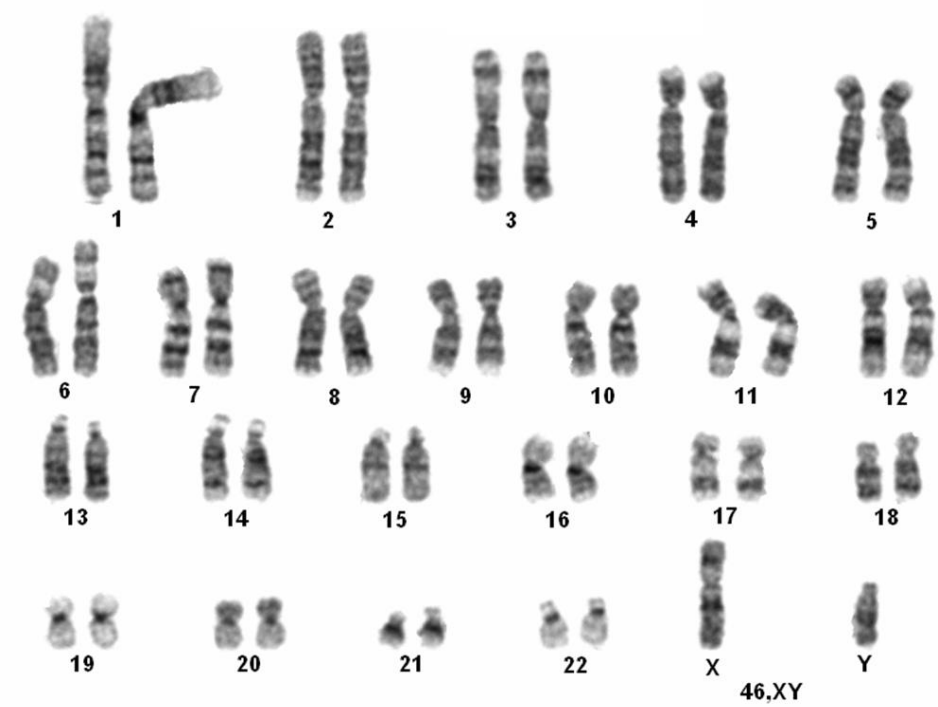


Figure 2: This karyogram displays the karyotype of a normal male. The homologous chromosomes are grouped together and positioned based on their chromosome number. All chromosomes have distinct features which distinguish them from each other. The karyogram was made by Mona Nystad, Divison of Child and Adolescent Health, Department of Medical Genetics, University Hospital of North Norway, Tromsø, Norway.

Owing to the fact that chromosomal banding techniques only detect aberrations larger than 3-5 Mb, there was a need for development of methods detecting smaller mutations. Fluorescence *in situ* hybridization (FISH) analysis and multiplex-ligation dependent probe amplification (MLPA) analysis are examples of higher resolution methods. These methods are considered to be a form of bridge between cytogenetics and molecular genetics. FISH and MLPA-methods detect microdeletions and microduplications, which are mutations of less than 2 Mb (5, 8). Their limitation is, however, that they require knowledge of the disease causing mutation`s genome position (8, 9). Furthermore, since only small parts of the genome are investigated in one test, it can be very time-consuming to diagnose patients with these methods.

During the last decade several methods for the detection of mutations in the whole genome in one single test have been developed. Whole genome sequencing, comparative genomic hybridization (CGH), array comparative genomic hybridization (aCGH) and single nucleotide polymorphism (SNP) arrays are all methods which screen the whole genome for mutations in one single test. Implementation of these methods has resulted in detection of many previously

unknown pathogenic mutations. Disadvantages with the whole genome approach are new ethical problems following the huge amount of information provided by these methods, and troubles with interpretation of the clinical importance of all the detected mutations.

1.5 Copy number variation

Many regions in the human genome have variations in number of copies (CNVs) of a specific sequence as a result of deletions or duplications of this sequence. Copy number variations could be inherited or arise *de novo (10)*. According to Choy *et al*, a *CNV is defined as stretches of DNA larger than 1000 base pairs which normally are found only once on each chromosome in one individual, but can be duplicated or triplicated in some individuals resulting in a variation in number of copies between different individuals* (10). About 12 % of the human genome consists of CNVs, and over 41 % of these overlap with known genes (10).

CNVs of specific regions on human chromosomes are often involved in the development of human diseases (8, 11) and are increasingly found to be associated with risk for development of several diseases, such as neurological disorders, several types of cancer, autoimmune disorders and several syndromes (table 1) (10). Standard chromosome analysis, CGH, FISH, MLPA, Bacterial artificial chromosomes (BAC) arrays, Southern blot and quantitative PCR (qPCR) are all analyses used to detect CNVs (8, 11). CGH has a resolution of 5-10 Mb and is not able to detect smaller CNVs than 5Mb (12). As a consequence array comparative genomic hybridization was developed which detects CNVs down to 5kb (13).

Table 1: CNVs associated with human diseases, modified from reference (10).

Disorder	CNV	Gene	Effect	Risk associated
Infectious Disease				
HIV-1/AIDS susceptibility	Common	*CCL3L1*	Dosage	Low CNV
Autoimmune Disorder				
Systemic lupus erythematosis (SLE)	Common	*FCGR38*	Dosage	Low CNV
Psoriasis	Common	*DEFB*	Dosage	High CNV
Crohn`s Disease	Common	*HBD-2*	Dosage	Low CNV
Neurological Disorders				
Autism Spectrum Disorders	unknown	Multiple	Unknown	De novo CNVs; Multiple CNVs
Parkinson`s disease	Rare	*SA/CA*	Dosage	Duplication/Triplication
Bipolar disorder	Rare	*GSKSb*	Dosage	Deletions and duplications
Schizoprenia	Rare	Multiple	Positional	Deletions and duplications; de novo CNVs
Cancers				
Breast cancer	Rare	*MTTUS1* (exon 4)	Positional	Exon deletion (decrease risk)
Prostate cancer	Common	*UGT2B17*	Positional	Gene deletion
Neuroblastoma	Common	*NBPF23*	Dosage	Deletions and duplications

1.6 Array comparative genomic hybridization

Array comparative genomic hybridization is a method for detection of CNVs and is based on CGH developed by Kallinomi *et al* in 1992 (14). aCGH identifies submicroscopic CNVs in the whole genome at high resolution in one single test (12). The aim of performing aCGH is to detect pathogenic chromosomal aberrations or CNVs (12).

aCGH is based on hybridization of differently labelled DNA from a patient and a normal control added to a microscope slide containing probes. The amount of hybridized DNA is measured by fluorescent signals detected by a scanner. These signals are then converted into fluorescence ratio profiles (figure 3), identifying deleted and duplicated areas in the patient`s genome.

Principle

For the detection of CNVs in the patient's genome the patient`s DNA is compared to DNA from a normal control. This is accomplished by labelling an equal amount of genomic DNA from the patient (green) and normal control (red) with different fluorescent colours. The differently labelled DNA are mixed together and added to probes on a microscope slide (10), where they compete for hybridization to the probes (14). After hybridization the fluorescent signals are detected by a laser scanner (14). If the patient`s DNA lacks CNVs, there will be an equal amount of both fluorescent signals and the colour on the array is yellow. If the patient`s DNA contains duplications, more of the patient`s DNA than the normal control will hybridize to the probes, creating a higher intensity of the green fluorescent signal and this area on the array will be green. In cases where patient`s DNA contains deletions, less of the patient`s DNA will hybridize to the probes compared to the normal control, producing a stronger red fluorescent signal. The colour of this location on the array will be red (14). A computer software is used to convert the fluorescent signals into a fluorescent ratio profile (figure 3), where duplicated and deleted regions are presented. The Y axis presents the log2 ratio which is a measure of the size of the CNV (number of copies) while the X axis documents its chromosomal position. Genome positions with no detected CNVs are contained at the zero line for the log2 ratio. Duplications are localized above the zero line and deletions below. Mutations with large copy number variations are located further from the zero line than small mutations. Duplications of one copy number result in a log2ratio around 0.5, while deletions of one copy number result in a log2ratio around - 0.5

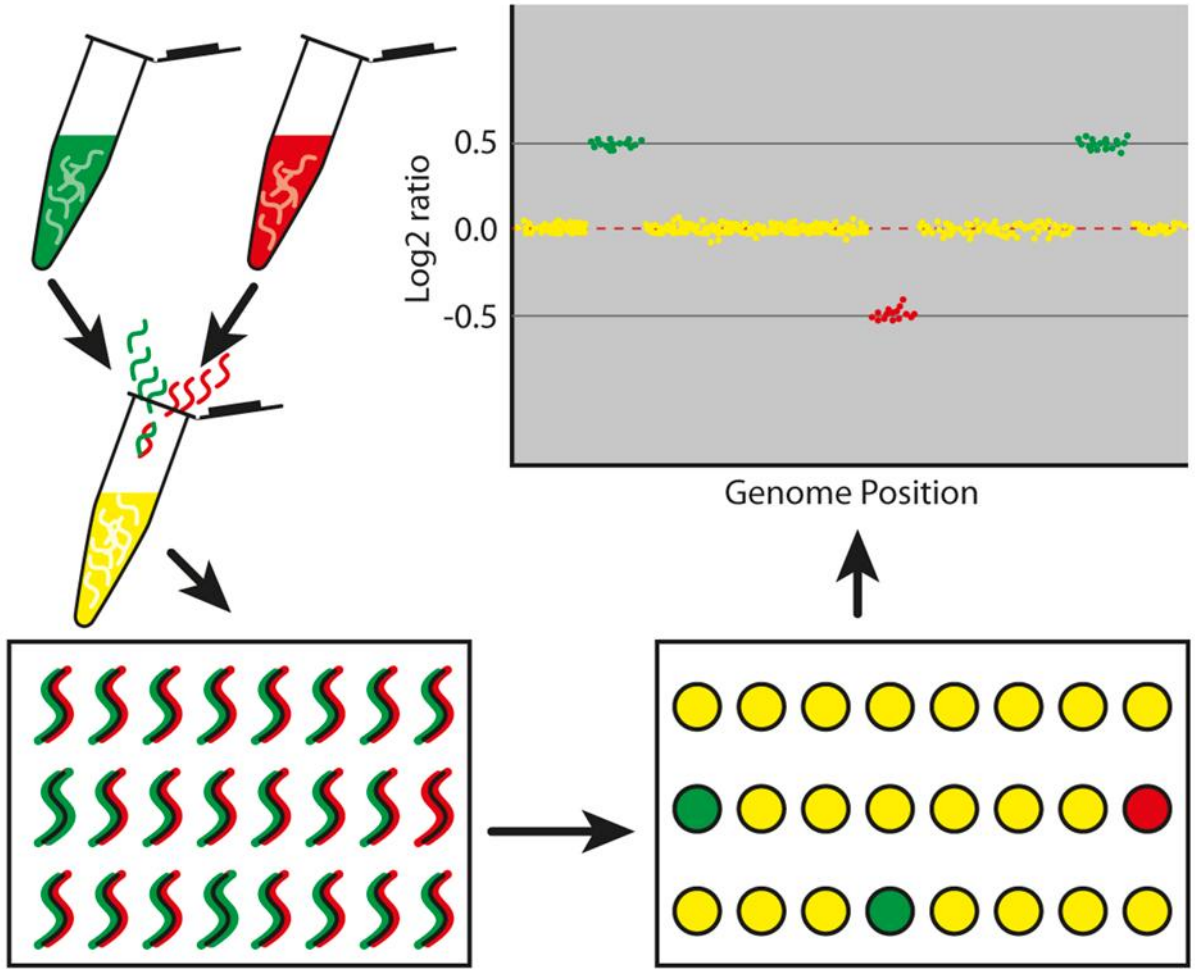


Figure 3: Illustration of the array CGH principle. The patient`s DNA is labelled with green fluorescence and the normal control with red fluorescence. The coloured DNAs are mixed and hybridized to probes on the array. After hybridization the array is scanned and the light intensities are analyzed using specialized software. Areas in the genome which are deleted in the patient`s DNA are represented as red circles, while duplicated areas are represented by green circles. The software converts the fluorescent signals into a fluorescence ratio profile where the log2 ratio on the Y axis indicates the size of the CNV. The X axis shows the detected CNVs position in the genome.

The number of CNVs detected by aCGH depends on the resolution and type of array. The resolution is dependent on the number of BAC or oligonucleotides on the array. The higher the number of probes better the resolution (5). Targeted arrays focus on known deletion and duplication in the genome (12, 15) ,while tiling arrays screen the whole genome for CNVs (12, 16).

Several studies have indicated that when aCGH is performed after obtaining a normal karyotype analysis, the diagnostic yield is increased by 8-17% (12). However aCGH only detects copy number variations, and therefore polyploidies, balanced translocations and inversions are not detected with this method (17). One disadvantage of aCGH is that CNVs of

unknown clinical significance are often detected when the whole genome is screened. These findings may be difficult to interpret and therefore guidelines for evaluation of the aCGH results have been developed (figure 4) (12). Detection of known pathogenic CNVs should always be verified by another independent method. When CNVs of unknown clinical significance are detected the Database of Genomic Variants (DGV) should be screened. The DGV is a database containing information on known structural variants in healthy individuals. If the detected CNV is not documented in DGV or found in healthy parents of the tested individual, it is considered to be a potential pathogenic mutation and should be verified by another method (12). FISH and qPCR are often used to verify aCGH results (12). If the detected CNV is present in the DGV or found in healthy parents of the tested individual it is considered to be benign.

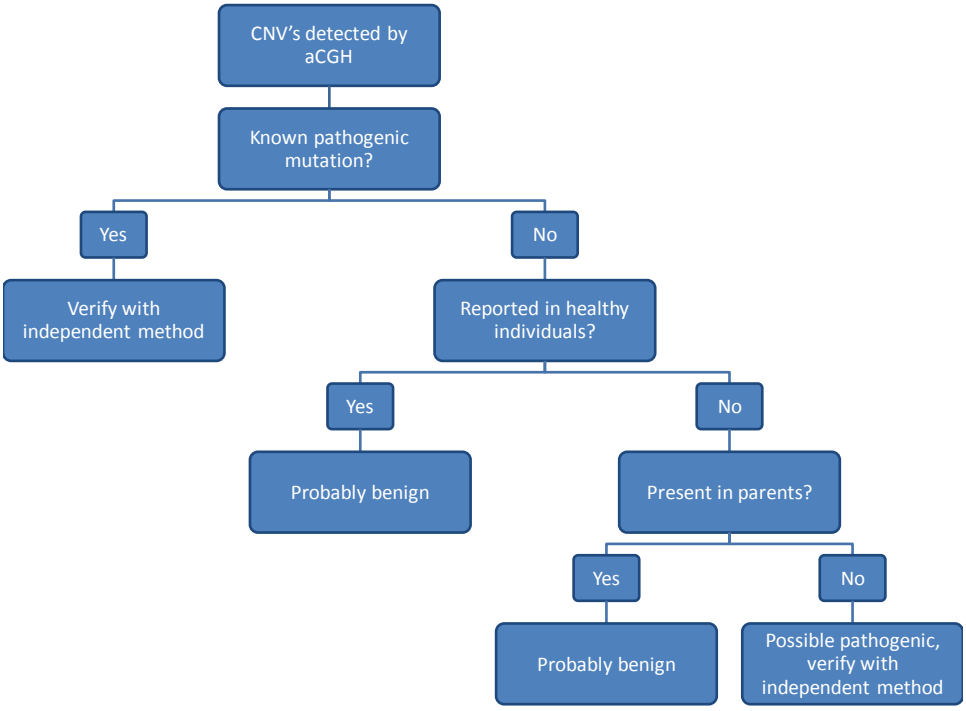


Figure 4: Flowchart representing the guidelines for the interpretation of aCGH results. Modified from reference (12).

CNVs are increasingly linked to several diseases and it is possible that CNVs in the fetal genome contribute to the pathophysiology of preeclampsia.

1.7 The Placenta

The placenta is a temporarily organ which develops and functions during pregnancy. Its function is to support normal growth and development of the fetus (figure 5A) (18).

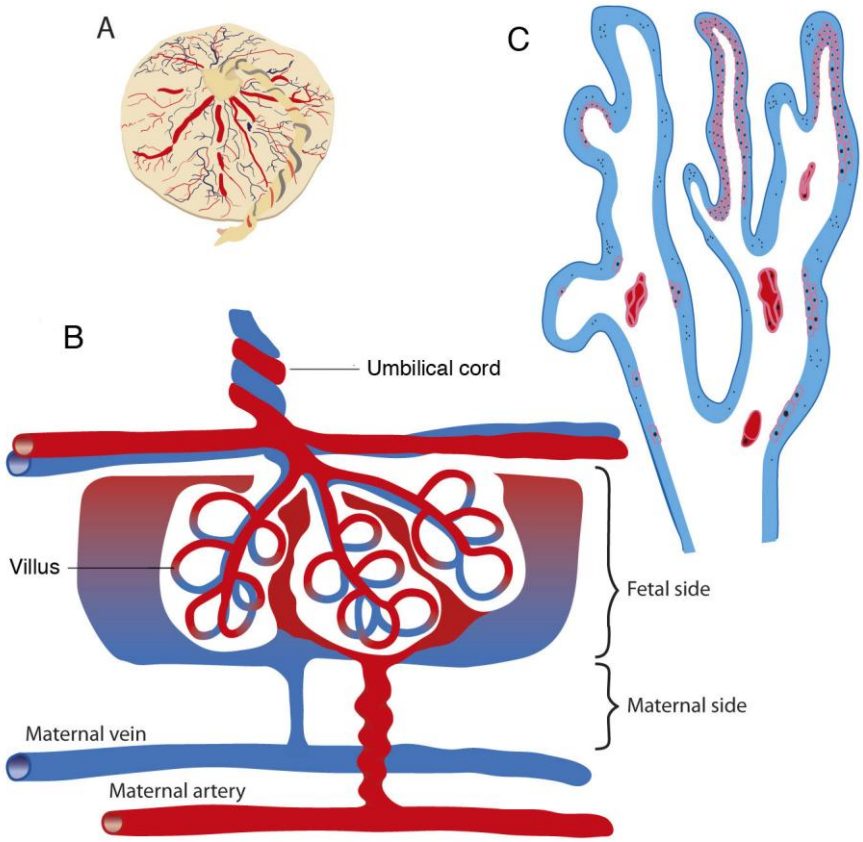
1.7.1 Development and morphology

The placenta (chorion) develops from a fertilized zygote which undergoes several mitotic cell divisions and differentiates into different cell types (19). After a certain number of mitotic cell divisions it becomes a blastocyst and implants into the endometrium in the uterine wall. The formation of the placenta starts with a reaction called the decidual reaction in the uterine wall. Here the stroma cells adjacent to the implanted blastocyst differentiate into metabolically active decidual cells (19). After implantation the trophoblast cells proliferate and differentiate into two pathways, the extravillous and villus pathway (18). The trophoblasts in the villus pathway are responsible for the creation of the fetal part of the placenta, the villus tree formation (figure 5B). Extravillous trophoblasts (EVT) connects the placenta to the maternal circulatory system by remodelling the maternal spiral arteries and develops the utero-placental circulatory system (18, 19). The spiral arteries are remodelled by invasive EVT cells that invade and replace the muscular layer of these vessels and makes the spiral arteries unresponsive to maternal vasomotor control (18, 20).

The villus tree consists of a large amount of villi (figure 5C). The villus consist of a fetal blood vessel surrounded by mesenchymal tissue, and an inner and outer layer of trophoblasts (21). The mononucleated cytotrophoblasts constitute the inner layer (illustrated in pink in figure 5C) of trophoblasts and are formed by invading cytotrophoblasts in the uterine wall. Non-migratory villous cytotrophoblasts fuses together and forms a growing layer of multinucleated syncytiotrophoblasts (illustrated in blue on figure 5C) which constitutes the outer layer and gradually covers the whole surface of the placenta (18, 21).

The umbilical cord connects the placenta to the fetus. It consists of two arteries carrying deoxygenated blood and waste substances from the fetal circulation to the placental villi and into the maternal circulation. A single umbilical vein carries oxygenated blood from the placenta to the fetus. (figure 5B) (18). The villi in the villus tree project into the intervillous

space which separates the maternal side of the placenta from the fetal side (figure 5B). Exchange of nutrients gas and waste products happens in the intervillous space. Here spiral arteries soak the intervillous space with maternal blood and substances from maternal blood enter fetal blood by diffusion through the villi (18).



Figur 5: A: Illustration of a placenta with the umbilical cord. **B:** The villus tree formation on the fetal side of the placenta. The intervillous space surrounds the villus tree which is soaked in maternal blood. The villi in the villus tree are illustrated in red and blue circles surrounded by a white space which represents the placental membrane that separates fetal blood from maternal blood. **C:** The villi structure with different cells. Cytotrophoblasts (pink cell layer) syncytiotrophoblasts (blue) and capillaries (red).

In the first trimester of pregnancy many spiral arteries are totally occluded as a result of trophoblast plugging causing a reduced blood flow to the placenta. The fetus receives a reduced amount of oxygen and is protected from excessive amounts of oxygen which could be damaging during critical stages in early development (21, 22). Between 8 and 12 weeks of gestation there is a significant increase in blood flow and the occlusions caused by trophoblast plugs are gradually removed (22).

1.7.2 Function of the placenta

The placenta has several functions and its main functions can be categorized into transport, metabolism, defence and endocrine function (18).

The placenta supplies the fetus with nutrients (oxygen, water, lipids, carbohydrates, amino acids, vitamins and minerals) from the maternal blood and eliminates waste substances (carbon dioxide, urea etc) from the fetal blood (18). It is also able to metabolize substances and release metabolic products into the maternal and fetal circulations (18). Supply of antibodies from the maternal blood circulation to the fetus helps to protect it from infections caused by virus and bacteria. Hormones affecting pregnancy (metabolism and growth) are produced and secreted from the placenta to the fetal and maternal circulations (18).

1.8 Preeclampsia

Preeclampsia is a clinical syndrome affecting only pregnant women and is defined as a new onset of hypertension and proteinuria in pregnant women after 20 weeks of gestation (23). In addition to hypertension and proteinuria systemic endothelial cell activation and inflammatory response are characteristic (21). Preeclampsia complicates 2-7 % of all pregnancies in healthy nulliparous women in developed countries and the incidence might be even higher in the developing countries (24). For women with previous pregnancies affected by preeclampsia the prevalence rate is approximately 18% and 14% for twin pregnancies (24).

Preeclampsia is a major cause of maternal and perinatal morbidity and mortality and can only develop in the presence of a placenta. The only cure for preeclampsia is therefore removal of the placenta by delivering the fetus (21). There are two types of preeclampsia, early onset and late onset. Early onset is the most severe and is associated with abnormal villus and vascular structure. Onset is before 34 weeks of gestation and the fetus is often affected by growth restriction (21). Furthermore the fetus can be affected by nutritional and respiratory insufficiency, asphyxia and death (25). Late onset is the most common form of preeclampsia and constitutes approximately 80% of all cases around the world. It usually has normal placental morphology and is not associated with fetal growth restriction (21). Women affected by severe preeclampsia may develop haemolysis, elevated liver enzymes, and low platelet

count syndrome (HELLP). Up to 20% of women with severe preeclampsia develop HELLP syndrome (26).

The pathophysiology of preeclampsia is still not fully understood, but it involves a complicated network of maternal and fetal interacting factors (21).

1.8.1 Development of preeclampsia

Development of preeclampsia is categorised into two stages, preclinical and clinical. The preclinical stage is before the maternal symptoms start to develop, and the clinical is after (25). Poor placentation is typical for the preclinical stage and is caused by an aberrant development of the early placenta and maternal blood supply (25). Defects in trophoblast differentiation may cause a reduced amount of extravillous trophoblasts, resulting in a reduction of the number of remodelled arteries and their depth of remodelling (21). Inadequate remodelling leads to reduced blood flow to the placenta in second and third trimester, inducing a gradually more and more hypoxic placenta (25, 27). The hypoxic placenta releases factors into the maternal circulation which are believed to induce the maternal symptoms like hypertension, proteinuria, clotting and liver dysfunction (25, 28, 29). These factors originate from an inflammatory response with endothelial dysfunction as a main component (25, 28).

Involvement of the immune system in the development of preeclampsia has been shown by several research groups. Decidual natural killer cells (dNK) and decidual macrophages are essential cell types implicated in immunological theories. The dNK cells infiltrate the decidualized uterus before implantation and remain in large quantities throughout the first trimester. Evidence suggests that they play a direct role in the remodelling of spiral arteries and tropohoblast invasion. There is however limited data on how dNK cells are involved in preeclampsia and research performed in this field provides conflicting results (21). One theory presented is that poor placentation and preeclampsia is less common in individuals where trophoblast stimulates dNK cells strongly by binding to their receptors (25). Decidual macrophages are believed to have the ability to disrupt vascular smooth muscles in spiral arteries before invasion of extravillous trophoblast, and may therefore influence the remodelling of spiral arteries and development of preeclampsia (21).

1.9 Genetics of preeclampsia

Despite extensive research and a large number of published studies the aetiology of preeclampsia remains unknown. Studies have revealed that there is a clear familial predisposition and a high recurrence risk for severe preeclampsia in pregnant women with earlier pregnancies affected by preeclampsia. Both of these factors are indications for a genetic contribution (23, 29-31). However, so far no specific genes have been identified that cause preeclampsia. Preeclampsia is in most cases considered to be a complex genetic disorder and conflicting results seems to be characteristic for the genetic studies performed (30).

Many factors contribute to the difficulties in establishing the aetiology of preeclampsia. Diagnosis of preeclampsia is often based on different thresholds for blood pressure and proteinuria. Several hypertensive disorders linked to pregnancies can be mistaken for being preeclampsia because of similar symptoms (32). Severity and involvements of different organs are often also highly variable between individuals, and could result from different molecular mechanisms (30). Early and late onset preeclampsia have different clinical characteristics and many studies lack distinction between these two types, that potentially could originate from different molecular mechanisms. Another problem is the difficulty in comparing and interpreting the findings of studies performed on placental samples obtained at different stages of development. Gene expression profiles from placentas differ between different gestational ages and are a source of conflicting results when they are compared to each other. Both the paternal and maternal genotypes may be responsible for development of preeclampsia (30, 33). Furthermore, fetal and maternal genomes may have genes that by themselves or together may be responsible for preeclampsia (32, 33). The paternal genotype should therefore also be investigated. Genetic variation between populations are also reasons for contradicting results since different molecular mechanisms may be present in different ethnic populations (33). The majority of studies have focused on maternal genes and therefore the Genetics of Pre-eclampsia Consortium has highlighted the need for studies performed on all genotypes (29, 33).

1.9.1 Inheritance pattern

Twin studies

Twin studies have been performed to investigate the relationship between genetics and environmental factors in preeclampsia (33). These studies show conflicting results (29). The first study reported a low concordance of development of preeclampsia between monozygotic twins indicating a low heritability by maternal genes. A more recent study performed based on material from the large Swedish Twin, medical Birth and multigeneration register reported the heritability of preeclampsia to be around 55 % with contribution from both maternal and fetal genes(33). Another twin study performed in Sweden reported the preeclampsia penetrance between twin sisters to be less than 50 %, representing a diversity in the inheritance pattern (33).

Paternal contribution

There are several indications for paternal contribution in the development of preeclampsia. Research documents that women with subsequent pregnancies with the same partner have a lower risk for developing preeclampsia, while women with subsequent pregnancies fathered by different partners have a higher risk (21, 31). Long term exposure to paternal antigens may induce a form for immunological memory (21). For women with preeclampsia in their first pregnancy, a change of partner may lower the risk for preeclampsia since the new partner could have more compatible HLA C molecules on their trophoblast cells and receptors (21). A higher rate of preeclampsia is also documented in pregnant women with a partner who himself was born of a woman with preeclampsia (27, 29, 33, 34). Another evidence for paternal contribution is that men who earlier have fathered pregnancies with preeclampsia, have a higher risk of fathering a new pregnancy with preeclampsia with a new partner (29).

Complex genetic disorder

Based on previous research, preeclampsia is considered to be a complex genetic disorder in the general population with some exceptions following the Mendelian inheritance pattern (29, 33). A complex genetic disorder is an inherited disease which is not caused by a single gene and that deviates from the Mendelian inheritance pattern (4). In such cases relatives who share

the same genotype may have different phenotypes depending on gene-gene interactions and gene-environmental interactions (4). Since most of the preeclampsia cases are considered to have a complex genetic cause, it is likely that no single gene or variant will account for all the occurrences of preeclampsia (33).

1.9.2 Genetic methods applied to study preeclampsia

Research studies on preeclampsia are performed on different population groups and with varying methods, such as the candidate gene approach, genome wide linkage studies, imprinting studies, and expression profiles studies (34).

Candidate gene approach

The candidate gene approach is a widely used method in preeclampsia studies and involves selection of one single gene for investigation based on already known information (33). Until now more than 70 single genes categorized into different groups based on pathological mechanisms such as oxidative stress, lipid metabolism, thrombophilia immunogenetics, and endothelial injury have been investigated (33). All these studies have provided contradicting results, and no genes have so far been universally accepted as the cause of preeclampsia (33). A selection of the most predominant functional candidate genes are presented in table 2.

Table 2: Presentation of the most predominant functional candidate genes. Modified from reference (33).

Pathophysiological mechanism group	**Gene name**	**Gene symbol**
Trombophilia	*Factor V Leiden*	*F5*
	Methylenetetrahydrofolate	*MTHFR*
	Protrombin	*F2*
	Plasminogen activator factor-1	*SERPINE1*
	Integrin glycoprotein IIIa	*GPIIIA*
Endothelial function	*Endothelial nitric oxide synthase 3*	*eNOS3*
	Vascular endothelial growth factor receptor 1	*VEGFR1*
	Vascular endothelial growth factor	*VEGF*
Vasoactive proteins	*Angiotensinogen*	*AGT*
	Angiotensin converting enzyme	*ACE*
Oxidative stress and lipid metabolism	*Apolipoprotein E*	*APOE*
	Microsomal epoxide hydrolase	*EPHX*
	GluthationeS-transferse	*GST*
Immunogenetics	*Tumor necrosis factor α*	*TNF*
	Interleukin 10	*IL10*

Genome- wide linkage screens

Genome-wide linkage screens (GWLS) is a method where linkage between genetic markers with known position distributed throughout the whole genome are investigated. Some linkage studies have been performed on preeclampsia, where samples from preeclamptic women and their relatives were compared to each other to find possible linkage (33). One disadvantage with this approach is that only relatively large areas which can contain hundreds of genes are identified (33). If potential linkage is found, further analysis needs to be performed to find possible candidate genes in the linkage region.

Arngrimsson *et al* (1999) found a maternal susceptibility locus for preeclampsia on chromosome 2p13 (29, 33, 35) with consistent results found by an Australia/ New Zeland research group (table 3) (29, 30, 36, 37). Extended work was performed to reveal associations between the potential candidate genes in the 2q22-23 locus (33). *Activin A* receptor type IIa (*ACVR2A*) was identified as a strong candidate gene, but different research groups published contradicting results (33). GWLS studies performed in Netherland and Finland found other susceptibility locus (table 3) (37, 38). The Finish research group investigated the *ROCK2* gene in the 2p25 locus further, but with disappointing results. No association with preeclampsia was documented (33).

Table 3: Outline of performed GWLS studies in the field of preeclampsia, modified from reference (33)

Country	Number of families	Chromosome loci
Iceland	124 (343 women)	2p13
Australia/New Zealand	34 (366 women)	2q23 11q23
The Netherlands	38 (332 women)	10q22 22q12
Finland	15 (174 women)	2p25 9p13 4q32 9p11

Genomic Imprinting

Many genes are located in conserved clusters where only one allele is expressed, either the maternal or paternal allele. This condition is called genetic imprinting. Indications for involvement of genetic imprinting with preferential expression of the maternal allele in

development of preeclampsia are documented (29) . Several studies have been performed on one gene found within the locus 10q22.1 (table 4). This gene encodes for the *STOX1* transcription factor which is believed to be involved in the transformation of invasive trophoblasts into non invasive. The data from these studies are contradicting, and further studies are necessary to elucidate its role in preeclampsia (33).

Table 4: Studies performed on one gene encoding the STOX1 transcription factor (39-42)**.**

Research group and year	**Imprinted locus**	**Potential genes**	**Conclusion**
Oudejans *et al* 2004	10q22.1		Downregulated expression in hydratiform molar placentas (only paternal contribution)
Djik *et al* 2005	10q22.1	*STOX1*	Found identical missense mutations in *STOX1* between affected sisters
Iglesias *et al* 2007	10q22	*STOX1*	Detected biallelic expression of *STOX1* in both normal and preeclamptic placentas and found no evidence for an imprinted mechanism.
Rigourd 2008	10q22	*STOX1*	Overproduction of *STOX1* reproduces transcriptional effects of preeclampsia.

Another gene examined for possible imprinting status is the H19 gene which is involved in regulation of growth and development of the embryo, and differentiation of cytotrophoblast cells (43). L. Yu *et al (2009)* documented that biallelic expression of *H19* existed in some cases in early stages of normal pregnancy and changed to monogenic expression around 10 weeks, while preeclamptic placentas had biallelic expression in third trimester. (33, 43). Bourque *et al* (*2010*) found no reduced methylation at ICR1 which is an imprinting region regulating the *H19* gene indicating that there is no biallelic expression in placentas from women with preeclampsia and contradicted Yu`s study (44).

Microarrays and gene expression

Gene expression varies between different cell types and cell stages (4). Several gene expression studies have compared profiles in preeclamptic placentas with normal placentas with conflicting results (4). Sitras *et al* compared his results from a gene expression study on preeclamptic placentas with 16 previously performed microarray studies and found some similarities of differentially expressed genes in women with preeclampsia compared to normal pregnancies. *Leptin, Human chorionic gonadotropin* (*hCG*), *Vascular endothelial growth factor* (*VEGF*), *Insulin- like growth factor 2* (*IGF2*), *Laeverin* and *Matrix metalloproteinases* (*MMPs*) were found to be differentially expressed in more than one study (23).

Several genes found to be either up or down regulated in microarray studies have been further investigated by the candidate gene approach to gather more information of the origin of preeclampsia.

1.9.3 Genes involved in early versus late preeclampsia

Sitras *et al* (2009) documented differences in gene expressions between early onset and late onset preeclampsia in his study. By comparing their gene expression profiles indications for involvement of different pathways in early and late onset of preeclampsia was found (23).

Enquobahrie *et al* (2011) compared gene expression profiles from early pregnancy peripheral blood from preeclamptic patients with gene expression profiles of placental tissue from preeclamptic women at delivery. They found gestational age and tissue specific differences in pathophysiological processes such as vasculature development in early preeclampsia versus hypoxia response in late preeclampsia (45). Numerous other studies have been performed on early and late preeclampsia on both maternal blood and tissue samples from placenta. In these studies inconclusive results are reported. A review based on current and earlier gene expression studies suggests that early preeclampsia is associated with alterations in angiogenesis and immune inflammatory response, while late preeclampsia is associated with alterations in response to hypoxia or oxidative stress and subsequent endothelial dysfunction (45).

1.10 The Laeverin gene

Sitras *et al* (2009) documented that *laeverin* (LVRN) was 10 fold up- regulated in preeclamptic placentas compared to normal controls (23). *Laeverin* encodes for a membrane-bound cell-surface metallopeptidase isolated from chorion lave and is therefore named *laeverin* (46) Matrix metalloproteinases degrade the extracellular matrix and are secreted from extravillous trophoblasts. They are considered to regulate EVT invasion in cooperation with their inhibitors (47). *Fujiwara 2007* reported that the metalloproteinase encoded by *laeverin* is only expressed on EVTs, indicating that it is a specific marker of EVT (47). The

metallopeptidase encoded by *laeverin* have some similarities with the CD13/aminopeptidase N. CD13/aminopeptidase is necessary for endothelial cell invasion (48) and controls the endothelial cells motility by rearranging the cytoskeleton to create filopodia formation (23). Petrovic *et al 2007* performed a study were they documented that inhibition of CD13 activity during the cell recovery phase followed by depletion of cholesterol or trypsinization prevented formation of filodopia and disrupted the distribution of membrane proteins (48).

Furthermore Maruyama *et al* documented that the cDNA from *laeverin* encodes for a protein consisting of 990 amino acid residues containing one motif specific for the M1 family of aminopeptidases (49). According to this study the protein encoded by *laeverin cleaves the N-terminal amino acid of many peptides, ex angiotensin III, kisspeptin-10, and endokinin C* (49, 50) which are substances expressed in large quantities in the placenta. Based on this documentation it is suggested that *laeverin* has an important role in placentation through regulation of important peptides in this process (49). Goto *et al* has recently indicated that *laeverin* is important in maintenance of normal pregnancies in humans (50).

After Sitras *et al* (2009) documented that laeverin was tenfold up- regulated in preeclamptic placentas, Nystad *et al* performed a immunefluorescence study on placental tissue (un-published study). Fluorescence marked antibodies directed against the laeverin protein were applied to tissue sections of placenta to reveal the localisation of the protein within the cell. In normal placenta laeverin was localised to the cell membrane and in preeclamptic placentas laeverin was localized to the cytosol. Further electronmicroscopic studies will reveal the subcellular localisation of the protein. Based on this information it was suggested that the role of *laeverin* in the pathophysiology of preeclamptic placentas needed to be further investigated (23).

1.11 Aims

The aims of this thesis were to:
Sequence the *laeverin* gene in blood samples from patients affected by preeclampsia and normal controls to investigate if there were any pathogenic mutations.

Sequence the laeverin gene in placental tissue sample from one patient with preeclampsia and investigate whether different mutations are present in the placental tissue sample compared to the blood sample from the same patient.

Screen the whole fetal genome for pathogenic deletions or duplications using array aCGH analyses on umbilical cord blood samples obtained from the newborns of preeclamptic women and compare with umbilical cord blood obtained from the newborns of women with healthy pregnancies.

Determine the chromosomal location and genes included in the CNVs identified.

Examine the relevance of the detected gene mutations and CNVs in the pathogenesis of preeclampsia.

2.0 Material and methods

2.1 Patient samples and normal controls

2.1.1 Normal controls

As normal controls ten samples from healthy individuals were used and consisted of EDTA blood from four women (K1-K4) and from six men (M1-M6). DNA was extracted both manually with QIAamp DNA Blood Mini Kit (50) and with the automatic method. The DNA concentration was within the range 16.0 – 42.5 ng/µl. Based on the DNA concentration and the results from NanoDrop quantifications, M1, M4, M6, K3, and K4 were selected for use in gradient PCR while, M6, and K4 were used in sequencing.

2.1.2 Patient samples used for sequencing

Eight samples from women with preeclampsia were sequenced for all exons and parts of introns in the *laeverin* gene, while 20 samples from preeclampsia patients were sequenced for exon 7 and 10. The sample material consisted of DNA extracted from EDTA blood samples from these women. DNA was extracted automatically with Biorobot EZ1, and had concentrations within the range from 43.5– 112.1 ng/µl. All patients had severe preeclampsia, but in addition some had HELLP syndrome and foetuses with intrauterine growth restriction (IUGR). Samples from women with only preeclampsia are not presented with sample number. All samples from women with preeclampsia are labelled with P after sample number. Samples 11P, 20P and 21P are from women with preeclampsia and HELLP syndrome. 1P are from a woman with preeclampsia, HELLP and IUGR, while 9P is collected from a woman with preeclampsia and IUGR.

One sample from placental tissue was sequenced for all exons and parts of intron in the *laeverin* gene. This tissue sample was collected from the placenta right after delivery and frozen in liquid nitrogen. DNA extraction was performed by an employee at the Division of Child and Adolescent Health, Department of Medical Genetics, University Hospital of North Norway, Tromsø, Norway with an extraction kit from Qiagen.

2.1.3 Samples from normal pregnancies used for sequencing

Three samples from normal pregnancies were sequenced for all exons of the *laeverin* gene and parts of introns, while 18 samples were sequenced for exon 7 and 10. Furthermore 23 samples were sequenced for only exon 7. The sample material consisted of DNA extracted automatically from EDTA blood samples with Biorobot EZ1. The DNA concentrations were in the range 24.0–142.0 ng/µl. Which samples were sequenced for all exons and which were sequenced for only exons 7 and 10 were randomly chosen since all had high enough concentrations. Some samples were only marked with a sample number while others were marked with an N after sample number or NK. N is a symbol for normal pregnancy and NK for delivery of the foetuses by caesarean-section.

2.1.4 Patient samples used in aCGH analysis

Nineteen samples from umbilical cord blood from preeclamptic pregnancies were selected for array comparative genomic hybridization based on DNA concentrations and purity after NanoDrop quantification. DNA was extracted from EDTA blood samples with Biorobot EZ1, and had concentrations in the range from 32.0-104.1 ng/µl. All samples were from patients with severe preeclampsia. In addition some patients had HELLP syndrome and IUGR. Samples from women with only preeclampsia are not presented with sample number. Sample 9P had IUGR in addition to preeclampsia. Sample 20P, 21P and 30P also had HELLP syndrome

2.1.5 Normal controls used in aCGH analysis

Nineteen samples from umbilical cord blood from normal pregnancies were selected for aCGH analysis on the same basis as the samples from preeclamptic patients. The DNA was extracted from EDTA blood samples with Biorobot EZ1, and had concentrations in the range from 27.0 – 104.5 ng/µl. The samples were labelled with sample number followed by N or NK.

2.2 Manual DNA extraction with QIAamp DNA blood Mini Kit (50)

The DNA from blood obtained from normal controls was extracted with the QIAamp DNA blood Mini Kit (50).

Principle

DNA extraction with QIAamp DNA blood Mini Kit (50) was performed by use of QIAamp mini spin columns. These columns contain a silica membrane which is essential in the extraction process.

The first step in this extraction method is cell lysis, which is performed by addition of protease and buffer to the blood samples extracted. The combination of these two reagents destroys the cell membrane and release the DNA from the cells. After cell lysis the solution containing the DNA is added to QIAamp mini spin columns. DNA is negatively charged and will bind to the silica membrane in the columns under high salt conditions. Addition of wash buffers and several centrifugation steps removes proteins and other cell components which are not negatively charged. These components are not able to bind to the columns and are washed away during centrifugation. For elution of DNA from the silica membrane sterile water is added (51).

Material and reagents

Table 5: Kit used for DNA extraction

Kits	**Producer**
QIAamp DNA blood Mini Kit (50)	Qiagen

Table 6: Equipment used for DNA extraction

Equipment	**Model**	**Producer**
Pipette	Pipet Lite	Ranin
Centrifuge	Centrifuge 5415D	Eppendorf
Centrifuge	Biofuge 13	Heracus sepatech
Vortexer	VF2	Janke og Kunkel, KA laboratorietechnik
Water bath	DC 100	Grant

Method

Twenty µl of Qiagen protease K was added to 1.5 ml microcentrifugation tubes. Then 200 µl blood and 200 µl buffer Al were added, before vortexing for 15 seconds and incubation at 56°C for 10 minutes. After incubation the tubes were centrifuged for a short time to remove prospective droplets on the inside of the lid. Two hundred µl of ethanol was added and vortexed for 15 seconds before a short centrifugation. The tube contents were transferred to QIAamp Mini Spin Columns, placed in 2 ml collection tubes and centrifuged at 6000 X g for one minute. After centrifugation the columns were put into new clean 2 ml collection tubes and the old ones were discarded. Five hundred µl of buffer AW1 was added to the columns and centrifuged at 6000 X g for one minute. Subsequent to centrifugation the columns were placed in new collection tubes and the old ones were discarded. Five hundred µl of buffer AW2 were added followed by centrifugation at 20 000 X g for three minutes. The centrifugation step was repeated two times and the columns were placed in new 2 ml collection tubes while the old ones were discarded. Two hundred µl of sterile water was added to the columns followed by incubation for 5 minutes at room temperature. Subsequent to incubation the columns were centrifuged at 6000 X g for one minute and the contents in the collection tubes were transferred to microcentrifugation tubes and stored.

2.3 Automatic DNA extraction from blood samples

Automatic extraction was performed on all patient samples and samples from normal pregnancies with Biorobot EZ1 and Biorobot M48.

Principle

DNA extraction with Biorobot EZ1 and Biorobot M48 is based on magnetic-particle technology. The magnetic-particles contain silica membranes and binds negatively charged DNA. The first step is cell lysis where cell walls are lysed by addition of guanidine hydrochloride. Guanidine hydrochloride also prepares the DNA molecules for binding to the silica membranes by providing a high salt concentration. Bound DNA molecules are separated from the lysate by a magnet which draws the magnetic particles to it. Proteins and other cell components will remain in the lysate. Additions of two different wash buffers removes the

cell lysate and cleanse the DNA molecules on the magnetic-particles. The first wash buffer removes guanidine chloride from the solution while the second one is a low salt buffer containing ethanol. Elution of the DNA molecules is performed by addition of Tris-EDTA (TE) buffer (52).

Material and reagents

Table 7: Kits used for DNA extraction.

Kits	Producer
EZ1 DNA Blood 350 µl Kit (48)	Qiagen
Sample Prep Plates 42 well, M48 100	Qiagen

Table 8: Equipment used for DNA extraction.

Equipment	Model	Producer
DNA extractor	BioRobot EZ1	Qiagen
DNA extractor	BioRobot M48	Qiagen

Method

The EDTA blood samples were mixed by inversion and vortexed before 350 µl were transferred to new tubes following the extraction kit. DNA extraction was performed with biorobot EZ1 or Biorobot M48 and 200µl was selected as elution volume.

2.3 Quantification of DNA by NanoDrop 2000

All extracted DNA samples were quantified by NanoDrop2000

Principle

NanoDrop 2000 is a spectrophotometer which measure DNA/RNA concentrations at the A260 ratio and proteins at A280. Sample purity are evaluated by the A260/280 ratio (53). The amount of light absorbed by the quantified DNA is measured and computer software calculates the concentration of DNA molecules present based on the amount of light absorbed. Absorption values are measured at several different wavelengths to evaluate the purity and concentration of DNA (54). Measurements at A260 are quantitative for relatively pure DNA molecules. To determine the purity of the DNA molecules the A260 and 280 ratios

are used. A260/A280 ratios of 1.8 -2.0 indicates highly purified DNA (54). Proteins have an absorption peak at 280 nm and will reduce the A260/A280 ratio. Substances with peptide bonds or aromatic moieties such as proteins and phenol absorb light at A230 nm (54).

Material and reagents

Table 9: Reagents used for DNA quantification.

Reagents	Purpose
dH2O	Sample blank
DNA	Quantification

Table 10: Equipment used for DNA quantification.

Equipment	Model	Producer
Pipette	Reference	Eppendorf
NanoDrop	NanoDrop 2000 Spectrophotometer	Thermo Scientific

Method

Before quantification of DNA the NanoDrop instrument was calibrated with water as sample blank. After calibration 1 µl water was measured to control the calibration before 1 µl of the extracted DNA was quantified.

2.4 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was performed on all samples used for sequencing to produce the necessary PCR products.

Principle

PCR is an *in vitro* method for specific amplification of a selected sequence of DNA, into millions of copies. The selected sequence (template) is amplified by DNA polymerase in a thermal cycling process which consists of three different steps, denaturation, annealing and elongation. These steps are repeated 20-30 times to generate millions of copies. Denaturation is the first step and is performed at 94°C, the high temperature separates the double stranded DNA into single stranded DNA. The annealing step follows after denaturation and the temperature is decreased to 45-60°C. At this temperature range primers which are short DNA sequences hybridizes to complementary sequences in the single stranded template. Different

primers have different annealing temperatures based on the length and sequence of the primers. Elongation is the third step and is performed at 72°C where DNA polymerase binds to the free 3 ` end of the primers and produces a new growing chain by incorporation of new nucleotides which are complementary to the template DNA. When these tree steps are repeated 20-30 times, millions of copies of the template sequence are produced (2).

Material and reagents

The primers were received from Mona Nystad and were designed in the primer3 (v 0.4.0) software at http://frodo.wi.mit.edu/. In the primer design the length of the primers, annealing temperatures, G-C contents and complementarities to other sequences were all factors taken into consideration. The primer design was checked by in silico in the UCSC Genome Bioinformatics database at http://genome.ucsc.edu/. The gene sequence of the *laverin* gene used in primer design was retrieved from the NCBI database and the following version and number was used NM_173800.4. The complete gene sequence is presented in appendix C.

Table 11: Primers used for PCR and sequencing. All primers are produced by Sigma-Aldrich. Forward primers are marked with F and reverse primers with R.

Name	Sequence	Product size	Lot. no
LVRN ex1_1F	GGGGTCTGTCTCTTCGAACC		-HA01606306
LVRN ex1_1R	TGGTCGTCACCGCTAGCTC	548	-HA01606307
LVRN ex1_2F	GGACTCAGGGACTTGGAAGC		-HA01606308
LVRN ex1_2R	ACCTGGACACGGTAGTCAGC	637	-HA01606309
LVRN ex2-F	CCAGGCTAACTTACAAGCCATC		-HA01606310
LVRN ex2-R	ACACTGCCTAGAAAAGCTACTGG	324	-HA01606311
LVRN ex3-F	GCTGGGTGACAATTGACTTGA		-HA01606312
LVRN ex3-R	CAGCCATAACCACTGGATGC	315	-HA01606313
LVRN ex4-F	TGCAGCTTCTTCTGGAAAGG		-HA01606314
LVRN ex4-R	ACCAAGGGCCCAAATAACAG	303	-HA01606315
LVRN ex5-F	GCCTTCTCTCCGTTCAGGTT		-HA01606316
LVRN ex5-R	TGAATCTGAGTCTTCCCCATTTT	451	-HA01606317
LVRN ex6-F	TTACTTAACATGCAATTACAACGAA		-HA01606318
LVRN ex6-R	TAGATAAAAAGCCTGGAATTAAAAA	373	-HA01606319
LVRN ex7-F	TTGGCATATTTCAACTCAAAGACA		-HA01606320
LVRN ex7-R	TGGATATTTGCAAGGTGAGAATTA	467	-HA01606321
LVRN ex8_9-F	GCAAAATGCAATCAGGAGCA		-HA01606322

Name	Sequence	Product size	Lot. no
LVRN ex8_9-R	CAAAGCTGTTTTCCTGTGGAAC	441	-HA01606323
LVRN ex10-F	CGAGTTTCTTTTTGGAGATTGC		-HA01606324
LVRN ex10-R	AGCGGGGGTAACGCTTCT	349	-HA01606325
LVRN ex11-F	TGAGTTCTGTGTGCTATTTCATCA		-HA01606326
LVRN ex11-R	CAAGAATGATATTTTAGTTTGAAATGA	295	-HA01606327
LVRN ex12-R	GGGCGACAGAGCAAGACTC	889	-HA01606328
LVRN ex13-F	GAAGGTGCTTTATAGGCAGCAG		-HA01606329
LVRN ex13-R	ATGGCCTGTATAGACTAAATAATCAAA	268	-HA01606330
LVRN ex14-F	TCATGCATTGAAACATGGAA		-HA01606331
LVRN ex14-R	ATGGAACTAAGACTATTACTTTGGAA	358	-HA01606332
LVRN ex15-F	TCTGCCACTTGTCTATCAATGTC		-HA01606333
LVRN ex15-R	GCATATAAAATCTCTCAATGAATAACA	315	-HA01606334
LVRN ex16-F	CCCATCTTTTTATGAAACACATATTC		-HA01606335
LVRN ex16-R	GCAAAACACTCATTCCCACA	295	-HA01606336
LVRN ex17-F	TCCTACTTTTGACCTTTATATCTGTGA		-HA01606337
LVRN ex17-R	CGCAACGTTTTAATACTTACATTAGC	391	-HA01606338
LVRN ex18-F	CCTTAGAACCATGGGATTTTGA		-HA01606339
LVRN ex18-R	CTGAGCTTCCAGTGCCAAAC	291	-HA01606340
LVRN ex19-F	GGGGCCATACACTTGACCTT		-HA01606341
LVRN ex19-R	CCCTGCATTCAGGTGAGAGA	399	-HA01606342
LVRN ex20_1F	CAAACTCATGTTGCTACTTAGCATTTA		-HA01606343
LVRN ex20_1R	TCAGCAAAGAATGTGCTCCT	600	-HA01606344
LVRN ex20_2F	CCACAGAATTTACTTTAAATGTCACG		-HA01606345
LVRN ex20_2R	TTTGACATCATTGTTTTCACTCG	571	-HA01606346
LVRN ex20_3F	AGTGGGATGGACACCCTTTC		-HA01606347
LVRN ex20_3R	CCTGACACTCAGGGAGAATCC	590	-HA01606348
LVRN ex20_4F	CATACGCTCAAAGCAATGTGA		-HA01606349
LVRN ex20_4R	TGACAGCAGCACAAAATGGA	446	-HA01606350
LVRN ex11_F1	TGAGTTCTGTGTGCTATTTCATCA		HA-01886724
LVRN ex12_R2	CGGGAATGAAGGTGGAATCT	754	HA-01886725

Table 12: Reagents used in PCR reactions.

Reagents	Source
Primer F and R	Sigma-Aldrich
Jump Start REDtaq Ready MIX	Sigma Life Science
dH_2O	Braun

Table 13: Equipment used for PCR.

Equipment	Model	Producer
Pipette	Pipet Lite	Ranin
Centrifuge	Mini Star Silver	VWR
Centrifuge	Spectrafuge Mini	Labnet
Thermal Cycler	MBS Satellite 0,2 Thermal Cycler	Thermo Electron Corporation
Thermal Cycler	MBS 0,2S	Thermo Hybaid
Thermal Cycler	MBS 0,2 G	Thermo Hybaid

Method

First gradient PCR was performed with all primers documented in table 11. The aim with the gradient PCR was to find the best suited annealing temperature for all primers. The components and volumes used in the gradient PCR program are presented in table 14 and were combined in 0.2 ml PCR tubes. Before analysis a quick spin were performed on all samples before they were analysed at the program presented in table 15. After completion of the gradient PCR, two PCR programs with different annealing temperatures were designed, program 1 and 2 (table 16, and 17).

For all PCR experiments performed with PCR program 1 or 2, one negative control were added for each primer to check for contamination. In this control DNA was replaced with sterile water. The primers used in the experiments are presented in table 11. For these PCR experiments the components and volumes presented in table 14 were combined in 0.2 ml PCR tubes. After combining the different volumes, a quick spin were performed before they were analyzed at the conditions presented in table 16 and 17.

Table 14: Reagents used for PCR.

Components	Concentration stock solution	Volume
DNA template	20-100 ng/ µl	1-2 µl
Primer F	10 pmol/µl	1 µl
Primer R	10 pmol/µl	1 µl
Jump Start REDtaq Ready MIX		5 µl
dH_2O		6-7 µl

Table 15: The Gradient PCR program.

Step	Annealing temperature
1	55.5ºC
2	55.8ºC
3	56.5ºC
4	57.8ºC
5	59.5ºC
6	61.4ºC
7	63.4ºC
8	65.4 ºC
9	67.6 ºC
10	69.0 ºC
11	69.8 ºC
12	70.3 ºC

Table 16. PCR program 1 with annealing temperature at 57.8°C

Step	Temperature	Time	Cycles
Denaturation	95°C	1 min	1
Denaturation	95°C	30 sec	30
Annealing	57,8°C	30 sec	
Extension	72,0°C	30 sec	
Final extension	72,0°C	8 min	1
Storage	4°C	Infinite	Infinite

Table 17. PCR program 2 with annealing temperature at 65.4°C.

Step	Temperature	Time	Cycles
Denaturation	95°C	1 min	1
Denaturation	95°C	30 sec	30
Annealing	65,4°C	30 sec	
Extension	72,0°C	30 sec	
Final extension	72,0°C	8 min	1
Storage	4°C	Infinite	Infinite

2.5 Agarose gel electrophoresis

Agarose gel electrophoresis was performed on all PCR reactions to confirm the presence of PCR products and to control that PCR products were absent from the negative control.

Principle

Agarose gel electrophoresis is a method where DNA molecules are separated according to size and charge by their migration pattern in an agarose gel.

DNA molecules are loaded on an agarose gel, made by combining agarose powder with an aqueous buffer. The agarose gel is then covered in buffer and constant voltage is applied. In the electric field produced by the applied voltage, negatively charged DNA molecules will migrate toward the positive pole. During the migration the DNA pass through pores in the agarose gel. The size of the pores varies according to the amount of agarose in the gel. The higher contents of agarose the smaller the pores are. Small DNA molecules moves faster trough the pores than larger ones and migrates further in the gel, separating the molecules by size. How fast the molecules travel through the gel is also dependent on the voltage applied. To visualize the migration pattern, the gel is soaked in a dye after completion of the gel electrophoresis. The dye intercalates with the DNA and creates visible bands on the gel when enlightened with UV light. The size of the DNA molecules can then be determined by comparing the samples migration pattern with a ladder. The ladder is loaded on the gel together with the samples and each band on the ladder has a known size (2).

Materials and reagents

Table 18: Reagents used in electrophoresis.

Components	**Producer**
UltraPure Agarose	Invitrogen
1 X TBE buffer	Made at the lab (table 19)
GelRed Nucleic Acid Gel Stain	Biotium
1 kb plus ladder	Invitrogen

Table 19: Components for the 10 X TBE and 1 X TBE buffers.

Contents	**10 X TBE**	**1 X TBE**
TRIS/BASE	1080 g	108 g
Boric acid	550g	55g
0.5 M EDTA (pH 8)	400 ml	40 ml
Ad aqua dest	10 000 ml	5000 ml
Total	10 000 ml	5000 ml

Table 20: Components for the 1 kb plus ladder.

Contents	**Volume**
Kb-ladder	20 µl
Loading buffer	50 µl
1 X TE buffer	330 µl
Total	400 µl

Table 21: Components for the loading buffer.

Contents	**Volumes**
0.25 % Bromphenol Blue	0.05 g
0.25% Xylene cyanol	0.005 g
40% w/v Sucrose	8.0 g
Ad aqua dest	20 ml
Total	20 ml

Table 22: Components for the TE buffer

Contents	**10 X TE**	**1 X TE**
0.5 M EDTA (pH 8)	20 ml	2 ml
TRIS/HCL	15.8 g	1.58 g
Diluted	1000 µl	100µl

Table 23: Equipment used for electrophoresis.

Equipment	Model	Producer
Geltray, comb and tray	Wide Mini-slub cell GT	BioRad
Power supply	Power Pac 300	BioRad
Power supply	Power Pac basic	BioRad
Imaging system	BioDoc-it 220 Imaging System	UVP

Method

A 1.5 % agarose solution was made by combining 7.5 g agarose with 500 ml of 1 X TBE buffer, which was mixed and warmed in a microwave oven until all of the agarose powder was completely dissolved, and then stored at 65°C. To produce a firm gel the agarose solution was poured into a gel tray containing a comb and left to polymerize for 20 minutes. The comb produces the wells in the gel were the samples are loaded. After polymerization the gel was placed in a tray containing 1 X TBE buffer. Ten µl of 1 kb plus ladder and 2 µl of PCR products were then loaded on the gel and run at 95 V for 30-60 minutes before sealing the gel in dye. In some analysis 4 µl of ladder and PCR products were applied on the gel. After completed electrophoresis the gel was soaked in a dye (GelRed) for 20 minutes to stain the DNA molecules. Before it was placed under UV light in a BioDoc-it imaging system for picture taking it was discoloured in water. By comparing the band on the 1 kb plus ladder with the bands from the PCR products, the size of the sample DNA was determined.

2.6 ExoSAP treatment of PCR products

ExoSAP was performed on all PCR products prior to sequencing.

Principle

ExoSAP is a method for cleaning of PCR products before sequencing by removal of primer leftovers and deoxynucleotides (dNTPs). Exonnuclease 1 removes primer leftovers while Shrimp Alkaline Phosphatase removes excess dNTPS (55).

Material and reagents

Table 24: Components in the ExoSap.

Reagents	Concentration	Volume per reaction	Producer
Exonuclease 1	10 U/µl	0,5	USB
Shrimp Alkaline Phosphatase	2 U/µl	0,5	USB

Method

1µl of ExoSAP was added to the PCR products and run at the PCR program presented in table 25.

Table 25: PCR program for ExoSAP.

Steps	Temperature	Time
Activation	37°C	1 hour
Inactivation	85°C	15 minutes
Soak	4°C	Infinite

2.7 DNA sequencing

DNA sequencing was performed on a selection of samples from preeclamptic patients, normal controls and samples from normal pregnancies.

Principle

Sanger sequencing is a method where the correct order of bases in the DNA sequence is determined. The basis for Sanger sequencing is use of DNA polymerase to synthesise a large amount of DNA fragments of variable length, by incorporation of 2`deoxynucleotides (dNTPs) and fluorescent labelled 2`3 `-dideoxynucleotides (ddNTPs). Primers are used to define the parts of the genome to be sequenced. The first step in the sequencing process is to denaturize the template DNA, followed by annealing of primers to complementary regions in the template. After hybridization of primers, DNA polymerase attaches to the primers 3 `end and starts to synthesis new fragments of different lengths by incorporating dNTPs and ddNTPs. When a dNTP is incorporated the synthesis will continue and creates a growing DNA strand. Incorporation of a ddNTP terminate the synthesis because the ddNTPs lacks a hydroxyl group at the 3` end of the molecule. This hydroxyl group is essential for addition of

new molecules. Since the incorporation of dNTPs and ddNTPs is random, DNA fragments differing by one base pair (bp) from each other are produced.

After completion of the sequencing reaction the DNA fragments are separated by size by capillary gel electrophoresis. The shortest molecules travel fastest through the gel and will first be illuminated by UV light from a laser. All DNA fragments have a ddNTP at the end which emits a fluorescent signal upon illumination. Each of the four ddNTPs (A, C, G and T) are labelled with different fluorescent colours and emits fluorescent lights with different spectres, revealing which type of base that is present at the end of the fragment. A computer program converts the emitted light into a complete sequence where all the bases in the sequence are presented in correct order (2, 56, 57).

Material and reagents

Table 26: Components used in sequencing reactions.

Components	Concentration	Volume	Producer
Primer F or R	10 pmol/µl	0.5µl	Sigma-Aldrich
BigDye V.3.1		0.5µl	Applied Biosystems
5 X buffer to BigDye. V.3.1		3 µl	Applied Biosystems
dH_2O		14	Braun
PCR product		2 µl	

Table 27: Equipment used in sequencing reactions.

Equipment	Model	Producer
Centrifuge	Mini Star Silver	VWR
Centrifuge	Spectrafuge Mini	Labnet
Thermal Cycler	MBS Satellite 0,2 Thermal Cycler	Thermo Electron Corporation
Thermal Cycler	MBS 0,2S	Thermo Hybaid
DNA sequencer	ABI PISM 3130 XL Genetic Analyzer	Applied Biosystems
Software	SeqScape	Applied Biosystems

Method

The components presented in table 28 were combined in a 0.2 ml PCR tube. The amount of PCR product and water varied between different analyses but the total volume of the sequencing reaction was always 20 µl. The cyclic amplification of PCR products in the sequencing reaction was performed in an ordinary thermo cycler at the program presented in

table 29. Separation of the sequencing products were performed by capillary electrophoresis with ABI PRISM 3130 xl Genetic Analyzer from Applied Biosystems.

Table 28: Components used in PCR.

Components	Volume
Primer F or R	0.5µl
BigDye V.3.1	0.5µl
5 X buffer to BigDye. V.3.1	3 µl
dH_2O	14
PCR product	2 µl

Table 29: The PCR sequencing program.

Step	Temperature	Time	Cycles
Denaturation	96°C	10 sec	
Annealing	50°C	5 sec	30
Extension	60°C	4 min	
Storage	4°C	Infinite	Infinite

2.8 Array comparative genomic hybridization

Array comparative genomic hybridization (aCGH) was performed on an assortment of 19 samples from women with preeclampsia and 19 samples from normal pregnancies.

Principle

The aCGH principle is presented in section 1.6 in the introduction (page 10-11)

Material and Reagents

Table 30: Kits used in aCGH analysis

Kit	Lot. no	Producer
NimbleGen Dual colour DNA labelling Kit	12888300, 12209400, 12537900	NimbleGen
NimbleGen labelling and hybridization control Kit.	12110700	NimbleGen

Table 31: Arrays and mixers used in the aCGH analysis.

Arrays and Mixers	Lot. no	Array number	Array format
Array	090819_HG18_WG_CGH_V2_HX3	398843	3X 720K
Array	100718_HG18_WG_CGH_V.3,1 HX3_HX3	461641	3X 720K
		461642	3X720K
		461643	3X720K
		461678	3X720K
Array	100718_HG18_WG_CGH_V3.1_HX12_HX12	461424	12 X135K
		509338	12 X135K
		514596	12 X135K
Mixer	0014 HX3		
	0016 HX3		
	0024 HX12		
	11002197 HX 12		
	11002197 HX12		

Table 32: Equipment used in aCGH analysis.

Equipment	Model	Producer
Pipette	Pipet Lite	Ranin
Centrifuge	Mini Star Silver	VWR
Centrifuge	Spectrafuge Mini	Labnet
Centrifuge	Capselfuge PMC-060	Tommy
Centrifuge	Centrifuge 5415D	Eppendorf
Vortexer	VF2	Janke og Kunkel, KA laboratorietechnik
Thermal Cycler	PTC-200 Pelter Thermal Cycler	MJ Research
MiniVac	MiVac	GeneVac
Dryer	NimbleGen Microarray Dryer	NimbleGen
Hybridization system	Hybridization System 4	NimbleGen
Skanner	GenePix 4000B	Molecular dynamics
Water bath	Ecoline Star Edition E100	Lauda
Gas supply	NimbleGen Compressed Gas Nozzle	NimbleGen
Software	GenePix version 6	NimbleGen
Software	GenePix version 7	NimbleGen
Software	NimbleScan	NimbleGen
Software	Excel	Microsoft

Method for 3X720K and 12X135K arrays

The aCGH procedure consisted of several steps, involving, preparation of Cy3 and Cy5 sample labelling, hybridization, washing, and scanning and data analysis (figure 6). How these steps were performed is documented in the following procedure where they were divided into different sections. A more detailed procedure regarding scanning and data analysis can be found in reference (58).

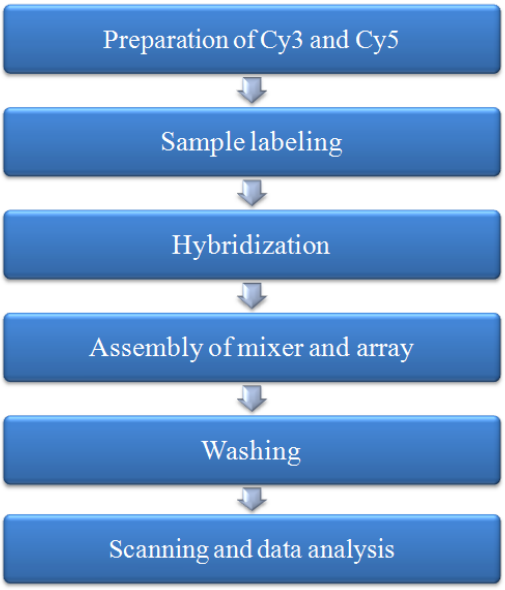


Figure 6: Illustration of the different steps in the aCGH protocol presented below.

Preparation of Cy3 and Cy5 for use in labelling of DNA.

2,200 µl Random primer Buffer and 4 µl B-mercaptoethanol were combined in one vial in fume hood. The vials containing Cy3 and Cy5 were spun briefly and then each of them were diluted in 1,050 µl of the mixture of random primer buffer and B-mercaptoethanol. The contents in the tubes were transferred to 0.2 ml PCR tubes. Each of the tubes contained 40 µl of the solution and was stored at -20°C protected from light.

Labelling of patient and normal control DNA

In this step the patient DNA and reference DNA were labelled with different colours (Cy3 and Cy5). The labelling reaction was performed with oligo nucleotide-primed synthesis by the klenow enzyme. For labelling the components in table 33 were combined in PCR tubes.

Amount of water were adjusted after the amount of DNA added. The total volume of water and DNA were always 40 µl.

Table 33: Components used for labelling of the DNA.

Contents	Patient sample	Normal control
DNA	1µg	1µg
Cy5	40 µl	
Cy3		40 µl
H_2O	Variable	Variable
LHC 1	2 µl	
LHC 2		2 µl

The samples were then heat-denatured in a thermocycler at 98°C for 10 minutes and chilled on ice for 2 minutes. A dNTP/KLENOW master mix for each of the samples was prepared on ice and in a fumehood as described in table 34.

Table 34: Components in the dNTP/Klenow master mix.

dNTP/ klenow Master Mix	Volume per sample
10 mM dNTP Mix	10 µl
PCR Grade Water	8 µl
Klenow fragment	2 µl
Total	20 µl

Twenty µl master mix was added to each samples and the contents was mixed well by pipetting up and down ten times. A quick spin was performed to collect the contents in the bottom of the tube before the samples were incubated for 2 hours at 37°C in a thermo cycler protected from light. After incubation 21.5 µl stop solution was added to the samples to stop the labelling reaction and precipitate DNA. Hundred and ten µl isopropanol was added to the same number of new clean 1.5 ml tubes as the number of samples. The samples were vortexed briefly and spun before the contents were transferred to the 1.5 ml tubes containing isopropanol. After transferral to ispopropanol the tubes were vortexed well and incubated for 10 minutes in room temperature, protected from light. To create a supernatant and form a pellet the samples were centrifuged at 12,000x g for 10 minutes. The supernatant was removed with a pipette, and 500 µl 80 % ice cold ethanol was added. Dislodge the pellet by pipetting a few times up and down. The samples were centrifuged at 12,000 X g for 2 minutes

and the supernatant was removed. A speedVac were used to dry the contents in the tubes on low heat protected from light until they were dry, approximately 5 minutes. After drying the samples were either stored at -20°C for up to 1 month or it was proceeded to the next step.

Rehydration prior to hybridization

Before opening the tubes were spun briefly, and the pellets were rehydrated by adding 35 µl of PCR Grade Water to each tube containing a pellet. The tubes were vortexed for 30 seconds and a quick spin was performed before 5 minutes incubation in room temperature protected from light. When the pellets were completely rehydrated they were vortexed again and a quick spin was performed. The concentrations of the rehydrated samples were then quantified by NanoDrop.

Hybridization

In this step a hybridization buffer was added to the labelled samples. The hybridization buffer contains formamide which denaturizes the DNA and creates single stranded DNA for hybridization. It also prevents unspecific binding at temperatures lower than 37°C. In addition the hybridization buffer contains Tween 20 which minimizes unspecific background noise, and Dextran Sulfate which increases the DNA concentration and improves hybridization for optimization of the results (59). Cot1 DNA is another component in the hybridization buffer and hybridizes to repetitive sequences in the labelled DNA and prevents false fluorescent signals (59).

The hybridization system was programmed to 42°C. Thirty one µg of test samples and reference samples were combined in new clean 1, 5 ml tubes for the 3X720K array and 20 µg for the 12X135K array. The tubes were dried on low heat and protected from light on a SpeedVac. Subsequently the samples were either stored at -20°C for up to one month or it was preceded to the next step.

The dried pellets were re-suspended in 5.6 µl sample tracking control (STC) for the 3X720K array and 3.3 µl for the 12X135K array and vortexed before a quick spin was performed. Afterwards one hybridization master mix was made containing the reagents presented in table 35.

Table 35: Components for the hybridization master mix.

Components	Volume for one slide (3X720K)	Volume for one slide (12X135K)
2 X hybridization buffer	35 µl	88.5 µl
Hybridization Component A	14 µl	35.4 µl
Alignment Oligo	1.4 µl	3.6 µl
Total	50.4 µl	127.5 µl

14.4 µl for the 3X720K array and 8.7 µl for the 12X135K array of the hybridization mix were added to each sample before vortexing for 15 seconds and a quick spin. Followed by incubation for 5 minutes protected from light at room temperature. Afterwards the tubes were placed in a PCR thermal cycler (42°C) for 5 minutes and then vortexed and spun before loading.

Assembly of mixer and array

Before loading of samples the mixer and array were attached to each other by use of the PMAT (figure 7).

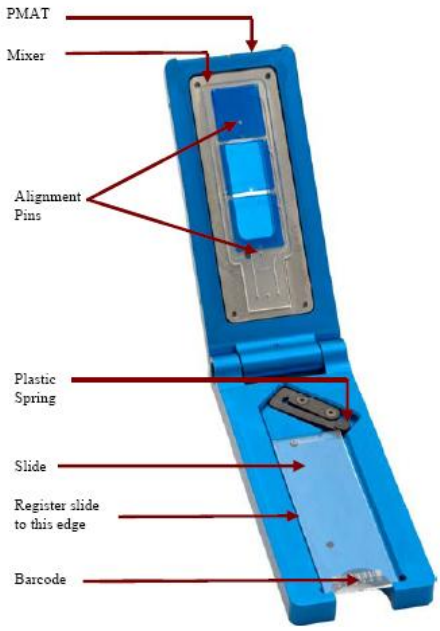


Figure 7: Preparation of mixer and array. The PMAT is illustrated in blue, and the array and mixer are aligned in the PMAT as demonstrated. After alignment of mixer and array in the PMAT, the PMAT is closed and the array attatches to the mixer. Retrived from reference (58).

After assembly of the mixer and array they were removed from the PMAT and incubated in the hybridization system for 5 minutes. For removal of potential air bubbles the mixer Brayer

was rubed over the mixer. The samples were vortexed and spun before loading 18 µl in the fill port on each subarray for the 3X720K array and 6 µl for the 12X135K array. Before inserting the mixer and array into the hybridization system, the fill and vent ports were sealed with mixer ports and adhered properly with the Mixer Brayer. The hybridization reaction was carried out in the range from 68 to 72 hours.

Preparation of washing solutions

Excess labelled DNA located on the array after hybridization can cause false fluorescent signals. To prevent this, the arrays were washed in washing solutions containing PBS and Tween20. One of the washing solutions also contains formamide (59).

Before removal of the mixer and array from the hybridization system, the wash 1, 2 and 3 solutions were prepared as presented in table 36. Two wash 1 solutions were prepared, one for the bucket and one for the slide container.

Table 36: Components and volumes for the washing solutions.

Wash for one slide	**Wash 1 (bucket)**	**Wash 1,2, and 3 (slide container)**
VWR	243ml	24.3 ml
10 X wash buffer 1,2 and 3	27ml	2.7 ml
1M DTT	27 µl	2.7 µl
Total	270 ml	27 ml

The bucket containing wash 1 solution was preheated to 42°C while the others were kept in room temperature. After preparation of the washing solutions, the mixer was removed from the hybridization system and put into the mixer diassembly tool. The mixer disassembly tool was placed in the bucket containing the pre heated wash 1 solution and the mixer was removed from the array by carefully peeling it off.

Washing procedure

The array was shaken in the bucket for 10-15 seconds and then transferred to the container with room tempered wash 1 solution, and shaken for 2 minutes. Then the array was transferred to the container with wash 2 solution and was shaken for 1 minute. In the last

washing step the array was transferred to wash 3 solution and shaken for 15 seconds before it was dried in the NimbleGen Microarray dryer for 2 minutes.

Scanning and data analysis

Scanning and data analysis were performed after the washing procedure was completed.
For scanning the Genepix software was used. Five μm was used as pixel size and the whole array was scanned producing one image. In all scanning procedure the PMT settings for each colour (Cy3 and Cy5) were adjusted to produce histograms with overlapping curves. The curves for Cy3 and Cy5 should be overlapping for most of the histogram to produce good quality results.

After end scanning data analysis was performed by importing the scanned image from the GenePix software into Nimblescan CGH software. Here each subarray on the scanned image was burst (separated) into a separate image, creating three images (A01-A03) for the 3X720K array and twelve images (A01-A12) for the 12X135K array.

After separating the array into subarrays the images were grided by aligning gridmarks in correct position over the array. This was performed automatically by the software, but in some cases manual alignment was necessary.

In the next step in the data analysis an experimental integrity report was made. For each subarray one STC was added in the hybridization step after combining one patient sample with one normal control. In this step one experimental report was made which documented the STCs present in the different arrays, as a quality control for adding the correct sample to the correct subarray.

After control of the STCs, pair reports were created which, documents the signal intensity for each probe on the array.

In the final step of data analysis CGH-segMNT reports were created. Here the data were normalized and copy number analysis was performed. Parameters included 5x averaging window, minimum number of probes equal to 2, and a log2 max/min scaling from -3 to 3. The

gff files produced were used in later bioinformatical analysis (section 2.9. page 48 and 49). Fluorescence ratio plots were made for all chromosomes on each subarray.

2.9 Bioinformatical analysis on aCGH results

Material

Table 37: Summary over the databases used in the processing of aCGH results.

Database	Link
CGHWeb (Childrens Hospital Program webserver)	(http://compbio.med.harvard.edu/CGHweb/)
Ensembl (R biomart package against human genome)	(may2009.archive.ensembl.org)
UCSC Genome Browser on Human Mar.2006 (NCBI36/hg18) Assembly	http://genome.ucsc.edu/cgi-bin/hgGateway?db=hg18
Database of Genomic Variants (DGV)	http://projects.tcag.ca/variation/

Table 38: Summary of the algorithms used in the CGHWeB software presented above

CGHWeb algorithms	Link
Lowess	
Wavelet	http://www.mendeley.com/research/array-cgh-data-modeling-and-smoothing-in-stationary-wavelet-packet-transform-domain/
Quantreg	http://bioinformatics.oxfordjournals.org/content/21/7/1146.full
RunAvg	
CBS	http://bioinformatics.oxfordjournals.org/content/23/6/657.short
CGHseg	http://cran.r-project.org/web/packages/cghseg/
BioHMM	http://bioinformatics.oxfordjournals.org/content/22/9/1144.short
cghFlasso	http://cran.r-project.org/web/packages/cghFLasso/cghFLasso.pdf
GLAD	http://bioinformatics.oxfordjournals.org/content/20/18/3413.short
FASeg	http://www.sph.emory.edu/bios/FASeg/

Method

After generating the ggf- files they were used in bioinformatical analyses. First the gff-formatted reports were fed directly into the CGHWeb webserver for the analysis of CGH data. This service offers the distinct advantage of running several alogrithms in parallel offering users the ability to select results based on the consensus findings of the algorithms. The algorithms used by this server are, Lowess, Wavelet, Quantreg, RunAvg, CBS, CGHseg, BioHMM, cghFlasso,GLAD, and FASeg (figure 8).

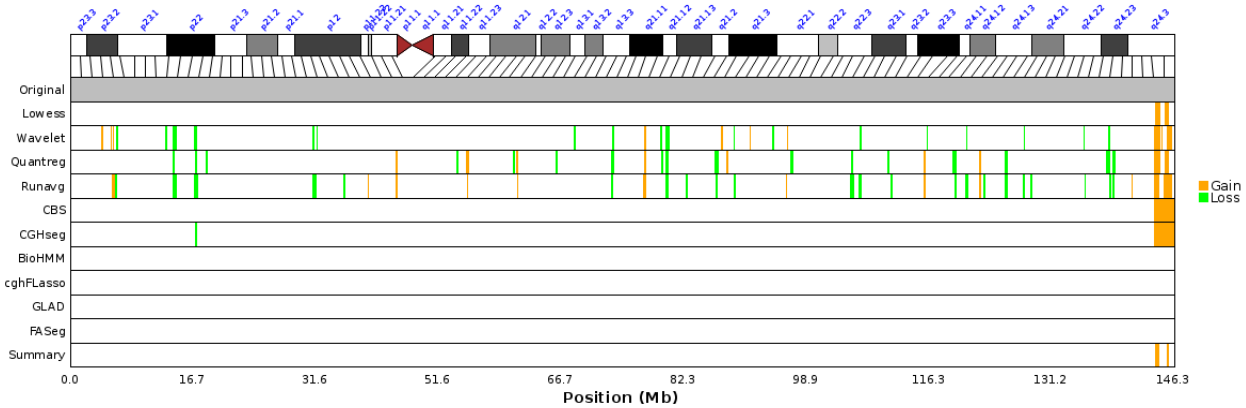


Figure 8: Presentation of the algorithms used in the CGHWeb server. On the left side of the figure the algorithms are documented. The chromosome illustrates the chromosomal positions for the detected deletionsand duplications. Green lines represent detected deletions and orange lines duplications.

Deletions and duplications detected in more than four of the presented algorithms were included in further analysis where consensus areas of deviation were parsed from the CGHweb analysis and mapped against known genes in area. Positions were mapped using R biomart package against human genome build 18 (may2009.archive.ensembl.org) in accordance with the design file supplied with the CGH arrays. Gene name, specifically HGNC gene name was taken from the Ensembl annotation based on chromosome name, probe start and probe stop. To better review and integrate results a html- page was generated for each array with direct hyperlinks to the UCSC browser for areas in question, from these the tables in appendix A (I-XX) were made. All deleted or duplicated areas involving 5 or more probes were included and were documented with chromosome position, start and end position for the probes, number of probes, genes involved and mean values. The mean value represent the average signal strength found within the window of probes.

The USCS hyperlinks were used to document the chromosomal position and gene contents for the areas in question (CNVs). This database also contained information on gene function and expression profiles for the detected genes and was used in the evaluation of their potential role in the pathophysiology of preeclampsia.

The database of Genomic Variants (DGV) was screened for the areas in question. DGV presents an outline over the structural variation in the human genome and structural variants involving more than 1kb in size are included in this database. All variants found in samples from healthy individuals are presented in this database.

3.0 Results

3.1 DNA extraction

Manual extraction was performed on normal controls from healthy individuals (K1-K4, M1-M6), while automatic extraction was carried out on all samples from preeclamptic patients and normal pregnancies used in PCR, sequencing and aCGH analysis. For comparison of the two extraction methods the controls from healthy individuals were also extracted automatically. Following the extraction process the DNA concentration was quantified by NanoDrop and all plots were examined before subsequent analyses were performed on the extracted DNA. The two extraction methods showed similar results regarding both concentrations and purity. The purity was determined after evaluation of the A260 and A280 ratios. The samples with the highest purity and concentrations were selected for succeeding analysis.

For the normal controls the DNA concentrations were within the range of 16.0–42.5 ng/µl. The patients samples used in sequencing (maternal EDTA blood samples) had DNA concentrations in the range from 43.5–112.1ng/µl while the samples from normal pregnancies had concentrations from 24.0–142.0 ng/µl. The patients samples used in aCGH had concentrations in the range of 32.0–104.1ng/µl and the samples from normal pregnancies from 27.0–104.5 ng/µl

3.2 Gradient PCR

Gradient PCR was performed on DNA extracted from healthy individuals for all the primers used in sequencing of the *laeverin* gene. The controls were selected based on their concentration and purity after quantification with NanoDrop. Sample M6 was used for the primers LVRN ex 1_1F/R-3F/R, K4 for primer LVRN ex 4F/R-10F/R, K3 for primers LVRN ex 11F/R-16F/R, M4 for primers LVRN ex 17F/R-20_3F/R and M1 for primers LVRN ex 20_4F/R and ex 11_F1/12_R2.

For all gel pictures presented in this section, the numbers on the left side of the picture represent the size of the fragments of the 1 kb plus ladder used in all experiments. All gel pictures have some weak bands above the first visible band on the ladder. These represent genomic DNA present in the PCR products. All primers were run at the same PCR gradient program which is presented in table 15. The PCR products from the different annealing temperatures were loaded on the gel in the same order as the gradient PCR program. The first well on the gel pictures always contains PCR products from annealing temperature 55.5°C continuing with increasing temperatures in the following wells.

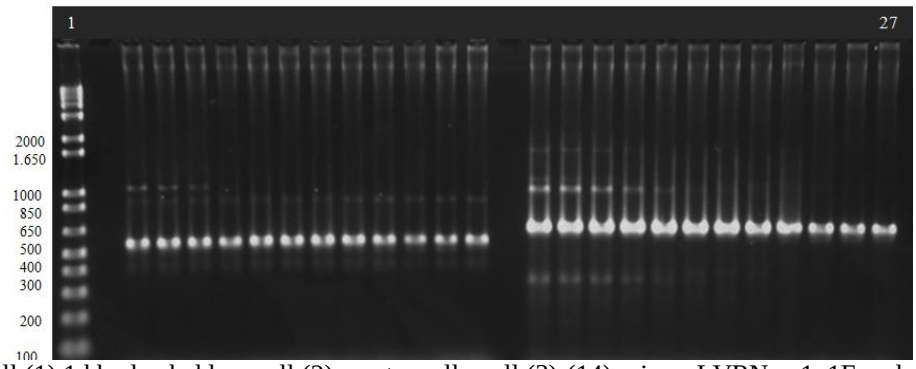


Figure 9: Well (1) 1 kb plus ladder, well (2) empty well, well (3)-(14) primer LVRN ex1_1F and ex 1_1R, Well (15) empty well, Well (16)-(27) primer LVRN ex 1_2F and ex 1_2R.

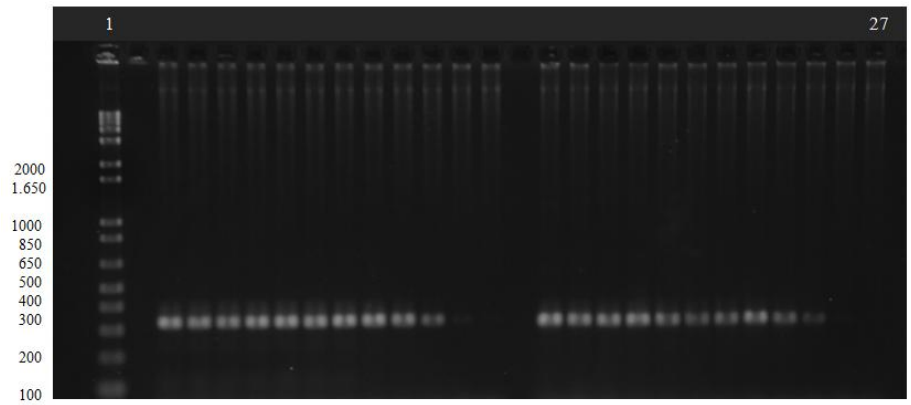


Figure 10: Well (1) 1 kb plus ladder, Well (2) empty well, Well (3-14) Primer LVRN ex 2F and ex 2R, Well (15) empty well, Well (16)-(27) Primer LVRN ex 3F and ex 3R.

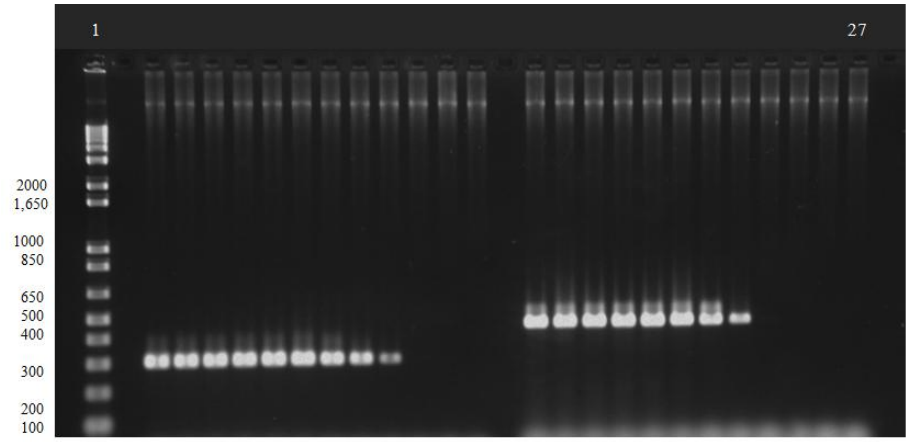


Figure 11: Well (1) 1 kb plus ladder Well (2) empty well , Wells (3)-(14) primer LVRN ex 4F and ex 4R, Well (15) empty well, Wells (16)-(27) primer LVRN ex 5F and ex 5R.

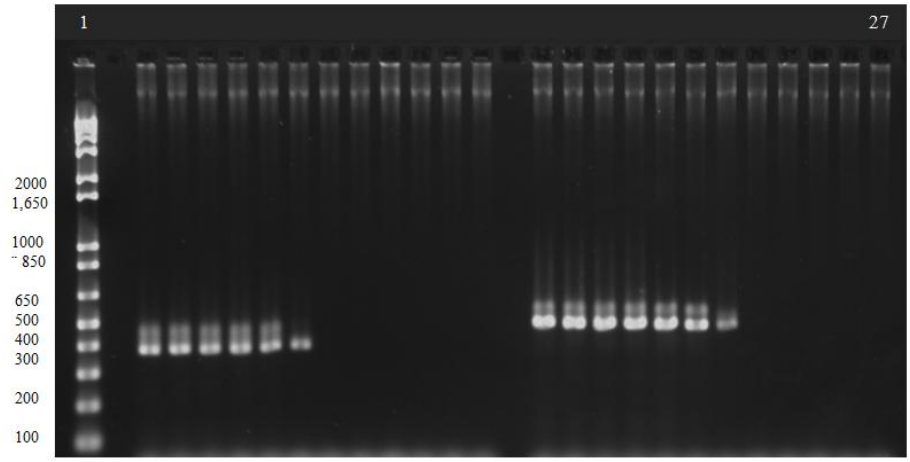


Figure 12: Well (1) 1 kb plus ladder well (2) empty well, wells (3)-(14) Primer LVRN ex 6F and ex 6R (15) empty wells (16)-(27) primer LVRN ex 7F and ex 7R.

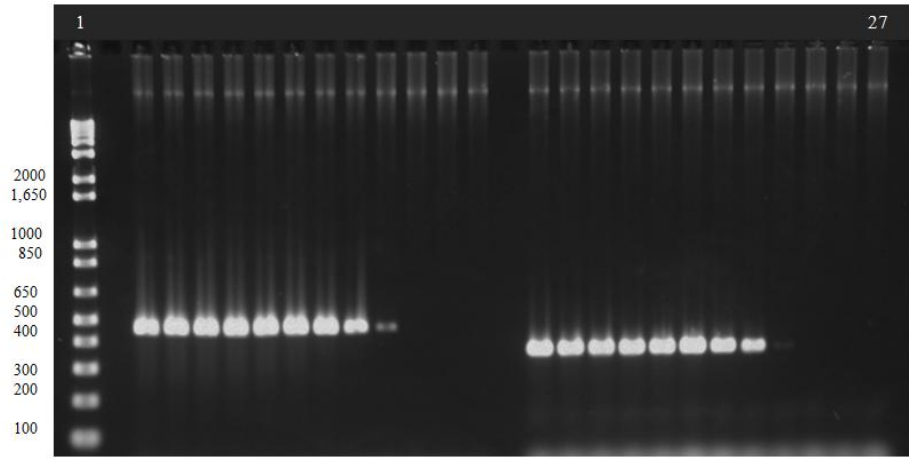


Figure 13: Well (1) 1 kb plus ladder well (2) empty well, wells (3)-(14) primer LVRN ex8_9F and ex 8_9R, Well (15) empty well, wells (16)-(27) primer LVRN ex10F and ex10R.

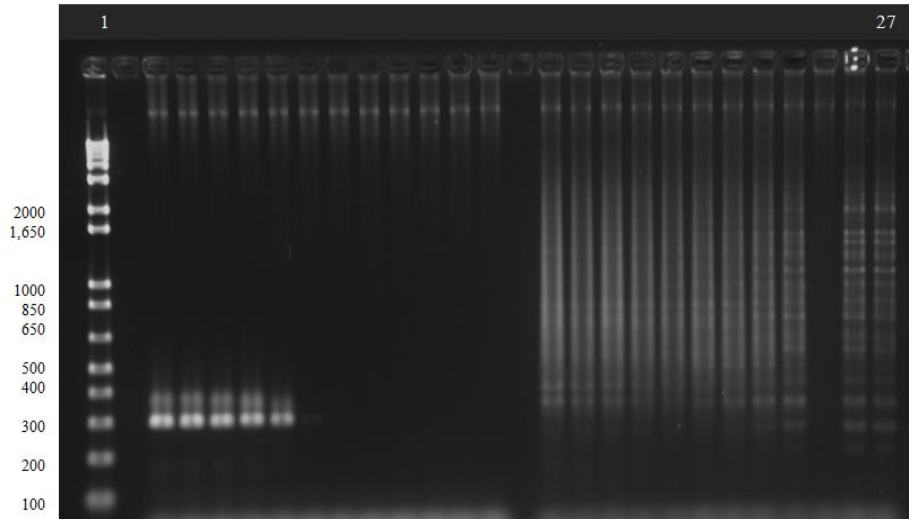


Figure 14: Well (1) 1 kb plus ladder, well (2) empty well, wells (3)-(14) Primer LVRN ex11F and ex 11R (15) empty well (16)-(27) primer LVRN ex 11F and ex 12R.

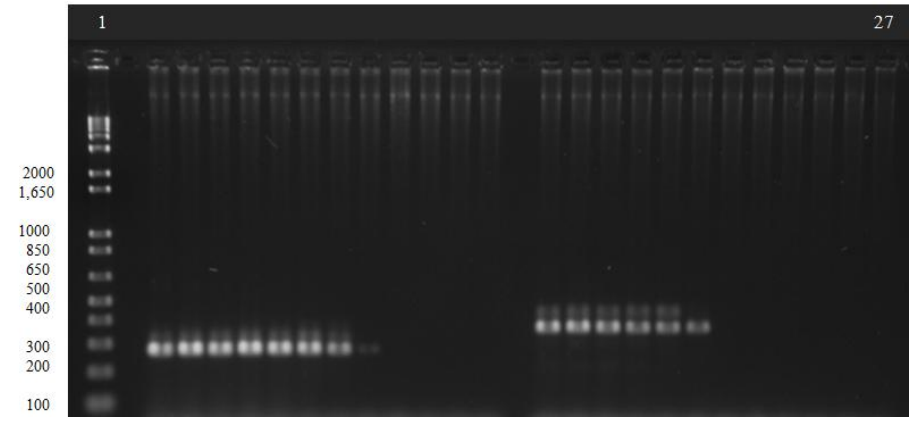


Figure 15: Well (1) 1 kb plus ladder, well (2) empty well, wells (3)-(14) primer LVRN ex 13F and ex 13R, well (15) empty well, wells (16)-(27) Primer LVRN ex 14F and ex 14R.

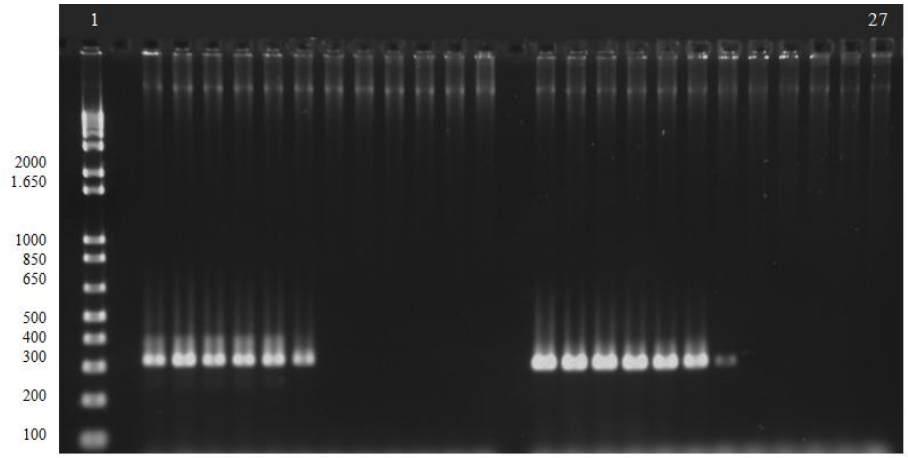


Figure 16: Well (1) 1 kb plus ladder , well (2) empty well, wells (3)-(14) primer LVRN ex 15F and ex 15R , well (15) empty well, wells (16)-(27) primer LVRN ex 16F and 16R.

Figure 17: Well (1) 1 kb plus ladder, well (2) empty well, wells (3)-(14) Primer LVRN ex 17F and ex 17R, well (15) empty well, wells (16)-(27) primer LVRN ex 18F + ex 18R

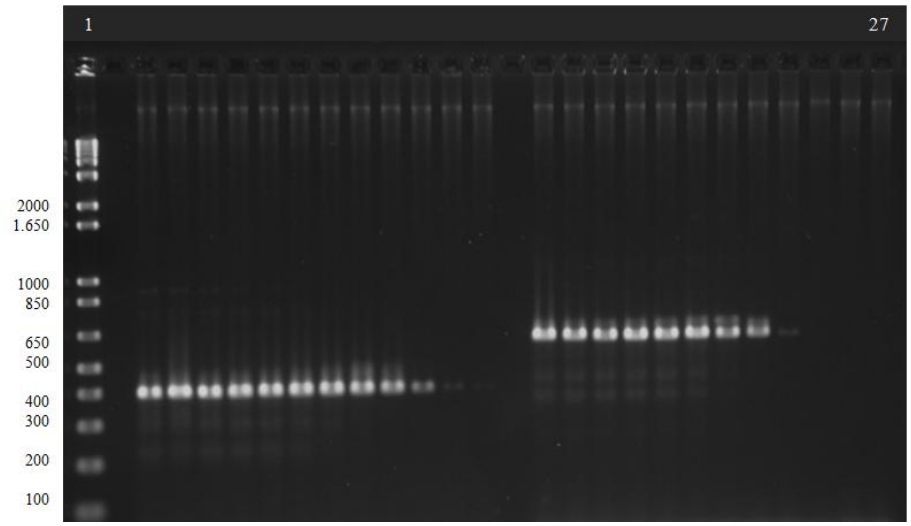


Figure 18: Well (1) 1 kb plus ladder, well (2) empty well, wells (3)-(14) primer LVRN ex 19F and ex 19R, well (15) empty well, wells (16)-(27) primer LVRN ex 20_1F and ex 20_1R.

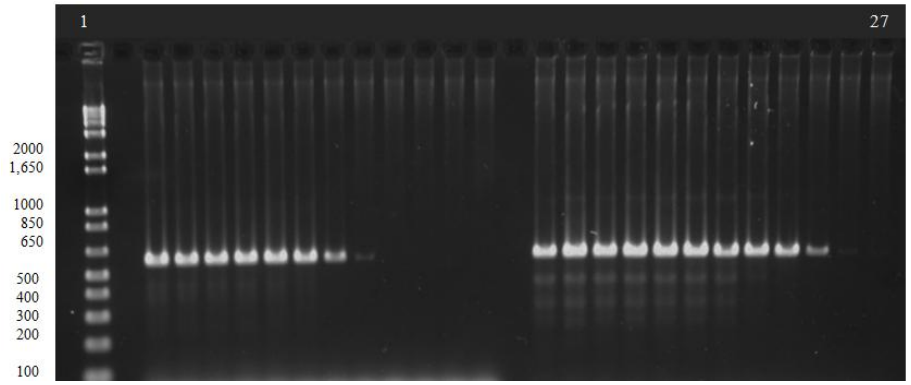


Figure 19: Well (1) 1 kb plus ladder, well (2) empty well, wells (3)-(14) Primer LVRN ex 20_2F and ex 20_2R well (15) empty well (16)-(27) Primer LVRN ex 20_3F + ex 20_3R.

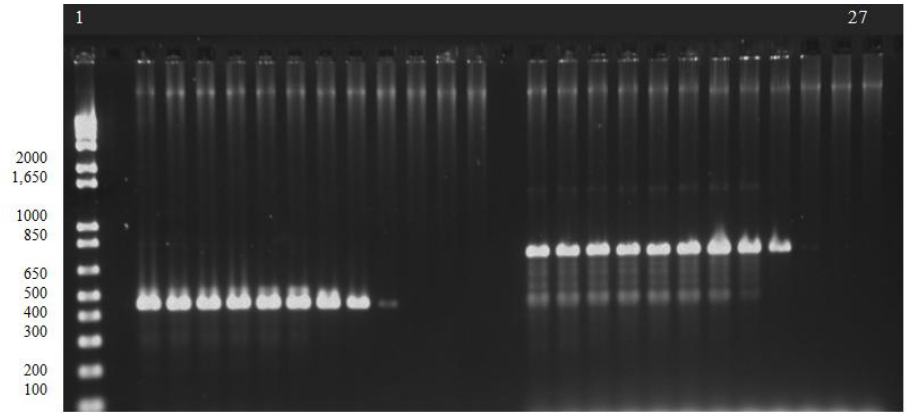


Figure 20: Well (1) 1 kb plus ladder, well (2) empty well, wells (3)-(14) primer LVRN ex 20_4F and ex 20_4R well (15) empty well, wells (16)-(27) Primer LVRN ex 11_F1 and ex12_R2

Table 39 was based on evaluation of all documented gel pictures. All primers were presented with the temperatures which produced their specific bands and the selected PCR program. In addition the temperatures which produced unspecific bands and to weak bands for subsequent analysis were included. Temperatures where no PCR products were produced were included in the column for weak bands.

Table 39: Summary of results from the gradient PCR experiments. Each primer is presented with the annealing temperatures providing their specific bands, unspecific bands and the temperatures where no bands were detected. The bp size for all bands is documented in addition to the PCR program selected for each primer. The annealing temperature for PCR program 1 is 57.8 °C and for program 2, 65.4°C

Primers	**Specific bands (Size)**	**Unspecific bands (size)**	**Weak bands**	**PCR program**
LVRN ex 1_1F/R	55.5°C -70.3°C (548 bp)	55.5ºC -56.5ºC (1000 bp) 55.5ºC -70.3ºC (850 bp)	-	2
LVRN ex 1_2F/R	55.5°C -70.3°C(637bp)	55.5ºC-56ºC (1,650bp) 55.5°C -57.8ºC (1000bp)	-	2
LVRN ex 2F/R	55.5ºC -67.6ºC (324 bp)	55.5-67.6 ºC (350bp)	69.0ºC -70.3ºC	1
LVRN ex 3F/R	55.5°C -67.6°C (313bp)	-	69.0ºC -70.3ºC	1
LVRN ex 4F/R	55.5ºC.65.4ºC (303 bp)	55.5°C -63.4°C (250bp)	67.6ºC-70.3ºC	1
LVRN ex 5F/R	55.5°C -65.4°C (451bp)	55.5°C -63.4°C (500bp)	67.6ºC -70.3ºC	1
LVRN ex 6F/R	55.5°C -61.4ºC (373)	55.5°C -59.5°C (450bp)	61.4ºC -70.3ºC	1
LVRN ex 7F/R	55.5ºC -61.4ºC (467)	55.5°C -61.4°C (550bp)	63.4°C -70.3ºC	1
LVRN ex 8_9F/R	55.5ºC -65.4ºC (441bp)	-	67.6°C -70.3ºC	1
LVRN ex 10/R	55.5ºC -65.4ºC (349bp)	-	67.6°C -70.3ºC	1
LVRN ex 11F/R	55.5ºC -59.5ºC (295bp)	55.5°C -59.5°C (350bp)	61.4ºC -70.3ºC	1
LVRN ex 11F /12R	-	55.5ºC -70.3ºC	-	-
LVRN ex 13F/R	55.5ºC -63.4ºC (295bp)	55.5°C -63.4°C (330bp)	65.4ºC -70.3ºC	1

Primers	Specific bands (Size)	Unspecific bands (size)	Weak bands	PCR program
LVRN ex14F /R	55.5ºC 5-61.4ºC (358bp)	55.5°C -59.5°C (400bp)	63.4ºC -70.3ºC	1
LVRN ex 15F/R	55.5ºC -61.4ºC (315bp)	55.5°C -59.5°C (400bp)	63.4ºC -70.3ºC	1
LVRN ex 16F/R	55.5ºC -61.4ºC (295bp)	-	63.4ºC -70.3ºC	1
LVRN ex 17F/R	55.5ºC -63.4ºC (391bp)	55.5°C -63.4°C (500bp)	65.4ºC -70.3ºC	1
LVRN ex 18F/R	55.5ºC -63.4ºC (291bp)	55.5ºC -63.4ºC (350bp)	65.4ºC -70.3ºC	1
LVRN ex 19F/R	55.5ºC -67.6ºC (399bp)	55.5ºC -59.5ºC (900bp) 55.5 ºC -61.4 ºC(200,300bp)	69.0ºC-70.3 ºC	2
LVRN ex 20_1F/R	55.5ºC -65.4ºC (590bp)	55.5ºC -65.4ºC (650bp) 55.5°C-61.4°C (400,500bp)	67.6ºC -70.3ºC	1
LVRN ex 20_2F/R	55.5ºC -63,4 ºC (571bp)	-	65.4ºC -70.3ºC	1
LVRN ex 20_3F/R	55.5ºC -67,6ºC (590bp)	55.5ºC -63.4ºC (400,500bp)	69.0-70.3ºC	2
LVRN ex 20_4F/R	55.5ºC -65.4ºC (446bp)	-	67.6ºC -70.3ºC	1
LVRN ex 11_F1/ 12R2	55.5ºC -67.6ºC (754bp)	55.5ºC -65.4ºC (500bp)	69.0ºC -70.3ºC	2

3.3 Polymerase chain reaction

The aim with the PCR experiments was to produce products for sequencing. After designing the two different PCR programs, PCR were performed on all samples prior to sequencing. The selected samples for PCR and sequencing were chosen after estimation of their concentrations and purity after quantification by NanoDrop. All samples had acceptable concentrations and ratios and the selected assortment was randomly chosen. PCR products for all exons and parts of introns in the *laeverin* gene were produced for the sample numbers 3P, 5P, 15P, 20P, 24P, 25P, 26P, 32P, M6, K4, 35N, 19 and 20. Furthermore PCR products for exon 7 and 10 were made from 20 samples from preeclampsia patients and 18 samples from normal pregnancies. In addition, PCR products for exon 7 were produced for 23 samples from normal pregnancies.

Negative water controls were applied for all primers in each experiment to examine if the samples were contaminated. After completion of the PCR experiments, gel electrophoresis was performed on all PCR products to investigate it there was PCR products in the negative control and to control that PCR products were present for all samples. Figure 21 documents

one representative gel electrophoresis performed on PCR products from preeclampsia patients with primer LVRN ex 10F/R.

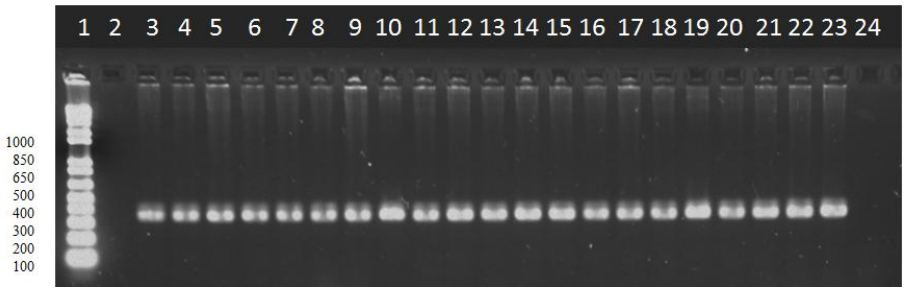


Figure 21: Well (1) 1 kb plus ladder well (2) empty well, well (3) 1P well (4) 2P well (5) 4P well (6) 6P well (7) 7P well (8) 8P well (9) 9P well (10) 10P well (11) 11P well (12) 12P well (13) 13P well (14) 14P well (15) 16P well (16) 17P well (17) 18P well (18) 19P well (19) 21P well (20) 22P well (21) 23P well (22) 20P well (23) 8N well (24) H_2O (negative control)

At the top of the wells some genomic DNA is present. PCR products were present in wells 3-23. In well 24 no bands are detected and the control is negative, indicating that there is no contamination of DNA.

3.4 DNA sequencing

3.4.1 Sequencing of DNA from maternal blood samples

An assortment of DNA samples extracted from maternal blood samples of preeclampsia patients and normal pregnancies were selected for sequencing. All samples had acceptable concentrations and purity and were randomly selected after DNA quantification with NanoDrop. Sample number 3P, 5P, 15P, 20P, 24P, 25P, 26P, 32P, M6, K4, 35N, 19 and 20 were sequenced for all exons in the *laeverin* gene including small parts of the introns. Twenty samples for preeclampsia patients and 18 samples from normal pregnancies were sequenced for exons 7 and 10. In addition 23 samples were sequenced for only exon 7. Detected variant from this sequencing analysis are presented in table 40. Normal sequences without variants are not included.

Table 40: Summary of detected variants from sequencing of the *laeverin* gene on DNA extracted from blood samples. (Het) is an abbreviation for a heterozygote variant while (Homo) is an abbreviation for a homozygote variant. The guidelines for the HGVS nomenclature retrieved from reference (60).

Location	HGVS nomenclature	Sample number	Protein change
LVRN Exon 1	g.-83 G›C (Het)	3P, 5P, 24, M6, 26P, K4	Untranslated region
LVRN Exon 1	g.-83 G›C (Homo)	15P, 20P, 32P, 35N, 25P, 19,20	Untranslated region
LVRN Exon 1	g. 64 C› T (Het)	3P, 5P, 24P, M6,26P, K4	None
LVRN Exon 1	g. 64 C› T (Homo)	15P, 20P, 32P, 35N, 25P, 19,20	None
LVRN Exon 1	g. 161 T› C (Het)	3P, 5P, 24P, M6, 26P, K4	p.Lys54Ser
LVRN Exon 1	g. 161 T› C (Homo)	15p, 20P, 32P, 35N, 25P, 19,20	p.Lys54Ser
LVRN Exon 1	g. 204 A› G (Het)	3P, 5P, 24P, M6, 26P, K4	None
LVRN Exon 1	g. 204 A› G (Homo)	15P, 20P, 32P, 35N, 25P, 19,20	None
LVRN Exon 1	g. 663 C›T (Het)	24P, M6, 26P, K4	None
LVRN Exon 1	g. 663 C›T (Homo)	15P, 32P, 35N, 25P, 19	None
LVRN Exon 7	g. 1459 G›A (Het)	20P	p.Glu487Lys
LVRN Exon 10	g. 1746 T› C (Het)	3P, 20P, 11P, 12P, 18P, 19P, 21P, 23P,1N, 5N, 6N, 12N, 14N, 18N, 21N	None
LVRN Exon 10	g. 1746 T› C (Homo)	16P, 2N, 10N, 17N	None
LVRN Intron 12	g.2037+9A(-A›G) (Het)	3P, 5P, 20P, 24P, M6	None coding
LVRN Intron 12	g.2037+9A (- A›G) (Homo)	15P, 32P, 35N	None coding
LVRN Exon 13	g. 2040 G›T (Het)	3P, 5P,20P,24P, M6	None
LVRN Exon 13	g. 2040 G›T (Homo)	15P, 32P, 35N, 26P,19, 20, K4, 25P	None
LVRN Exon 13	g.2067 G› C (Het)	3P, 5P, 20P, 24P, M6	p.Leu 689Phe
LVRN Exon 13	g.2067 G› C (Homo)	15P, 32P, 35N, 26P, 19, 20, K4,25P	p.Leu 689Phe
LVRN Intron 17	g. 2618 + 34- T›G (Het)	3P, 15P, 24P	None coding
LVRN Intron 17	g. 2618+ 34- T›G (Homo)	5P, 32P, 35N, M6	None coding
LVRN Exon 20	g.3306 (*3) A›G (Het)	3P, 24P, 32P, 35N, 20, K4, 25P, 20	None
LVRN Exon 20	g.3306 A›G (Homo)	5P, M6, 26P, 19	None
LVRN Exon 20	g.3377(*74) T›C (Het)	15P, 24P	None
LVRN Exon 20	g. 4439(*1136) T›G (Het)	25P	

After sequencing of all exons and parts of introns in the sample numbers presented above, only two of the variants documented in table 40 were found exclusively in samples from preeclamptic patients in translated exons. These variants were in exon 7, position g. 1459G›A and in exon 10 position g. 1746T› C. Electropherograms documenting these two variants are

presented in figures 22 and 23. Twenty samples from preeclamptic patients and 18 samples from normal pregnancies were screened for these variants. The variant in exon 10 was detected in seven more samples from preeclampsia patients and in 10 samples from normal pregnancies and was considered as a normal variant. The variant detected in exon 7 did not occur in any samples from normal pregnancies or controls. Furthermore 23 samples from normal pregnancies were sequenced for exon 7. None of these contained the variant and the variant is therefore a potential pathogenic mutation.

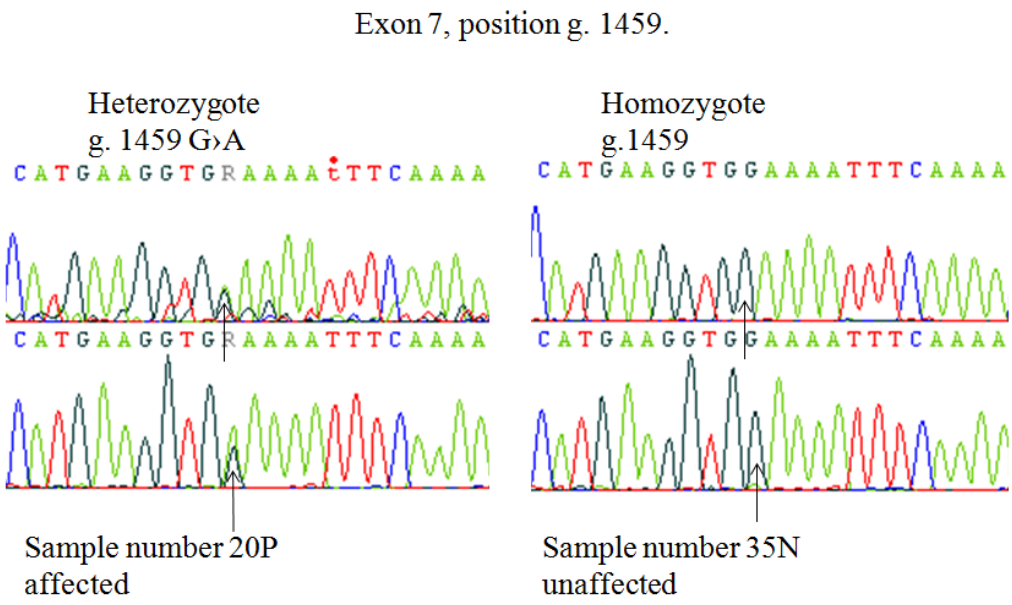


Figure 22: The left side of the figure is an electropherogram from Applied Biosystems seqscape software illustrating the heterozygote variant found in exon 7 position g.1459G>A in sample 20P. The two arrows illustrates that one A nucleotide is present at both strand in addition to one G nucleotide in this position. The G nucleotide is normally present alone at this position in the genome. Right side of the figure is an electropherogram from Applied Biosystems seqscape software illustrating the normal sequence of exon 7, position g.1459. The arrows document the presence of only one G nucleotide at this location.

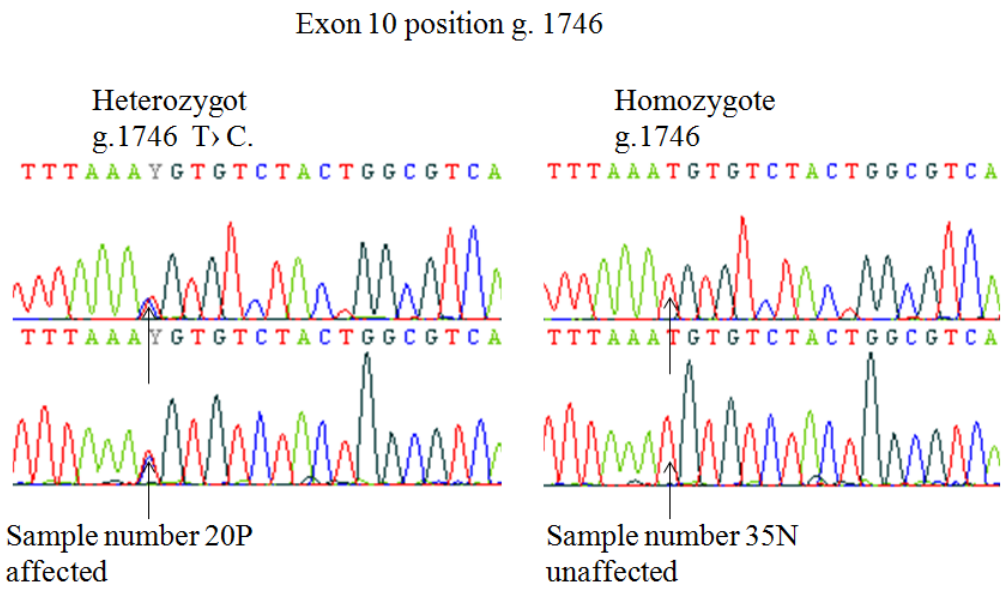


Figure 23: The left side of the figure is a electropherogram from Applied Biosystems seqscape software illustrating the heterozygote variant found in exon 10 position g.1746T>C in sample 3P and 20P. The two arrows indicate the presence of both one C and one T nucleotide at this position. The right side illustrates the homozygote normal sequence for exon 10 position g.1746. The two arrows document the presence of only one T nucleotide at this position.

To investigate the effect of the potential pathogenic mutation in exon 7 the amino acid substitution caused by the mutation was documented. Figure 24 presents the normal amino acid composite for the *laeverin* gene. According to this one Glutamic acid is replaced with Lysine for position 487 and is marked with red in figure 24. Both amino acids are hydrophilic.

10	20	30	40	50	60
MGPPSSSGFY	VSRAVALLLA	GLVAALLLAL	AVLAALYGHC	ERVPPSELPG	LRDLEAESSP
70	80	90	100	110	120
PLRQKPTPTP	KPSSARELAV	TTTPSNWRPP	GPWDQLRLPP	WLVPLHYDLE	LWPQLRPDEL
130	140	150	160	170	180
PAGSLPFTGR	VNITVRCTVA	TSRLLLHSLF	QDCERAEVRG	PLSPGTGNAT	VGRVPVDDVW
190	200	210	220	230	240
FALDTEYMVL	ELSEPLKPGS	SYELQLSFSG	LVKEDLREGL	FLNVYTDQGE	RRALLASQLE
250	260	270	280	290	300
PTFARYVFPC	FDEPALKATF	NITMIHHPSY	VALSNMPKLG	QSEKEDVNGS	KWTVTTFSTT
310	320	330	340	350	360
PHMPTYLVAF	VICDYDHVNR	TERGKEIRIW	ARKDAIANGS	ADFALNITGP	IFSFLEDLFN
370	380	390	400	410	420
ISYSLPKTDI	IALPSFDNHA	MENWGLMIFD	ESGLLLEPKD	QLTEKKTLIS	YVVSHEIGHQ
430	440	450	460	470	480
WFGNLVTMNW	WNNIWLNEGF	ASYFEFEVIN	YFNPKLPRNE	IFFSNILHNI	LREDHALVTR
490	500	510	520	530	540
AVAMKVENFK	TSEIQELFDI	FTYSKGASMA	RMLSCFLNEH	LFVSALKSYL	KTFSYSNAEQ
550	560	570	580	590	600
DDLWRHFQMA	IDDQSTVILP	ATIKNIMDSW	THQSGFPVIT	LNVSTGVMKQ	EPFYLENIKN
610	620	630	640	650	660
RTLLTSNDTW	IVPILWIKNG	TTQPLVWLDQ	SSKVFPEMQV	SDSDHDWVIL	NLNMTGYYRV
670	680	690	700	710	720
NYDKLGWKKL	NQQLEKDPKA	IPVIHRLQLI	DDAFSLSKNN	YIEIETALEL	TKYLAEEDEI
730	740	750	760	770	780
IVWHTVLVNL	VTRDLVSEVN	IYDIYSLLKR	YLLKRLNLIW	NIYSTIIREN	VLALQDDYLA
790	800	810	820	830	840
LISLEKLFVT	ACWLGLEDCL	QLSKELFAKW	VDHPENEIPY	PIKDVVLCYG	IALGSDKEWD
850	860	870	880	890	900
ILLNTYTNTT	NKEEKIQLAY	AMSCSKDPWI	LNRYMEYAIS	TSPFTSNETN	IIEVVASSEV
910	920	930	940	950	960
GRYVAKDFLV	NNWQAVSKRY	GTQSLINLIY	TIGRTVTTDL	QIVELQQFFS	NMLEEHQRIR
970	980	990			
VHANLQTIKN	ENLKNKKLSA	RIAAWLRRNT			

Figure 24: Presentation of the amino acid sequence for the *laeverin* gene. The amino acids in the M1 motif are presented in green. The underlined amino acid is hydrophobic residues corresponding to a transmembrane domain. The amino acid marked in red for position 487 is replaced with Lysine in the possible pathogenic mutation detected in exon 7. The amino acid marked in red in position 455 is the amino acid replaced with Arginine in exon 6 in the variant detected in the placenta sample. Modified from reference (46).

3.4.2 Sequencing of DNA from placental tissue

Sample 15P collected from placental tissue were sequenced for all exons including small parts of the introns for the *laeverin* gene. This sample originates from the same person as sample 15P collected from EDTA blood. The results from this sequencing are documented in table 41. Only detected variants are included in this table. Normal sequences without variants are not presented.

Table 41: Summary of variations found in sample 15P from placenta tissue sequenced for all exons and parts of the introns in the *laeverin* gene. The guidelines for the HGVS nomenclature is retrieved from reference (60) The normal controls which the placental samples are compared with are from maternal blood samples. (Het) is an abbreviation for a heterozygote variant while (Homo) is an abbreviation for a homozygote variant. Variants detected in normal controls are marked with +. Variants which are not detected in normal controls are marked with -.

Location	HGVS nomenclature	Protein change	Detected in normal controls
LVRN Exon 1	g -41 G›C (Het)	Untranslated region	
LVRN Exon 1	g.64 C›T (Het)	None	+
LVRV Exon 1	g.162T› C (Het)	p.Leu54Ser	+
LVRN Exon 1	g.204 A›G (Homo)	None	+
LVRN Exon 1	g.664 C›T (Het)	None	+
LVRN Exon 6	g.1365 A›G (Het)	p.Leu455Arg	-
LVRN Intron 12	g2037+9-A›G (Homo)	Possible splice site change	+
LVRN Exon 13	g.2040 G› T (Homo)	None	+
LVRN Exon 13	g.2068 G›C (Homo)	p.Leu689Phe	+
LVRV Exon 20	g.3307 (*3)A›G (Het)	After translation termination	+
LVRN Exon 20	g.3378 (75*)T›C (Het)	After translation termination	- (15P,24P)

Only the variants found in exon 6 at position g.1365A›G (figure 25) and exon 20 g.3378 (75*)T›C were found exclusively in placental tissue. The variant in exon 20 is after termination of translation and is therefore not considered to be significant. The placenta tissue contains fetal DNA, while the other sequenced preeclampsia samples, normal controls and samples from normal pregnancies contains maternal DNA. In the variant detected in exon 6 the amino acid Lysine is replaced with Arginine which both are hydrophilic amino acids.

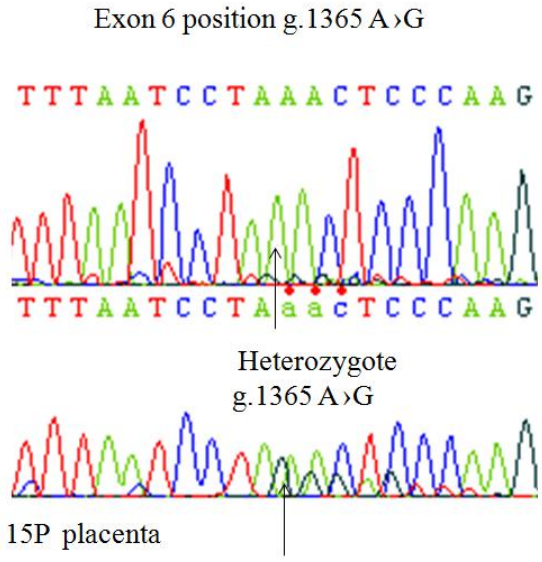


Figure 25: Electropherogram from Applied Biosystems seqscape software illustrating the difference between the forward and reverse strand in the sequence from the placenta tissue. At the bottom strand the arrow illustrates the presence of one G nucleotides in addition to one A nucleotide. In the upper strand this G nucleotide is absent. The area of the green spike in the bottom half of the electropherogram is only half the area of the one in the upper strand. When a mutation is said to be heterozygote the area of the spikes should be halved from the normal homozygote variation.

3.5 aCGH results

aCGH is sensitive to DNA quality thus an assortment of 19 samples from preeclamptic patients and normal pregnancies were selected based on DNA concentration and purity. In total eight aCGH experiments were performed according to the procedure presented in section 2.8 and 2.9 (page 47-49). The arrays used in these experiments are presented below and in table 31.

Array Formats 3X720K

The numbers on the arrays with the 3X720K format used in the experiments were, 39884, 461642,461643,461641, and 461678

Array format 12X135K arrays

The numbers on the arrays with the 12X135K format used in the experiments were, 461424,509338, and 514596

All analyzed 3X720K arrays contained a mixture of DNA from umbilical cord blood from 19 patients with preeclampsia compared with umbilical cord blood from 19 women with normal pregnancies. This DNA mixture was analyzed in triplicate on the A01, A02 and A03 subarrays. On the 12X135K arrays nineteen samples from preeclamptic women were analyzed separately against mixture of nineteen samples from normal pregnancies.

To measure the reproducibility between arrays and within one array parallels of some samples were included. Samples 12P and 21P were analyzed in duplex on array number 514596, while sample 9P was analyzed on array number 509338 and number 514596. The results from array number 461424 were excluded due to problems with hybridization and all samples were re-analyzed at array number 509338. Table 42 is a summary over samples analyzed on the different subarrays and arrays on the 12X135K format.

Table 42: Summary of samples analyzed on the different subarrays and arrays on the 12X135K arrays.

Subarray	Array 461424	Array 509338	Array 514596
A01	2P	2P	4P
A02	4P	4P	9P
A03	7P	7P	12P
A04	8P	8P	12P
A05	9P	9P	21P
A06	10P	10P	21P
A07	12P	12P	23P
A08	13P	13P	24P
A09	15P	15P	25P
A10	16P	16P	26P
A11	17P	17P	30P
A12	20P	20P	33P

3.5.1 Images of the arrays

The Arrays were scanned using the Genepix software as described in section 2.8 (page 47). Subsequent to scanning the images were inspected for occurrence of potential air bubbles or other artifacts, which may cause problems in the data analysis. Furthermore the distribution pattern for Cy3 and Cy5 were examined.

The scanned images from the experiments are presented below. Array numbers, 39884, 461642, and 461643 were analyzed at 10 µm in pixel size instead of 5 µm due to a default setting in the Genepix software. The results from these analyses were therefore excluded and are not presented in this thesis. The images from the 12x135K arrays are in this section separated into two images.

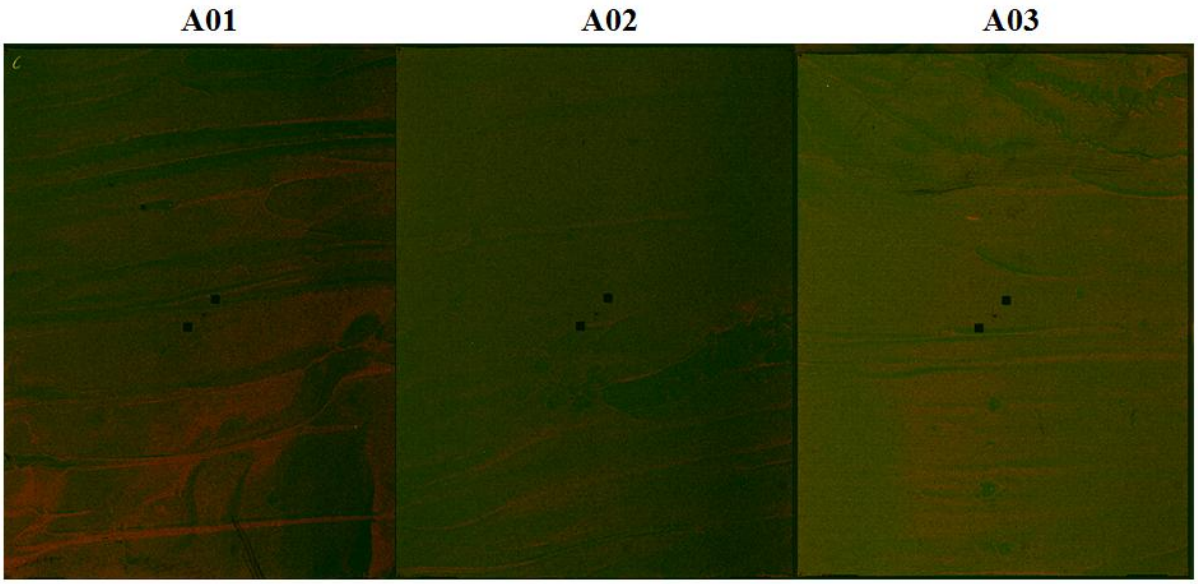


Figure 26: Scanned image of the 461641 array. The different subarrays are labelled with A01, A02 and A03. All three subarrays have an uneven distribution of Cy5, while Cy3 is distributed over the whole subarrays. Distinct red gradients and clear green areas are present at many locations on all subarrays. A01 has some black artefacts present at the bottom of the picture. A03 has the same kind of black artefacts present at the upper part of the subarray.

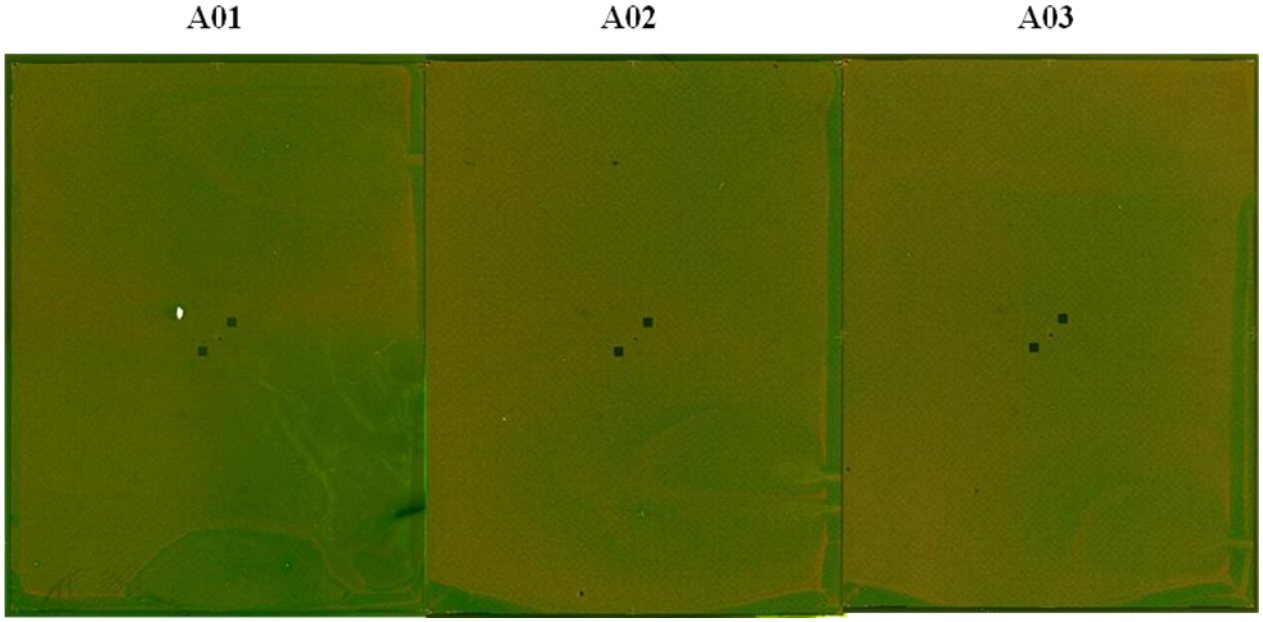


Figure 27: Scanned image from the 461678 array. The different subarrays are labelled with A01, A02 and A03. All three subarrays have a yellow colour and an evenly distribution of Cy3 and Cy5. The A01 and A03 subarrays have a red gradient present at the right side and the two bottom corners. A03 has a distinct red gradient and a clear green area at the bottom of the array. Some light green artefacts are also present.

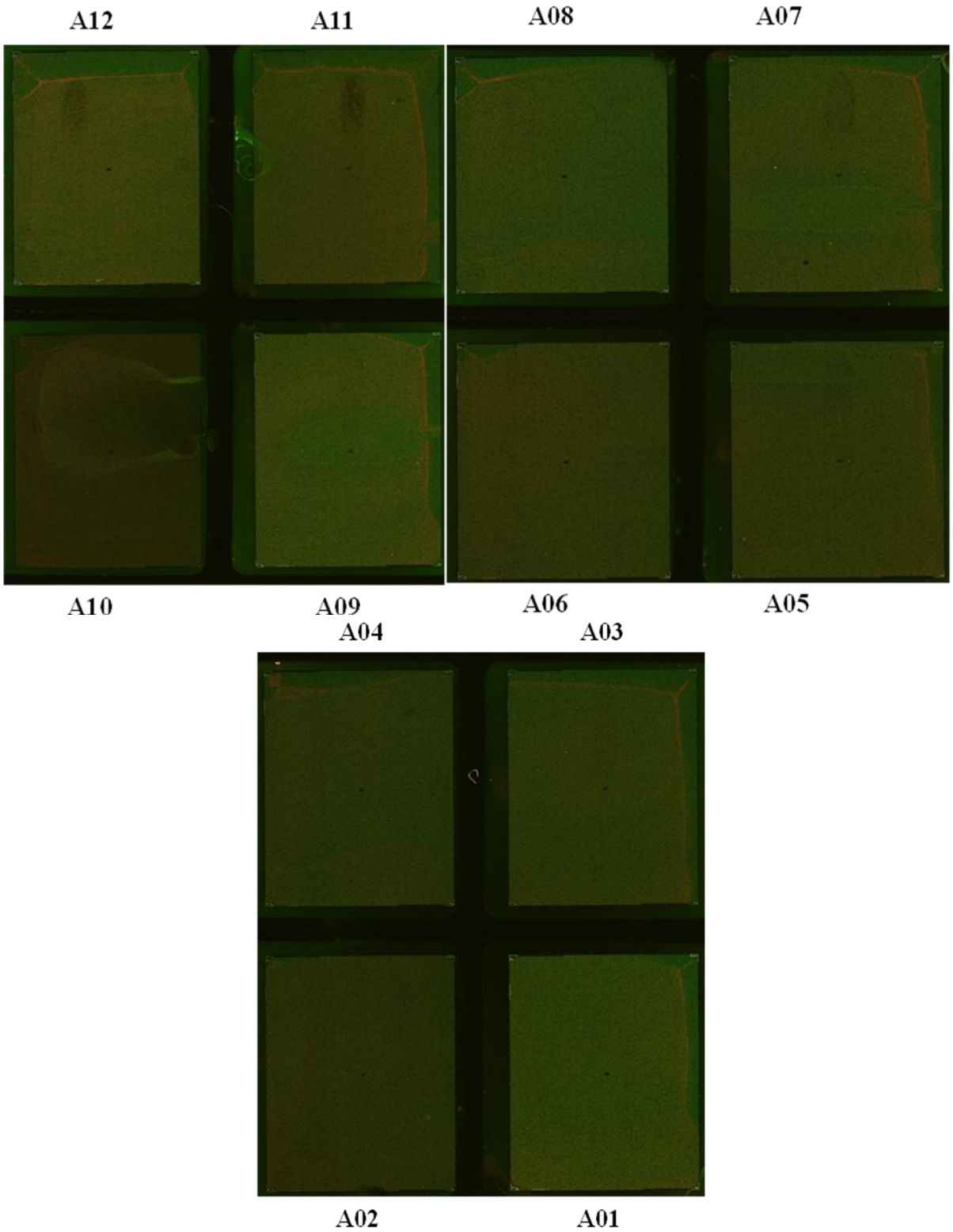
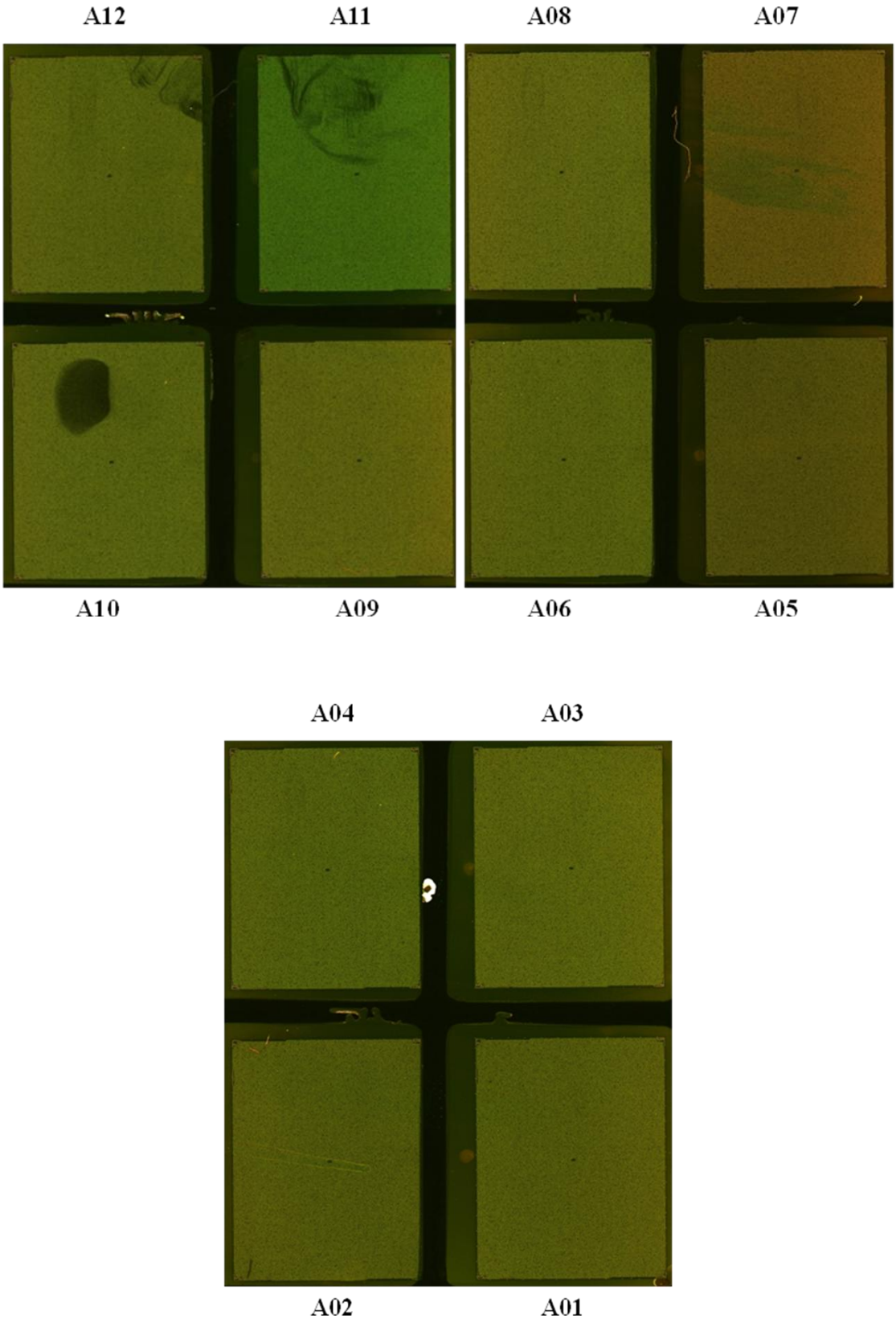


Figure 28: Scanned image with the A12-A01 subarrays from the 461424 array. A12 and A04 have a distinct red gradient at the top of the image and the array is greener above the gradient. A black mark is also present on A12. A11, A03 and A01 have a distinct red gradient at the top of the image and on the right side with green areas missing Cy5. A11 has also one black mark present. The A10 array has a darker colour than the others and has a distinct red gradient at the top left corner. One large air bubble is detected at the upper half. A09, A07 and A05 have a clear red gradient at the upper right corners and on the right side and seem to have one air bubble present. The A08 and A06 subarrays have more or less an even distribution of Cy3 and Cy5. The top left corner has a clear red gradient.



Figur 29: Scanned images of the A12-A01 subarrays of the 509338 array. All of the subarrays seems to have an evenly distribution of Cy3 and Cy5. The A11 subarray however has a more green like colour than the three others. A12 and A11 have black artifacts present at the top of the array. A10 has a black air bubble located at the upper half of the subarray. A07 has a more red like colour than the others. At the center of this subarray there is a distinct area which is more green. A02 has a black artefact at the bottom left corner.

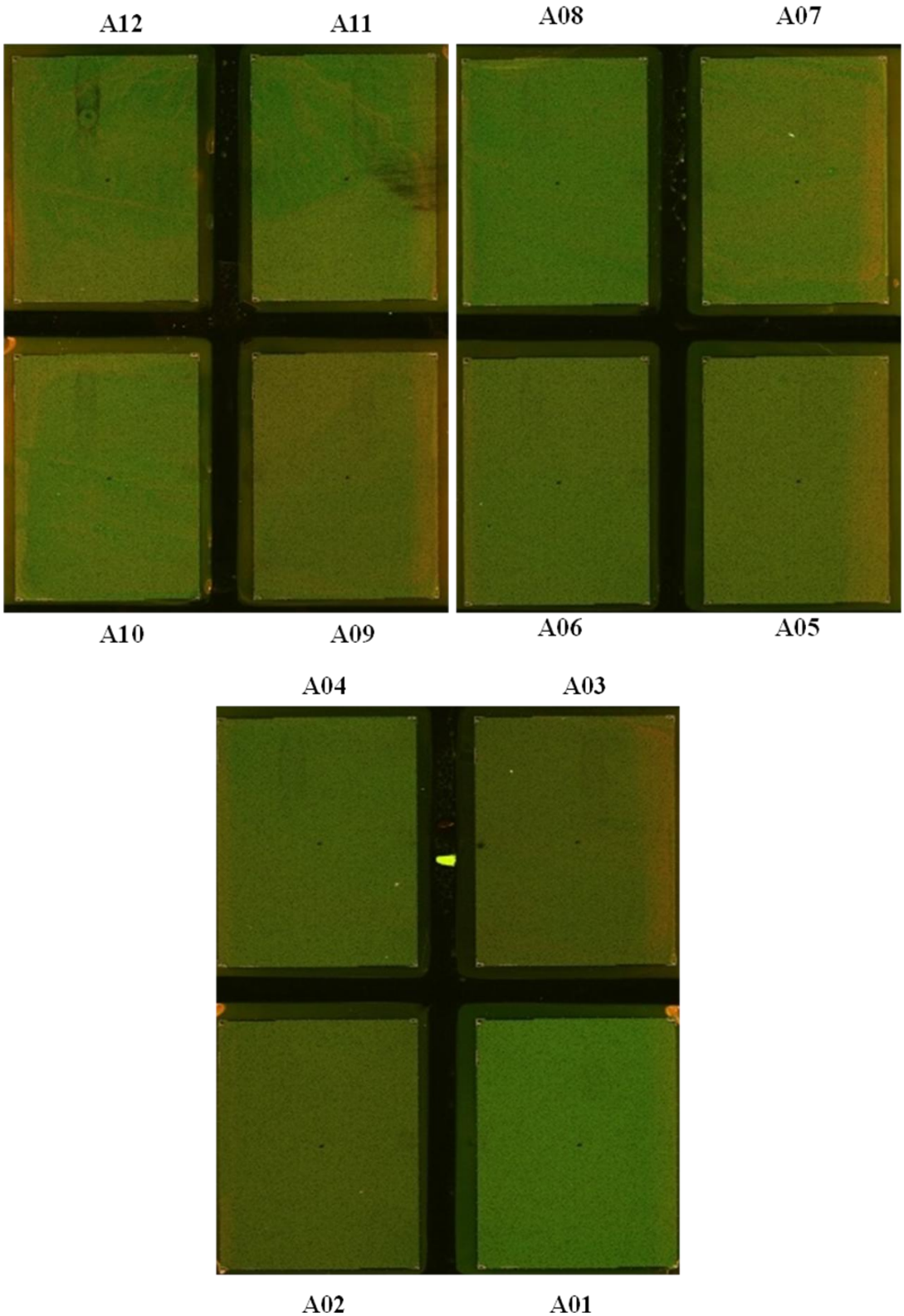


Figure 30: Scanned image of subarrays A12-A01 from array 514596.A12-A09 have an uneven distribution of Cy5. A red gradient is present at several locations. A12 and A11 have some black artefacts present. Array A08 and A06 have an evenly distribution pattern for both Cy3 and Cy5 over the whole subarray. A07 has a distinct red gradient on the right side, and over large parts of rest of the subarray the red colour are unevenly distributed. A05 has a large distinct red gradient to the right on the subarray. Over the rest of the subarray the colours seems to be evenly distributed. In A04, A02 and A01 Cy3 and Cy5 equally distributed over the whole array. A03 has a distinct red gradient on the right side of the array.

The results from array 461641 and 461424 were discarded after evaluation of the scanned images, (figures 26 and 28) due to hybridization problems with Cy5. Cy5 was absent from many areas on these arrays.

3.5.2 Processing of data from the aCGH analysis.

After scanning of the arrays data analysis was performed. Sample tracking, pair and CGH-segment reports were created as described in section 2.8 (page 47and 48). The sample tracking report was controlled to document correct loading of samples on the arrays and pair reports were used for evaluation of the signal intensity of Cy3 and Cy5. Arrays with large variations in signal intensities throughout the whole array were discarded. For arrays with acceptable intensity of Cy3 and Cy5 bioinformatics were performed as explained in section 2.9 (page 48 and 49). All detected copy number variants was screened in the DGV database and none of them were detected in this database. They were therefore considered to be potential pathogenic mutations. DGV contains information from healthy normal individuals, but it is not stated in this database whether some of the healthy individuals have had preeclampsia.

The amount of detected copy number variants was too large for presentation of all results in this section. Therefore all detected CNVs were documented in appendix A with results tables made from the html- files. The tables in appendix A (page I-XXI) includes the CNV chromosome position, probe position, gene contents and mean values. After thorough evaluation of all detected CNVs, those with the greatest interest were selected for presentation in section 3.5.3. For this assortment of CNVs, fluorescence ratio profiles and the UCSC link from the bioinformatical analysis were documented.

3.5.3 Presentation of the most interesting findings

In this section the CNVs which contain the most interesting genes documented by the aCGH analyses are presented (table 43). The CNVs and their genes were selected based on their gene contents, gene functions, and which organs the genes were expressed in. Genes

expressed in placenta, testis, ovaries and cancer were most interesting because they might be involved in pathways leading to development of preeclampsia. All CNVs contained more than one gene and an overview of the rest of the gene contents for these CNVs is presented in appendix A (page I-XXI). Some findings were detected in many of the preeclamptic patients, but with some variation regarding the start and stop position for the probes.

Table 43: Presentation of the detected CNV`s with the most interesting findings regarding their gene content. The red colour for the mean value column document that the copy number variant is a deletion and the green colour that it is a duplication.

Position	**Sample**	**Probe position**	**Gene**	**Gene function**	**Mean**
1q31.3	17P	19486999 195129999	*CFHR1*	Suggested involvement in complement and tissue invasion.	-0,41
7q11.23	17P	72019999 72409999	*FKBP36*	Member of the immunophilin protein family which participates in immunoregulation, protein folding and trafficking	-0,45
8q24.3	24P	143552499 144137499	*BAI1*	Angiogenesis inhibitor which may belong to the secretin receptor family.	0,205
8q24.3	24P	143552499 144137499	*PSCA*	Encodes for a glycosylphospahtidylinositol-anchored cell membrane protein.	0,205
15q11.1-q11.2	15P	18394999 21059999	*POTEB*	Belongs to the ankyrin domain family involved in maintenance of the cell membrane	-0,24
15q11.1-q11.2	15P	18394999 21059999	*GOLGA8E*	Encode for a golgi autoantigen	-0,24
17q21.31-q21.32	2P	41599999 42119999	*NSF*	Catalyzes the fusion of transport vesicles within the Golgi cisternae.	0,38
17q21.31-q21.32	9P	41794999 42119999	*NSF*	Catalyzes the fusion of transport vesicles within the Golgi cisternae.	-0,45
19q13.31	Mix	48162499 48299999	*PSG6*	Encodes for human pregnancy-specific glycoproteins (PSGs)	-0,24

In this section one fluorescence ratio profile is presented for all of the selected CNVs. In addition the results from the UCSCs genome browser are presented to document the involved genes and the size of the variants. Information about the UCSC database is presented in section 2.9. (page 48-49) The information presented regarding the gene functions in this section are also retrieved from the UCSC database.

CNV 17q21.31-q21.32.

The copy number variants arr17q21.31q.21.32(41599999-42119999)x1 and arr17q21.31q.21.32(41599999-42119999)x3 were documented in umbilical cord blood from fetuses of preeclampsia patients. Figures 32 and 34 presents the fluorescence ratio profiles for

samples 2P and 9P. The size of the CNV is 200 Kb for sample 2P and100 Kb for sample 9P (figure 31 and 33) and several genes are involved in these areas. According to the USCS genome browser *NSF* encodes for an N-Etylmaleimide-sensitive fusion protein which is essential in vesicle mediated transport. The function of this protein is to catalyze fusion of transport vesicles in the Golgi cisterna and between endoplasmatic reticulum and the Golgi stack. There are some indications for involvement of this protein in transportation of cargo proteins in the Golgi apparatus independent of their vesicle origin.

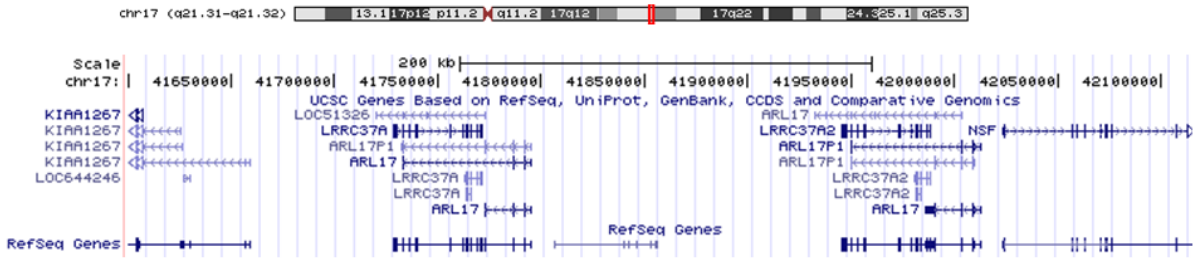


Figure 31: Illustration of genes involved in the CNV in chromosome position 17q21.31-q21.32 from sample 2P. Genes are linked to the UCSC- database and chromosomal position for each detected CNV. An illustration of the chromosome shown at the top of the figure with a red box marking the position of the CNV. Below the chromosome all genes in the detected CNV are documented. The function of these genes are presented by clicking on the gene name for the genes with a known function.

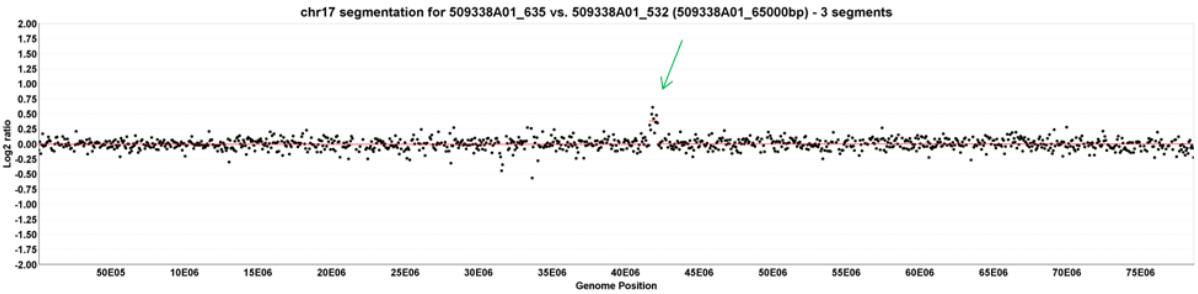
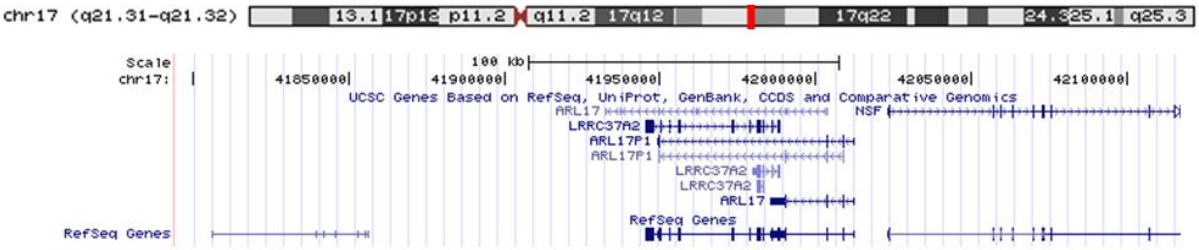


Figure 32: Fluorescence ratio profile over the CNVs detected for chromosome 17 in sample 2P analyzed with the 509338 array. To the left (y axis) the log2 ratio is presented, and on the bottom line the genome position is documented. The green arrow documents the detected CNVs position in the fluorescence ratio profile.



Figur 33: Illustration of genes involved in the CNV in chromosome position 17q21.31-q21.32 from sample 9P. Genes are linked to the UCSC- database and chromosomal position for each detected CNV. An illustration of the chromosome shown at the top of the figure with a red box marking the position of the CNV. Below the chromosome all genes in the detected CNV are documented. The function of these genes are presented by clicking on the gene name for the genes with a known function.

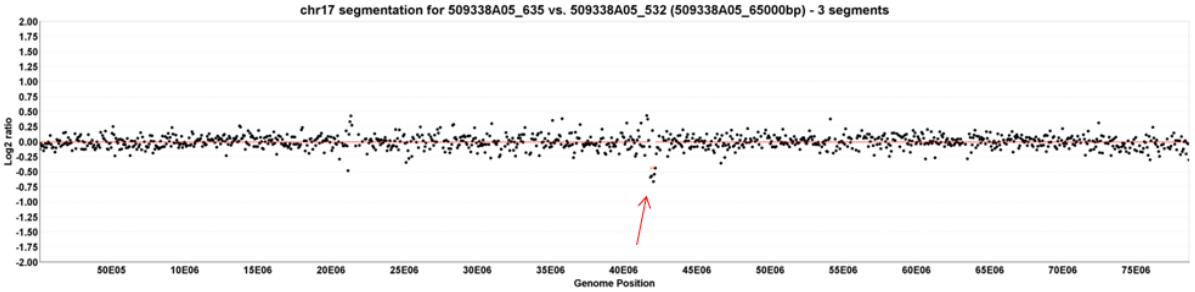
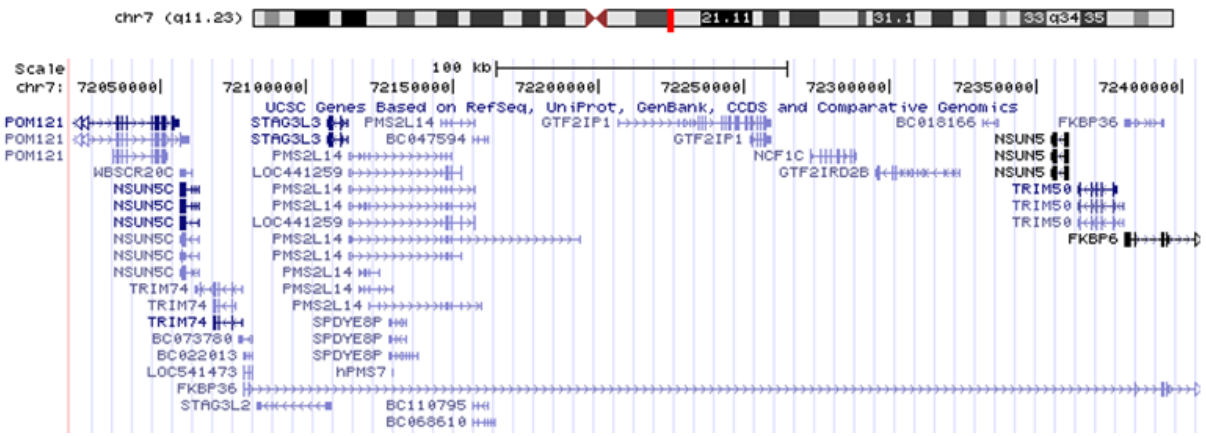


Figure 34: Fluorescence ratio profile over the CNVs detected for chromosome 17 in sample 9P analyzed with the 509338 array. To the left (y axis) the log2 ratio is presented, and on the bottom line the genome position is documented. The red arrow documents the deletion.

CNV 7q11.23

The CNV arr7q11.23(72019999-72409999)x1 was deleted in sample 17P while the CNV arr 7q11.23(72019999-72409999)x3 was duplicated in several of the fetuses of preeclamptic patients. In sample 17P the size of CNV is 100Kb and contains several genes (figure 35). The mean value for this CNV is -0.45 and figure 36 documents that the log2ratio is around 0.5 indicating one copy number variation. After examination of the documented genes by the UCSCs genome browser the *FKBP36* gene were considered to be of special interest. According to the UCSC genome browser *FKBP36* belongs to the immunophilin protein family which participates in immunoregulation and basic cellular processes such as protein folding and trafficking. The orthologous gene in mouse is essential in fertility and homologous pairing in male meiosis. In humans *FKBP36* is highly expressed in testis.



Figur 35: Illustration of genes involved in the CNV in chromosome position 7q11.23 from sample 17P. Genes are linked to the UCSC- database and chromosomal position for each detected CNV. An illustration of the chromosome shown at the top of the figure with a red box marking the position of the CNV. Below the chromosome all genes in the detected CNV are documented. The function of these genes are presented by clicking on the gene name for the genes with a known function.

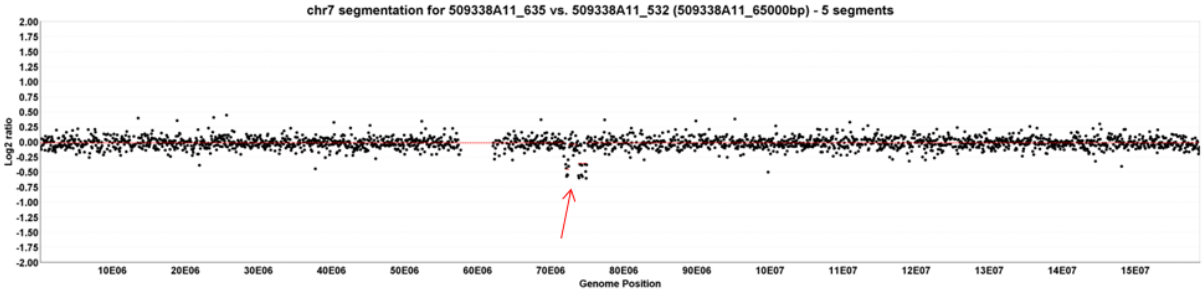
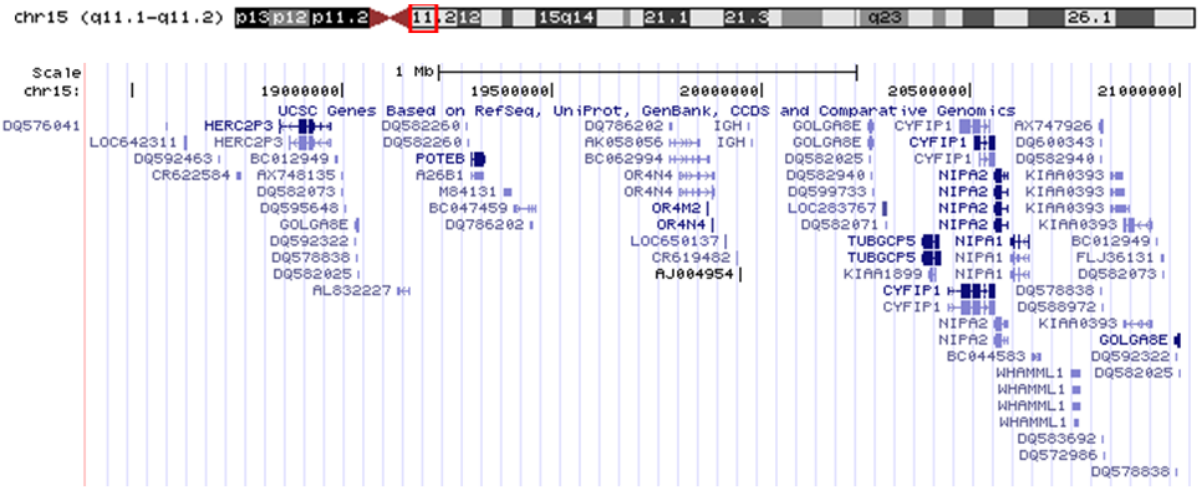


Figure 36: Fluorescence ratio profile over the CNVs detected for chromosome 7 in sample 17P with the 509338 subarray. To the left (y axis) the log2 ratio is presented, and on the bottom line the genome position is documented. The red arrow documents the deletion.

CNV 15q11.1-q11.2

The CNV arr15q11.1-q11.2(18394999-21059999)x1 was detected in umbilical cord blood of several of the fetuses of preeclamptic patients with varying mean values and log2ratios. In sample 15P this CNV involves a deleted region on 1Mb and has a mean value on -0.24. In sample 10P the mean value for this mutation is- 0.42 and the size of the mutation is 500 kb. After examination of all genes involved in the CNV in both samples, *POTEB* and *GOLGA8E* were selected for further investigations.

According to the UCSC genome browser *POTEB* is expressed in prostate, ovaries and testis, and belongs to the ankyrin domain family which are involved in the maintenance of the cell membrane (61). *GOLGA8E* is highly expressed in testis and encodes for a Golgi autoantigen.



Figur 37: Illustration of genes involved in the CNV in chromosome position 15q11.1-q11.2 from sample 15P. Genes are linked to the UCSC- database and chromosomal position for each detected CNV. An illustration of the chromosome shown at the top of the figure with a red box marking the position of the CNV. Below the chromosome all genes in the detected CNV are documented. The function of these genes are presented by clicking on the gene name for the genes with a known function.

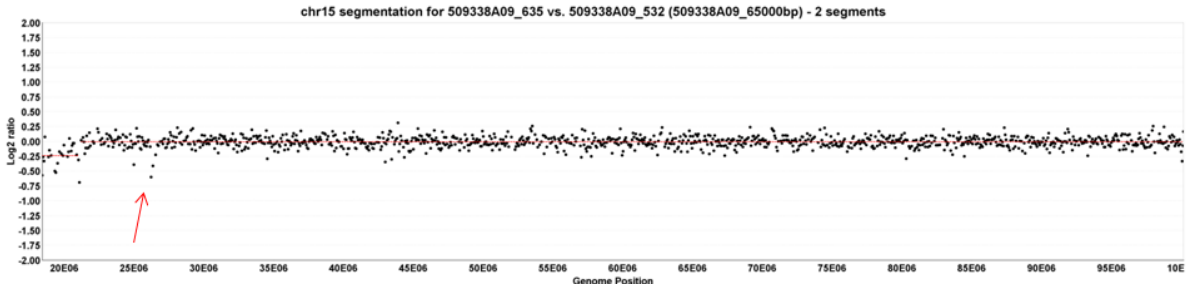
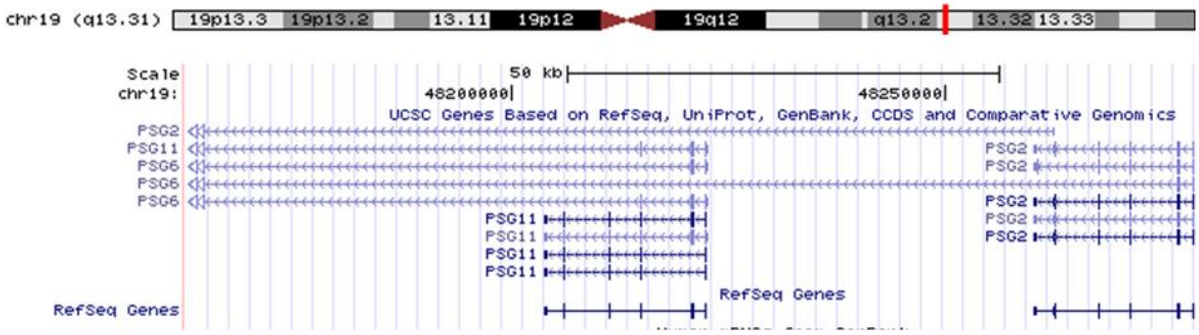


Figure 38: Fluorescence ratio profile over the CNVs detected for chromosome 15 on sample 15P analysed with array number 50933. To the left (y axis) the log2 ratio is presented, and on the bottom line the genome position is documented.

CNV 19q13.31

The CNV arr19q13.31(48162499-48299999)x1 was detected in one parallel on the 461678 array. This array contains a mixture of 19 umbilical cord blood from foetuses of preeclamptic patients analyzed in triplicate, see section 3.5. The CNV was detected in the A01 parallel. The size of the mutation is 50kb and the mean value is - 0.24. All genes involved in this CNV belong to the human pregnancy-specific glycoproteins (PSGs). According to the UCSC genome browser this group of proteins is mainly produced in large amounts by the placental syncytiotrophoblasts during pregnancy and is released into the maternal circulation. The syncytiotrophoblast are essential in placenta development. The pregnancy- specific glycoproteins includes a subgroup of the carcinoembryonic antigen (CEA) family that belongs to the immunoglobulin family.



Figur 39: Illustration of genes involved in the CNV in chromosome position 19q13.31from a mixture of nineteen preeclampsia patients. Genes are linked to the UCSC- database and chromosomal position for each detected CNV. An illustration of the chromosome shown at the top of the figure with a red box marking the position of the CNV. Below the chromosome all genes in the detected CNV are documented. The function of these genes are presented by clicking on the gene name for the genes with a known function.

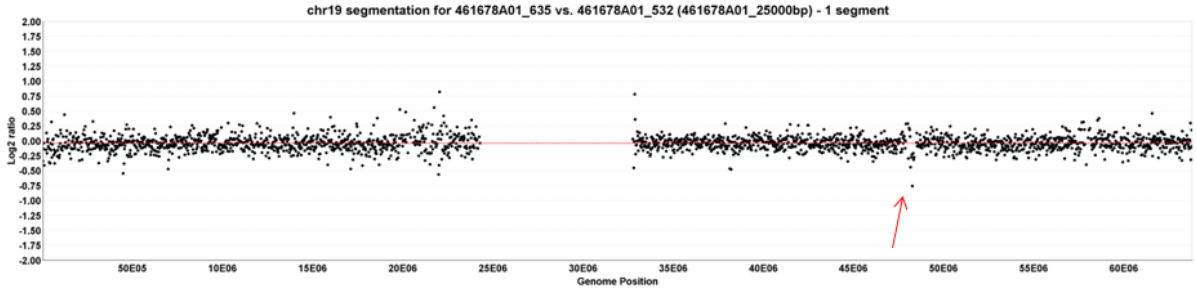
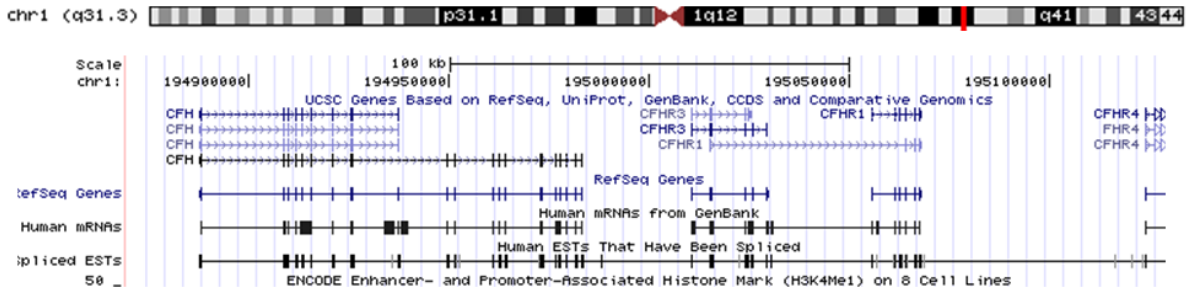


Figure 40: Fluorescence ratio profile over the CNVs detected for chromosome 19 on subarray A01 on with 461678 array. To the left (y axis) the log2 ratio is presented, and on the bottom line the genome position is documented. The red arrow documents the deletion.

CNV 1q31.3

CNV arr1q31.3(19486999-1951299999)x1 was detected in sample 17P. The size of the mutation is 100 kb and the mean value is -0.41. The genes included in this mutation are presented in figure 41. The *CFHR1* gene encodes for a protein belonging to the complement factor H protein family. In reference to the genome browser the protein binds to Pseudomonas aeruginosa elongation factor Tuf in collaboration with plasminogen. It is suggested that Tuf acts as a virulence factor by acquiring host proteins to the pathogen surface and in this way controls complement and tissue invasion. *CFHR1* is highly expressed in several organs included the fetal liver. The other genes represented in this deletion (figure 41) are involved in the complement system.



Figur 41: Illustration of genes involved in the CNV in chromosome position 1q31.3 from sample 17P. Genes are linked to the UCSC- database and chromosomal position for each detected CNV. An illustration of the chromosome shown at the top of the figure with a red box marking the position of the CNV. Below the chromosome all genes in the detected CNV are documented. The function of these genes are presented by clicking on the gene name for the genes with a known function.

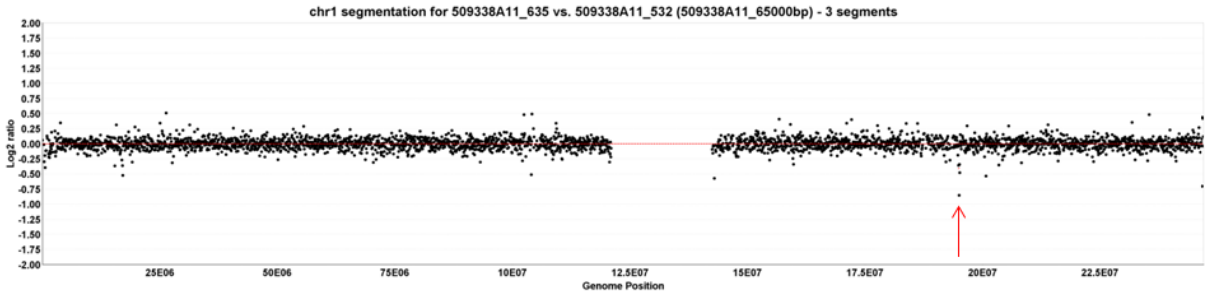
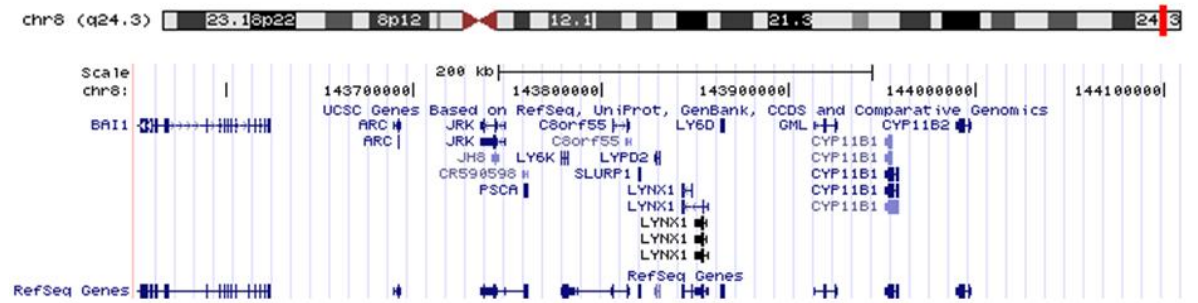


Figure 42: Fluorescence ratio profile over the CNVs detected on chromosome 1 on subarray A11 from the 509338 array. To the left (y axis) the log2 ratio is presented, and on the bottom line the genome position is documented. The red arrow documents the deletion.

CNV 8q24.3

CNV arr8q24.3(143552499-144137499)x1 was detected in sample 24P. The size of the mutation is 200kb (figure 43) and the mean value is 0.205. Of the genes involved in this mutation the *BAI1* and *PSCA* are of greatest interest. UCSC genome browser document that *BAI1* encode for a brain- specific angiogenesis inhibitor 1. *PSCA* encodes a glycosylphosphatidylinositol-anchored cell membrane glycoprotein which is highly expressed in prostate in addition to the bladder, placenta, colon, kidney and stomach. In many cases of prostate cancer this gene is up-regulated, and is also detected in bladder and pancreas cancers.



Figur 43: Documentation of the results in the USCS linked following the bioinformatics analysis for sample 24P on chromosome 8. The detected CNVs position is marked in red on the illustrated chromosome. The genes involved are presented below in blue writing. The size of the mutation are presented in black above the genes and are in this case 200 kb in size.

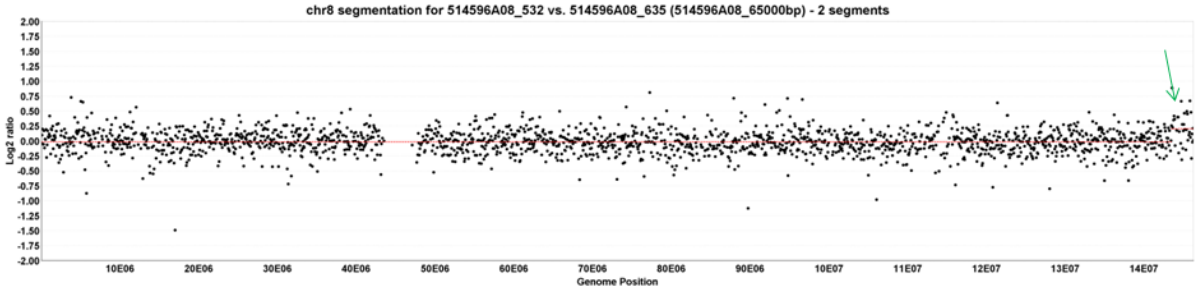


Figure 44: Fluorescence ratio profile over the CNVs detected for chromosome 8 on subarray A08 with the 514596 array. To the left (y axis) the log2 ratio is presented, and on the bottom line the genome position is documented. The green arrow documents the duplication.

3.5.4 Results from analysis of parallels

Parallels within one array and between arrays were analyzed to examine the methods reproducibility based on the hybridization problems with Cy5.

Parallels of one sample analyzed on two different arrays.

Sample 9P was analyzed on array number 509338 and 514596 and the results from these analyses are documented in table number 44 and 45. In these tables the green colour for the mean values columns documents that the detected CNVs are duplications and the red colour represent deletions. CNVs detected in both arrays are marked with an underline.

Table 44: Summary of the results from bioinformatic analysis on patient sample 9P applied on the A05 subarray with array 509338.

Position	Start position	End position	Probes	Genes	Mean	DGV
10q26.3	132014999	132404999	8		0.5	No
14q11.1-q11.2 Centromere	183949999	19304999	10		0.34	No
21q11.2-22.3	1345499	46853544	465	*The whole q arm is duplicated.*	0.44	No
17q21.31-21.32	41794999	42119999	7	*LRRC37A2,ARL17P1, ARL17,NSF*	-0.42	No

Table 45: Summary of the results from bioinformatic analysis on patient sample 9P applied on the A02 subarray with array 514596

Position	Start position	End position	Probes	Genes	Mean	DGV
9p11.2	44167499	45207499	12	*KGFLP1, AK126080, CR615666, DQ 594366, BC134347, AK131029, FAM27E3,*	0.21	No
9p11.2-11.1 centromere	46182499	46702499	7	*LOC554249, KGFLP1, AK126080, LOC643648*	0.21	No
10q26.3	132047499	132437499	7	*No genes found*	0.401	No
14q11.1-11.2 centromere	18752499	19142499	6	*BC016035, AK022914, BC041856, BX248778, BC017398, AY338954, DQ583164, DQ590589, DQ583610, A26C2, POTEG, AY338954*	0.516	No
17q21.31-q21.32	41827499	42152499	6	*ARL17, LRRC37A2, ARL17PI, NSF,*	-0, 515	No
21q11.2-22.3	13487499	46886044	464	The whole q arm is duplicated	0.48	No

Parallels of two samples run twice on the same array

Sample 12P and 21P were analyzed in duplex on array 514596 to measure the methods reproducibility within one analysis, (tables 46-49). The green colour for the mean value columns represents duplications and the red colour deletions. CNVs detected in both parallels of the samples are marked with an underline.

Table 46: Summary of the results from bioinformatic analysis on patient sample 12P applied on the A03 subarray on array 514596

Position	**Start position**	**End position**	**Probes**	**Genes**	**Mean**	**DGV**
7q11.23	74327499	74847499	6	*BC068610, BC047594, BC110795, PMS2L14, hPMS7, AL831977, DQ571357, BC068610, LOC441259, SPDYE8P, STAG3L1 ,WBSCR16*	0.214	No
8q11.21	48717499	48977499	5	*KIAA0146, CEBPD, PRKDC, AK095778,*	0.284	No
9p13.1-q13 (including several bands and the centromere)	39292499	70167499	122	*FAM75A2, DKFZp572C163, ZNF658B, KGFLP2, FAM74A3, CTNAP3B, DKFZp686115204, LOC401507, CR933660, DQ588135, DQ587539, CBWD5, CCDC29, AK054645, LOC554249, AK126080,CBWD3,FOXD4L4,++++++*	0.281	No
10q11.22	46117499	46832499	12	*BMS1P5, CR604707, SYT15 ,GPRIN2, ANXA8 ,PPYR1, ANXA8L1, LOC644054 ,AK309109, BC075841, AF130068, LOC728643, BC127745,AK3092109, AK309024,*	0.518	No
15q11.2	18427499	20247499	15	*DQ576041, CR622584, HERC2P3, AX748135, GOLGA8E ,POTEB ,DQ582073, DQ595648, DQ578838, DQ582025, A26B1, M84131, BC047459, DQ786202,IGH, OR4N4, DQ786202, AK058056, BC062994, ORAM2*	0.231	No
15q11.2	20637499	21092499	5	*NIPA1, BC044583, WHAMML1, AX747926, DQ600343, DQ582940, DQ578838, DQ588972, DQ583692, KIAA0393, BC012949, FLJ36131, DQ582073, GOLGA8E,DQ592322, DQ578838, DQ582025*	0.226	No

Table 47: Summary of the results from bioinformatic analysis on patient sample 12P applied on the A04 subarray on array 514596

Position	Start position	End position	Probes	Genes	Mean	DGV
10q11.22	46247499	47092499	12	*SYT15 , PPYR1, ANX8L1, ANXA8, LOC642826, LOC643650, GPRIN2, LOC728643, AK309109,AK309024, BC075841, AF130068, BC127745*	*0.369*	No
15q11.2	18427499	19987499	14	*LOC6446096, HERC2P3,GOLGA6L6, GOLGA8E, NBEAP1, POTEB, LOC646214, CXADRP2, LOC348120, NFIP2, LOC727924, ORAM2, OR4N4, OR4N3P, AX748135, A26B1, DQ595648, M84131, LOC650137, AJ004954 AL832227,*	*0.277*	No

Table 48: Summary of the results from bioinformatic analysis on patient sample 21P applied on the A05 subarray on array 514596

Position	Start position	End position	Probes	Genes	Mean	DGV
9p13.1-p12	39877499	40332499	7	*FAM75A2, DQ586551, DKFZp686115204,*	0.226	No
9p12-11.2	41697499	42997499	13	*MGC221881, KGFLP2, ANKRD20A2, ANKRD20A3, FOXD4L2, FOXD4L4, AQP7P3, KGFLP2, DQ588135, DQ587539, AK125850, BC064148, CBWD5, AK128231, AK310876, BC070322,BC052332,LOC554249,*	0.234	No
9p11.2	43192499	43582499	7		0.207	No
9p11.2-q12 (including several bands and the centromere)	44102499	65812499	39	*KGFLP1, AK126080, DQ594366, BC134347, AK131029, FAM27E3, CR615666, DQ594428, LOC100132167, BC134347, AK131029, LOC554249, KGFLP1, AK126080, LOC401507*	0.24	No
9q12	66787499	68412499	12	*BC070322, BC071803, AQP7P1, BC064331, FAM27E3, DQ786189, BC110369, AK308561, ANKRD20A3, ANKRD20A1 ,BC110369, CR626459, AK311167, PGM5P2*	0.231	No
14q11.1 centromere	18427499	19077499	7		0.331	No
17q21.31-q21.32	41762499	42152499	7	*LOC51326, LRRC37A,ARL17P1, ARL17, LRRC37A2, NSF, LOC51326*	-0.444	No

Table 49: Summary of the results from bioinformatic analysis on patient sample 21P applied on the A06 subarray on array 514596

Position	Start position	End position	Probes	Genes	Mean	DGV
9p11	44297499	46182499	17	*2KGFLP1, AK126080, DQ594366, BC134347, AK131029, FAM27E3, DQ594426, CR615666, LOC100132167,*	0.266	No
17q21.31-21.32	41827499	42152499	6	*ARL17, LRRC37A2, ARL17P1, NSF, LRRC37A,*	-0.323	No

3.5.5 Comparison of the 3x720K and 12X135K array formats

The summary table in appendix B (page XXI-XXIV) presents all the documented deletions and duplications found on array numbers 461678, 509338 and 514596 and compares the findings between the different arrays.

4.0 Discussion

4.1 Gradient PCR

Gradient PCR was performed on all primers used in PCR and sequencing experiments performed on the *laeverin* gene, to find the best annealing temperatures. By the use of only two PCR programs for all the primer sets the PCR experiments were more time efficient. As a consequence some of the primers were utilized at PCR programs without the best suited annealing temperature. However to be sure of correct analysis all PCR experiments were performed at annealing temperatures where the gradient PCR had documented production of distinct bands at the same size as the specific product of the primers. After thorough evaluation of the gel pictures from all gel electrophoresis analysis (figures 9-20), two PCR programs were designed. PCR program 1 had 57.8°C as annealing temperature and PCR program 2 had 65.4°C. Most of the primers produced specific bands at 57.8°C and were therefore applied in program 1 (table 39) while five of the primers were most suited for program two. For some of the primers unspecific bands were present at some or all temperatures. To avoid these, annealing temperatures where the unspecific bands were undetected were preferred. In some cases it was impossible to avoid unspecific bands because they either were present at all temperatures or were absent at the annealing temperatures utilized in program 1 and 2. None of the unspecific bands were so strong that they were considered to influence on the PCR product formed and the subsequent sequencing analysis.

4.2 Sequencing of DNA from maternal blood samples

The sequence variants detected in the sequencing analysis on DNA from maternal blood samples were documented in table 40. The variants found in exon 7 (figure 22) and 10 (figure 23) were the only variants detected exclusively in preeclampsia patients in translated exons. Further investigations detected presence of the variant in exon 10 in several normal controls and it was therefore ruled out as a possible pathogenic mutation. The variant in exon 7 was not detected in any normal controls and may therefore be a pathogenic mutation which should be further investigated. In this variant one Glutamic acid in position 487 of the laeverin

protein is replaced with Lysine and both are hydrophilic amino acids. To investigate the effect of this variant, site directed mutagenesis may be performed. Recombinants containing the variant in exon 7 could be constructed and their proteins enzyme activity may be compared to the enzyme activity of the wild type protein. Maruyama *et al* (2009) performed this type of study on one variant found in amino acid 379 in the *laeverin* gene (62).

4.3 Sequencing of DNA from placental tissue

Sample 15P collected from placental tissue was sequenced for all the exons in the *laeverin* gene and parts of the introns. The sequence variant detected in exon 6 was the only one of the variants that was not present in maternal blood for translated exons (table 41). In this variant the amino acid Lysine is replaced with Arginine. Both are hydrophilic amino acids.

In the endoplasmatic reticulum and the Golgi apparatus O-glycosylation of Lysine residues is used to mark some proteins for secretion from the cells (63). Replacement of Lysine with Arginine may result in a deficient marking of the protein and thus no secretion of the protein from the cells. Immune fluorescence studies performed by Nystad *et al* on placental tissue sections from the preeclamptic women revealed that laeverin is localized to the cytoplasm in preeclampsia patients and in the cell membrane in normal controls. This could be due to this variant. However, this variant was only present in one of the strands (figure 25). The placental tissue samples used for sequencing were collected taking care to avoid contamination with maternal blood. Therefore despite a theoretical possibility it is unlikely that one strand could have been sequenced on maternal DNA and the other one fetal DNA.

The gene sequence of the placental sample from the preeclamptic patient was compared with the gene sequences of blood samples from normal and preeclamptic women, but not with the gene sequence of normal placenta, maternal decidua or the fetal blood. Therefore, we do not know whether similar variant can be present in the normal placenta, maternal decidua or the fetal blood. The variant was not detected in maternal blood from the same preeclamptic patient that the placenta sample was collected from. For further investigations several samples from placental tissue from preeclamptic patients and normal controls should be screened for this variant. If it is detected in several samples from preeclamptic patients and not in any

normal controls this could be an indication of a potential pathogenic mutation. If the variant is not detected in any normal controls or more preeclamptic samples it may be a familial variant.

4.4 Evaluation of scanned arrays

Array number 461641

All three subarrays have an uneven distribution of Cy5 (figure 26). Some positions on the subarrays have distinct red gradients because more of the Cy5 has hybridized to the probes than to the other positions. The Cy3 is evenly distributed for all subarrays. Due to the poor hybridization of Cy5, the signal intensity for the labelled DNA with this dye was too low for obtaining correct results.

Array number 461678

The same DNA mixture as the one analyzed with array number 461641 were re-analyzed with array number 461678 with a new labelling kit. The data analyses of 461678 shows some variation between the A01, A02 and A03 parallels and are presented in appendix A (page I-IV). Ideally the three subarrays should show exactly the same regions duplicated and deleted since they contain the same sample material. A03 has a more uneven distribution of Cy5 than A01 and A02 (figure 27). At the bottom right corner there are some light green artefacts present and to the left some black artefacts. A01 and A02 have an even distribution of the two colours.

Array number 461424

Array 461424 has some of the same problems as 461641 regarding hybridization of Cy5 (figure 28). The signal intensity of Cy5 was too low for obtaining correct results and the analysis was excluded.

Array number 509338

All subarrays except for one have an even distribution pattern of Cy3 and Cy5 (figure 29). Subarray A10 contains one large air bubble and A11 and A12 have some black artefacts present. According to the protocol from NimbleGen these black marks could be washing artefacts. The Cy5 and Cy3 may have been damaged by the components in the washings

solutions in these areas. Since this array has an even distribution of Cy5 further analyses were performed on this array, and are presented in appendix A (page IV-IX).

Array number 514596

The distribution pattern of Cy5 varies between the different subarrays. A12-A09 have many clear green areas where the red colour is absent. The results from data analyses of these subarrays are presented in appendix A (page XIV-XX). The A12- A09 subarrays have a much higher number of duplications present compared to the others subarrays and arrays. Large parts of these four subarrays are mostly green, and seem to lack hybridization of Cy5 (figure 30). The normal controls were labelled with Cy5 and when these are absent due to poor hybridization the patient`s genome will be presented with a large amount of duplications and the results will be erroneous.

Hybridization problems

Poor hybridization of Cy5 was a problem for several of the arrays presented above. Cy3 and Cy5 were handled in the same way and in accordance to the protocol in all steps in the aCGH analysis. The hybridization reaction is sensitive to several conditions such as, salt concentrations, temperature, mixing of samples prior to loading, amount of labelled DNA and the duration of the hybridization reaction. High salt concentration improves the labelled DNAs ability to hybridize to the probes on the array. Temperatures influence on the stringency of hybridization, and high temperatures in the hybridization reaction and washing steps makes the hybridization reaction more stringent and reduce the amount of unspecific binding. Cy3 and Cy5 were handled under the same salt and temperature conditions and the same quantity of DNA labelled with Cy5 and Cy3 were combined before hybridization. Therefore both colours should have been evenly distributed. Before application on the arrays the patients samples were combined with normal controls and mixed and vortexed thoroughly. Cy5 was only present at some locations on the arrays and may indicate that it was more viscous than Cy3. However according to the manufactures both colours have the same viscosity. There could be some problems with the Cy5 fluorochrome but it doesn't explain why it is present in large quantities in some areas and are absent from others. If the flurochrome weren't properly attached to the nucleotide, Cy5 should be absent from all locations and not present in large quantities in some positions. aCGH is sensitive to DNA quality. After labelling all samples were quantified with

NanoDrop, and all samples had high concentration of labelled DNA, therefore the sample quality is not a reasonable explanation to the poor hybridization. The high quantity of the labelled DNA also rules out low fluorescence intensity from Cy5 as a possibility. In addition some parts of the subarrays have strong red signals indicating that the intensity from the fluorochrome is good enough. The problem could be due to poor mixing of the solutions in the hybridization instrument. However if the problem were poor mixture of the solutions during hybridization there should be an uneven distribution of both Cy3 and Cy5 and not only Cy5. The hybridization reaction was performed within the range of the recommended time.

4.5 Interesting findings

The most interesting findings from the aCGH analysis were presented in section 3.5.3 (page 69-76)

CNV 17q 21.31-q21.32.

The copy number variant arr17q21.31q.21.32(41599999-42119999)x1 was detected in the fetal genome of one preeclampsia patient while arr17q21.31q.21.32(41599999-42119999)x3 was documented in several. The fetal genome is closely related to placental genome and can influence maternal physiology through a variety of biological mechanisms. Previous research by Nystad *et al* documented that the transportation of laeverin protein from the Golgi apparatus to the cell membrane may not be functioning properly in preeclamptic placentas (un-published observations). From these experiments it is not clear whether laeverin is retained in the Golgi apparatus or if it is localised in other cell compartments. To clarify whether laverin is retained in the Golgi apparatus in preeclamptic placentas electron microscopic experiments needs to be performed. *NSF* is involved in transportation within the Golgi apparatus and it is therefore possible that there are some interacting effects between *laeverin* and *NSF*. If *NSF* is required for transportation of laeverin from the Golgi apparatus to the cell membrane this might explain the depletion of laeverin in the cell membrane. Reduction of the amount of the protein encoded by *NSF* due to deletions of the *NSF* gene may result in retention of laeverin. *NSF* was deleted in one preeclampsia patient and duplicated in several patients. Duplications and deletions may alter the protein`s structure and thereby its

function. If the protein function is altered then its transportation capacity may be reduced which can lead to depletion of laeverin in the cell membrane and development of preeclampsia.

CNV 7q11.23

The CNV arr 7q11.23(72019999-72409999)x1 was present in the fetal genome of one preeclamptic patient while the variant arr 7q11.23(72019999-72409999)x3 was documented in several. The orthologous gene to *FKBP36* in mouse is essential in fertility and homologous pairing in male meiosis and F*KBP36* may therefore be involved in fertility in humans. Documentation of high expression levels in the testis in human males is an indication for this. It's belonging to the immunophilin protein family which are involved in immunoregulation and basic cellular processes such as protein folding and trafficking makes it even more interesting. Contribution of the immune system in development of preeclampsia has been indicated in several studies. As for *NSF* its role in trafficking may be linked to depletion of laeverin in the cell membrane. Its function in protein folding is also interesting since incorrect folded proteins are not transported from the Golgi apparatus. *FKBP36* was only deleted in the umbilical cord blood sample from the newborn of a preeclamptic mother. Deletion of this gene might be a familial variant. Another possibility could be that both deletions and duplications of this gene influence protein folding and alter the protein function.

CNV 15q11.1-q11.2

The CNV arr15q11.1-q11.2(18394999-21059999)x1 was deleted in several umbilical cord blood samples from the neonates of preeclamptic patients. Expression of *POTEB* in prostate, ovaries and testis implies for involvement of this gene in fertility, and it may therefore be involved in processes leading to preeclampsia. *POTEB* belongs to the ankyrin family which is involved in the maintenance of cell membranes. Ankyrins are important for the correct localization of the ion channels in the cell membranes (61). Since the deletion is detected in several neonates of preeclamptic patients the protein product may be reduced resulting in malformations of cell membranes in important cells in the placenta development. If this gene is involved in maintenance of cell membranes in extravillous trophoblast the deletion may alter their cell membranes resulting in poor placentation. However, so far ankyrin has not

been characterized in extravillous trophoblast or other cells involved in the placental development. This should therefore be further investigated.

GOLGA8E is highly expressed in testis and encodes for a Golgi autoantigen. Since it is present in the Golgi apparatus it might be in involved in modifications of transport of proteins.

CNV 19q13.31

The CNV arr19q13.31(48162499-48299999)x1 involved genes that encodes for human pregnancy-specific glycoproteins (PSGS) secreted by the syncytiotrophoblasts. Syncytiotrophoblasts participate in the formation of the villus tree and are therefore essential for the utero placental circulatory system and formation of the placenta. The influence of the (PSGS) proteins in placental development should therefore be examined. One possibility is that a reduced amount of these glycoproteins may induce alterations which lead to development of preeclampsia. If these proteins are necessary for differentiation or formation of syncytiotrophoblasts a reduced amount may lead to malformation and poor development of the placenta. The maternal symptoms in preeclampsia are believed to be induced by factors released from the hypoxic placenta. If these pregnancy-specific glycoproteins are essential for the development of syncytiotrophoblasts then a reduced amount could lead to malformation of the placenta which in turn releases factors into the circulation and the maternal symptoms develops. In addition, these genes and their products involvement in the immune system is an interesting factor. dfvf

CNV 1q31.3

The CNV arr1q31.3(19486999-1951299999)x1 was detected in one umbilical cord blood sample from the neonate of one preeclamptic patient and involves several genes associated with the complement system. The *CFHR1* gene encodes for a protein belonging to the complement factor H protein family and may be involved in control of complement and tissue invasion. Since the development of the placenta and establishment of the utero placental circulatory system is dependent on invasive EVTs, deletions of this gene may have negative influence on the EVT invasion. *CFHR1* is highly expressed in several organs including the

fetal liver. Fetal genes may contribute to the development of preeclampsia and these findings are therefore interesting. The other genes represented in this deletion (figure 41) are also associated with complement system and may have some functions related to EVT invasions. There may be some interacting affects between these and other genes involved in tissue invasion.

CNV 8q24.3

CNV arr8q24.3(143552499-144137499)x3 contains one gene that encodes a brain-specific angiogenesis inhibitor. Preeclampsia is characterized by poor development of the early placenta and the maternal blood supply and angiogenesis inhibitors may therefore be involved. Duplications of *BAI1* could cause an overproduction of the angiogenesis inhibitor, resulting in reduced remodelling of the spiral arteries and thereby poor placentation. Tumor cells are dependent on angiogenesis for their growth and metastasis. Increased secretion of angiogenesis inducers and decreased production of inhibitors promotes angiogenesis. The protein encoded for by *PSCA* is also involved in cancer and is expressed in the placenta. These two genes should therefore be further investigated and confirmed by a second independent method

4.6 Evaluation of the analyzed parallels

Between runs

For sample 9P the cytobands 10q26.3, 14q11.1-q11.2, 21q11.2-q22.3 and 17q21.31-q21.32 were present on both arrays (table 44 and 45). In addition 9P analyzed on 514596 has detected the cytobands, 9p11.2, and 9p11.2-11.1. The chromosomal position 21q11.2-q22.3 involves trisomy 21 and is detected in both analyses. Since both aCGH analysis detects almost the same findings and taken into consideration that this is a screening method where findings should be verified by a second independent method, the reproducibility between runs are considered to be good enough. In addition the detected trisomy 21 has been verified by G-banding performed by the staff at the Divison of Child and Adolescent Health, Department of Medical Genetics, University Hospital of North Norway, Tromsø, Norway.

Within runs

The bioinformatical analysis of parallels of the samples 12P and 21P have varying results. (tables 46-49). Only some cytobands are present in both parallels. For sample 12P the Cytobands 10q11.22 and 15q11.2 are found in both parallels (tables 46 and 47). Cytobands 7q11.23,8q11.21,9q13.1-q13 are found in only the first parallel analyzed on subarray A03 not in the one analyzed on A04. The Cytobands 17q21.31-q21.32 is the only one in common for both parallels for sample 21P (tables 48 and 49). Cytoband 9p13.1-p12, 9p11.2, 9p11.2-q12, 9q12 and 14q11.1 are all detected in only one of the parallels.

The scanned images of A03-A06 (figure 30) which represents these samples seem to have an even distribution of both Cy3 and Cy5 for most parts of the arrays. A05 and A03 have both a distinct red gradient at the right side of the subarrays. Probes located in these areas could give false results and may to some extent contribute to the variation. However probes for the different genes are placed at several areas on the arrays and the red gradient is only present at the right side of the subarrays, and does not explain the degree of variation.

The variation between the parallels within one analysis is quite large, and ideally it should be lower. However, aCGH is a screening method and the results should always be verified by a second independent method. If the copy number variants detected by the aCGH analyses are confirmed by a second independent method then they are considered to be reliable.

4.7 Comparison of findings on the 3X720K and 12X135K arrays.

The 3x720K arrays were used as a pre-screening to investigate if some of the preeclampsia samples contained genes which could be involved in preeclampsia and to examine if it was possible to analyze mixtures of samples on one subarray. The summary table in appendix B (page XXI-XXIV) presents all variants detected for all arrays and documents the variation between the different arrays. Several of the CNVs detected by the 12X135K arrays are not represented in the 3X720K arrays and vice versa. The variation between the 12X135K arrays and the 3X720K arrays might be explained by resolution and sample material. The 3X720K

array format has a higher resolution than the 12X135K arrays and when single samples are run more CNVs should be detected with these arrays. However when a pool of 19 samples was analyzed the amount of DNA from each sample was much smaller than when one single sample was analyzed with the 12X135K arrays. Some CNVs might therefore be undetected. Therefore some of the copy number variants present on the 12X135K arrays might be undetected on the 3X720K arrays. In addition a large amount of the variants detected in the 12X135K arrays are duplications in array number 514596. As presented earlier the Cy5 had an uneven distribution for many of the subarrays, resulting in detection of many duplications. It is therefore likely that many of these findings are not real aberrations, and they should be evaluated by a second method. The fact that some of the copy number variations are documented in both array formats demonstrates that pooling of samples was possible.

4.8 Further investigations

All detected copy number variations should be verified by a second independent method. Most of the detected CNVs were too small to be detected by traditional G-banding. However FISH and qPCR are well suited methods for confirmation. In addition to copy number variation FISH analysis detects inversions and translocation that are undetected by aCGH or qPCR. FISH analysis is a more time consuming method compared to qPCR, where several samples could be analyzed in one run. If the copy number variations are confirmed, further investigations regarding their contribution to preeclampsia ought to be performed.

To investigate if the genes presented in section 3.5.3 (page 69-76) may contribute to preeclampsia gene expression profiles from fetal DNA should be performed and compared to gene expression profiles from fetal DNA from normal pregnancies. Western blot may also be used to analyze proteins encoded by these genes in preeclampsia patients. Experiments with antibodies directed against the proteins in question may be used to detect their distribution pattern in the cells and to gain insight into the amount of produced proteins. To investigate if the genes are essential in placental development, cell culture experiments where these genes are knocked-out by siRNA may be used. The methods presented above should be performed on a large number of preeclamptic patients to

investigate if there might be a common link between them or if some of the CNVs could be a familial variants.

NSF and *FKBP36* may be involved in transportation of laeverin to the cell membrane and the samples used in the immunofluorescense study performed by Nystad *et al* ought to be investigated for deletion or duplications of these genes to investigate if there is a link between these findings. Deletions of *FKBP36* and *NSF* could result in depletion of the protein products and thereby result in too few proteins for transportations. Duplications and deletions may alter protein structure and function, and in this way produce abnormal proteins.

All documented CNVs contain more than one gene and there might be interacting effects between them which should be taken into consideration in the interpretation of results. In addition many of the genes are presented with unknown functions and some of these may be essential in the disease development.

5.0 Conclusions

Two detected variants in the *laeverin* gene were found exclusively in preeclamptic patients in the sequencing analysis. The effect of these variants on enzyme activity should be examined. The variant detected in exon 6 is of special interest since Lysine is replaced with Arginine. O-glycosylation of Lysine residues marks some proteins for secretion from the cells. Replacement of Lysine may therefore result in retention of *laeverin* in the Golgi apparatus and may explain the depletion of laverin in the cell membrane of throphoblasts in preeclampsia patients.

The aCGH analysis revealed copy number variations for genes involved in vesicle transport, protein folding, maintenance of cell membranes, production of pregnancy-specific glycoproteins, immunological reactions, complement system, tissue invasion, angiogenesis inhibitors and cancer. The ability of aCGH to screen the whole genome in one single test makes it less time consuming and provides a large amount of information compared to other traditional methods. Using this method we detected several genes which may be linked to the pathogenesis of preeclampsia. However these findings should be confirmed by a second independent method before their potential roles in the development of preeclampsia can be further investigated.

6.0 References

1. Griffiths JFG, Wessler SR, Lewontin RC, Caroll SB. Introduction to genetic analysis. 9th edition ed. New York: FREEMAN; 2008.3-5.
2. Reece RJ. Analysis of Genes and Genomes. Manchester: Wiley; 2006.1,7,17,24-27,43,59-60,88-94,153-162,179-181,289,296-303.
3. Nussbaum RL, McInnes RR, Huntington FW. Thompson & Thompson Genetic In Medicine. 7 ed. Philadelphia: Sunders Elsevier; 2007.16,59-61,70-75,175-179,533,542.
4. Founds SA, Dorman JS, YP C. Microarray Technology Applied to the Complex Disorder of Preeclampsia. JOGNN. 2008;37:146-57.
5. Shaffer LG, Bejjani BA. A cytogeneticist`s perspective on genomic microarrays. Human Reproduction Update. 2004;10(3):221-6.
6. Shaffer LG, Bejjani BA. Medical applications of array CGH and the transformation of clinical cytogenetics. Cytogenetic and Genome Research. 2006;115:303-9.
7. Strachan T, Read AP. Human Molecular Genetics. 3th ed: Garland Science; 2004.46-49.
8. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequencs by multiplex ligation-dependent probe amplification. Nucleic Acids Research. 2002;30(12):1-13.
9. Wells D, Levy B. Cytogenetics in reproductive medicine: the contribution of comparative genomic hybridization (CGH). BioEssays. 2003;25.3:289-300.
10. Choy KW, Setlur SR, Lee C, Lau TK. The impact of human copy number variation on a new era of genetic testing. BJOG; An international Journal of Obstetrics and Gynaecology. 2010:1-7.
11. D'Haene B, Vandesompele J, Hellemans J. Accurate and objective copy number profiling using real-time quantitative PCR. Methods. 2010 Apr;50(4):262-70.
12. Edelmann L, Hirschhorn K. Clinical utility of Array CGH for the detection of chromosomal imbalances associated with mental retardation and multiple congenital anomalies. Annual New York Academy of Science. 2009;1151:157-66.
13. NimbleGen. Human CGH 3X720K Whole-Genome Tiling Arrays. 2011 [20.11.11]; Available from: http://www.nimblegen.com/products/cgh/wgt/human/3x720k/index.html.
14. Van den Veyer IB, Beuadet AL. Comparative genomic hybridization and prenatal diagnosis. Current Opinion in Obstetrics and Gynecology. 2006;18:185-91.
15. Beaudet AL, Belmont JW. Array-based DNA diagnostics: let the revolution begin. Annual Reviews in Medicine. 2008;59:113-29.
16. Gresham D, Dunham MJ, Botstein D. Comparing whole genomes using DNA microarrays. Nature Reviews. [Journal Article]. 2008;9:291-302.
17. Harper JC, Harton G. The use of arrays in preimplantation genetic diagnosis and screening. Fertility and Sterility. [Journal Article]. 2010;94(4):1173-7.
18. Gude NM, Roberts CT, Kalionis B, King RG. Growth and function of the normal human placenta. Thrombosis Research. 2004;114:397-407.
19. Larsen WJ. Human Embryologi. third edition ed. Livingstone C, editor. United States of America2001.1-3,20-22,37-45,481-487.
20. Huppertz B. The anatomy of the normal placenta. J clin pathol. 2008;61:1296-302.
21. James JL, GS W, JE C. Pre-eclampsia:Fitting together the placental,immune and cardiovascular pieces. Journal of Pathology. 2010;221:363-78.

22. Kliman HJ. The Story of Decidualization, Menstruation, and Trophoblast Invasion. American Journal of Pathology. 2000;157:1759-68.
23. Sitras V, Paulssen RH, Grønaas H, Leirvik J, Hanssen TA, Vårtun Å, et al. Differential Placental Gene Expression in Severe Preeclampsia. Placenta. 2009;30:424-33.
24. Sibai BM. Diagnosis and management of gestational hypertension and preeclampsia. Obstet Gynecol. 2003 Jul;102(1):181-92.
25. Redman CW, IL S. Latest Advances in Understanding Preeclampsia. Science. 2005;308:1592-4.
26. ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. Obstet Gynecol. 2002 Jan;99(1):159-67.
27. Nejatizadeh A, Stobdan T, Malhotra N, Pasha MAQ. The genetic Aspects of Pre-eclampsia: Achievements and limitations. Biochem Genet. 2008;46:451-79.
28. Bdolah Y, Karumanchi SA, Sachs BP. Recent Advances in Understanding of Preeclampsia. Croat Med J. 2005;46(5):728-36.
29. Mütze S, Schöneborn-Rudnik S, Zerres K, Rath W. Genes and the preeclampsia syndrome. JPerinat Med. 2008;36:38-58.
30. Lachmeijer AMA, Dekker GA, Pals G, Aarnoudse JG, Kate LPT, Arngrimmson R. Searching for preeclampsia genes:the current position. European Journal of obstetrics & Gynecology and Reproductive Biology. 2002;105:94-113.
31. Bezerra PCFM, Leao MD, Queiroz JW, Melo EMD, Pereira FVM, Nobrega MH, et al. Family history of hypertension as an important risk factor for the development of severe preeclampsia. Acta Obstetricia et Gynecologia. 2010:612-7.
32. Harrison GA, Humphrey KE, Jones N, Badenhop R, Guo G, Elakis G, et al. A Genome Wide Linkage Study of Preeclamspia/Eclampsia Reveals Evidence for a Candidate Region on 4q. AmJHumGenet. 1997;60:1158-67.
33. Wiliams PJ, Pipkin FB. The genetics of pre-eclampsia and other hypertensive disorders of pregnancy. Best practice & research Clinical obstetrics and Gynaecology. 2011:1-13.
34. Mohaupt M. Molecular aspects of preeclampsia. Molecular Aspects of preeclampsia 2007;28:169-91.
35. Arngrimsson R, Sigurard ttir S, Frigge ML, Bjarnad ttir RI, Jonsson T, Stefansson H, et al. A genome-wide scan reveals a maternal susceptibility locus for pre-eclampsia on chromosome 2p13. Hum Mol Genet. 1999 Sep;8(9):1799-805.
36. Moses EK, Lade JA, Guo G, Wilton AN, Grehan M, Freed K, et al. A genome scan in families from Australia and New Zealand confirms the presence of a maternal susceptibility locus for pre-eclampsia, on chromosome 2. Am J Hum Genet. 2000 Dec;67(6):1581-5.
37. Lachmeijer AM, Arngrimsson R, Bastiaans EJ, Frigge ML, Pals G, Sigurdardottir S, et al. A genome-wide scan for preeclampsia in the Netherlands. Eur J Hum Genet. 2001 Oct;9(10):758-64.
38. Laivuori H, Lahermo P, Ollikainen V, Widen E, Mällinen-Häivä L, Sundström H, et al. Susceptibility Loci for Preeclampsia on Chromosome 2p25 and 9p13 in Finnish Families. AmJHumGenet. 2003;72:168-77.
39. Oudejans CBM, Mulders J, Lachmeijer AMA, Dijk MV, Könst AAM, Westerman BA, et al. The parent-of-origin effect on 10q22 in pre-eclamptic females coincides with two regions clustered for genes with down-regulated expression in androgenetic placentas. 2004;10:589-98.

40. Iglesias-Platas I, Monk D, Jebbink J, Buimer M, Boer K, van der Post J, et al. STOX1 is not imprinted and is not likely to be involved in preeclampsia. Nat Genet. 2007 Mar;39(3):279-80; author reply 80-1.
41. Rigourd V, Chauvet C, Chelbi ST, Rebourcet R, Mondon F, Letourneur F, et al. STOX1 overexpression in choriocarcinoma cells mimics transcriptional alterations observed in preeclamptic placentas. PLoS One. 2008;3(12):e3905.
42. van Dijk M, Mulders J, Poutsma A, Konst AA, Lachmeijer AM, Dekker GA, et al. Maternal segregation of the Dutch preeclampsia locus at 10q22 with a new member of the winged helix gene family. Nat Genet. 2005 May;37(5):514-9.
43. Yu L, Chen M, Zhao D, Yi P, Lu L, Han J, et al. The H19 gene imprinting in normal pregnancy and pre-eclampsia. Placenta. 2009 May;30(5):443-7.
44. Bourque DK, Avila L, Penaherrera M, von Dadelszen P, Robinson WP. Decreased placental methylation at the H19/IGF2 imprinting control region is associated with normotensive intrauterine growth restriction but not preeclampsia. Placenta. 2010 Mar;31(3):197-202.
45. Enquobahrie DA, Qiu C, Muhie SY, Williams MA. Maternal peripheral blood gene expression in early pregnancy and preeclampsia. Int J Mol Epidemiol Genet. 2011 Jan 1;2(1):78-94.
46. Fujiwara H, Higuchi T, Yamada S, Hirano T, Sato Y, Nishioka Y, et al. Human extravillous trophoblasts express laeverin, a novel protein that belongs to membrane-bound gluzincin metallopeptidases. Biochem Biophys Res Commun. 2004 Jan 23;313(4):962-8.
47. Fujiwara H. Membrane-bound peptidases regulate human extravillous trophoblast invasion. Placenta. 2007 Apr;28 Suppl A:S70-5.
48. Petrovic N, Schacke W, Gahagan JR, O'Conor CA, Winnicka B, Conway RE, et al. CD13/APN regulates endothelial invasion and filopodia formation. Blood. 2007 Jul 1;110(1):142-50.
49. Maruyama M, Hattori A, Goto Y, Ueda M, Maeda M, Fujiwara H, et al. Laeverin/aminopeptidase Q, a novel bestatin-sensitive leucine aminopeptidase belonging to the M1 family of aminopeptidases. J Biol Chem. 2007 Jul 13;282(28):20088-96.
50. Goto Y, Yoshioka R, Arisaka N, Hattori A, Tsujimoto M. Involvement of glutamine-238 in the substrate specificity of human laeverin/aminopeptidase Q. Biol Pharm Bull. 2011;34(1):24-7.
51. Qiagen. DNA blood Mini Kit (50). 2011 [24.11.11]; Available from: http://www.qiagen.com/products/genomicdnastabilizationpurification/qiaampsystem/qiaampdnabloodminikit.aspx#Tabs=t1.
52. Qiagen. EZ1 DNA blood 350 Kit. 2011 [24.11.11]; Available from: http://www.qiagen.com/products/genomicdnastabilizationpurification/biorobotez1system/ez1dnablood350ulkit.aspx#Tabs=t2.
53. NanoDrop. NanoDrop 2000. 2011 [25.11.11]; Available from: http://www.nanodrop.com/Productnd2000overview.aspx.
54. Gallagher S, Desjardins P. Quantitation of DNA and RNA with absorption and Fluorescensce Spectroskopy. Curr protoc Hum Appendix 3D 53 A3D1-A3D2. 2007.
55. Nucleics. Exonuclease I-Shrimp Alkaline Phosphatase clean up of PCR products. 2011 [25.11.11]; Available from: http://www.nucleics.com/DNA_sequencing_support/exonucleaseI-SAP-PCR-protocol.html.
56. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463-7.

57. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008 Nov;92(5):255-64.
58. NimbleGen. NimbleGen Arrays User's Guide CGH and CNV Arrays Version 8.1. 2011 [24.11.11]; Available from: http://www.nimblegen.com/products/lit/NG_CGHCNV_Guide_v8p1.pdf.
59. Redon R, Fitzgerald T, Carter NP. Comparative genomic hybridization: DNA labeling, hybridization and detection. Methods Mol Biol. 2009;529:267-78.
60. den Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat. 2000;15(1):7-12.
61. Mohler PJ. Ankyrins and human disease: what the electrophysiologist should know. J Cardiovasc Electrophysiol. 2006 Oct;17(10):1153-9.
62. Maruyama M, Arisaka N, Goto Y, Ohsawa Y, Inoue H, Fujiwara H, et al. Histidine 379 of human laeverin/aminopeptidase Q, a nonconserved residue within the exopeptidase motif, defines its distinctive enzymatic properties. J Biol Chem. 2009 Dec 11;284(50):34692-702.
63. Patel J, Pathak RR, Mujtaba S. The biology of lysine acetylation integrates transcriptional programming and metabolism. Nutr Metab (Lond). 2011;8:12.

Appendix A

In this appendix all detected CNVs by aCGH are presented with their chromosome position, start and end positions for the probes, number of probes involved in the variant, mean values, genes and if they are documented in DGV.

Array number 461678

Subarray A01

Position	Start position	End position	Probes	*Genes*	Mean	DGV
4q22.3	98549999	98762499	17	*C4orf37*	0,26	No
7q11.23	72087499	72349999	20	*GTF2IRD2P, FKBP36 STAG3L2, GTF2IPI, GTF2IRD2B, NCF1C, PMS2L14, SPDYE8P*	-0.28	No
7q11.23	74812499	74937499	11	*PMS2L14, LOC541473, NSUN5B, POM121-2,POM121-C, STAG3L1, TRIM73*	-0,28	No
7q22.1	101849999	102062499	15	*PRKRIP1, ALKBH4, RASA4, POLR2J3, UPLP, AK311374, BC085014,ORA12, RASA4, KIAA0538, SPDYE2, AK097289*	-0,55	No
8p23.3	0	74999	7	*BC071667*	-0,57	No
15q13.1	26112499	26649999	18	*HERC2, GOLGA8G, FLJ32679, LOC283767, DQ579249*	-0.34	No
16p11.2	31812499	33712499	78	*TP53TG3, ZNF267, IGHV, IG VH, IGHV3-30, HERC2P4, X69637, DQ571479, TP53TG3, TP53TG3b, LOC440366, DQ574674, BC038215, SLC6A10P, BC041879, SLC6A8, Z14218*	-0.22	No
19q13.31	48162499	48299999	5	*PSG2,PSG11,PSG6*	-0.24	No
22q11.21	17012499	17224999	10	*AK129567,GGT2, POM121-like 1, DKFZp434K191, USP18, AK748067,*	-0.39	No
22q11.21	18787499	19024999	9	*LOC375133, AK129567, RIMBP3,*	-0.38	No
22q11.21	19824999	20024999	11	*AK129567, BC039313, AK128837, GGT2,BC053652, BC112340, DQ571494,DKFZp434K191,BCR*	-0.35	No

Subarray A02

Position	Start position	End position	Probes	*Genes*	Mean	DGV
1p36.33	24999	787499	24	*AK311604, DQ580039, AK026901, X64709, LOC643837, OR4F16*	-0.28	No
1p36.13	16749999	16962499	7	*NBPF10, NBPF1, FLJ00313, MST1, AL035288, KIAA0445, MSTP9,*	-0. 23	No
4q28.3	131337499	131449999	9		0.2	No
7q11.23	72062499	72337499	22	*FKBP36, STAG3L3, PMSL214, LOC441259, GTF2IPI, NCF1C, GTF2IRD2B, NSUN5C, TRIM74, LOC541473, SPDYE8P,*	-0.25	No
7q11.23	73774999	74949999	71	*GTF2I,GTF2IRD2, STAG3L2, PMSL25, AK095583, WBSCR16, TRIM73, NSUN5B, BC070376, TFII-I,NCF1, BC047594,BC110795, hPMS7, BC068610, GTF2IRD2b,NCF1C, GTF2IP1, hPMS7, BC110795, PMS2L14, DQ579684, AK095583, AL832384, WBSCR16, AL831977, DQ571357, DQ579684, LOC441259, SPDYE8P, STAG3L1, LOC541473, TRIM73, NSUN5B, POM121C, POM121-2,*	-0.24	No
7q22.1	99662499	99762499	10	*STAG30S, SPDYE3, PMS2L13*	-0.25	No
7q22.1	101849999	102099999	17	*PRKRIP1, ORA12, LRWD1, POLR2J, KIAA0538, UPLP, RASA4, AK31174, BC085014, ALKBH4, DQ601342, SPDYE2,*	-0.37	No
8p23.3	0	7499	7		-0.54	No
9p21.1	31274999	31349999	6		0.2	No
11p15.5	62499	224999	7	*IFITM5, DQ584937, AK310751, BX647595, DQ577244, SCGB1C1, ODF3, BET1L, RIC8A, SIRT3,*	-0.3	No
15q13.1	26049999	26812499	29	*HERC2, FLJ36131, LOC28376, FLJ32679, GOLGA8G, DQ592322, LOC283767, AK311660, BC150524*	-0.24	No
16p11.2	31862499	33712499	75	*TP53TG3b, HERC2P4, IGHV, IGH, I GHV3-30, X69637, BC041879, DQ571479, TP53TG3, LOC440366, DQ574674, BC038215, SLC6A10P, BC038215, SLC6A8, Z14218*	-0.22	No

Position	Start position	End position	Probes	*Genes*	Mean	DGV
21p11.2-q11.2 Centromer	9887499	13512499	13		-0.26	No
22q11.21	17074999	17224999	7	*GGT2, POM121-1 like, AK129567, DKFZp434K191, AX748067,*	-0.21	No
22q11.21	18762499	19024999	11	*LOC375133, AK129567, RIMBP3*	-0.28	No
22q11.21	19824999	20237499	23	*AK129567, GGT2, BC039313, DKFZp434K191, PI4KAP2, AK094490, AK128832, BC053652, BC112340, DQ571494, BCR, RIMBP3B, HIC2, AK094490, AK308704, AK094309, PI4KAP2, LOC375133, PI4KAP2, RIMBP3C*	-0.26	NO

Subarray A03

Position	Start position	End position	Probes	*Genes*	Mean	DGV
7q22.1	101849999	102062499	15	*PRKRIPI, ALKBH4, LRWD1, RASA4, POLR2J3, POLR2J, UPLP, KIAA0538, AK311374, HSPC047, SPDYE2, BC085014*	-0.32	NO
8p23.3	0	49999	6	*BC071667*	-0.59	No
15q11.2	18412499	20274999	82	*HERC2P3, POTEB, GOLA8E, OR4N4, OR4M2, DQ576041, LOC642311, DQ592463, AX748135, DQ582073, DQ595648, DQ592322, AL832227, DQ582260, POTEB, A26B1, M84131, BC047459, DQ786202, AK058056, BC062994, OR4N4, OR4M2, LOC650137, CR619482, OR4N4, DQ582025, DQ582940, DQ599733*	-0.22	NO

Position	Start position	End position	Probes	*Genes*	Mean	DGV
15q13.1	26112499	26824999	25	*HERC2, FLJ36131, GOLGA8G, DQ579249, AK095014, DQ592322, DQ5788838, FLJ32679, DQ59973, LOC283767, DQ594306, DQ582071, AK309255, AK307870, BC047911, AK095014, DQ582025, DQ582940, DQ597560, BC150525, BC150524, BC066982, AK091254, BC035099, AK311660*	-0.24	NO
22q11.21	17074999	17224999	7	*AK129567, GGT2, POM121-like 1, DKFZp434K191, AX748067*	-0.44	No
22q11.21	18774999	19024999	10	*LOC375133, RIMBP3, AK129567*	-0.36	No
22q11.21	19924999	19987499	5	*BC112340, DQ571494, BCR, DKFZp434K191, AK129567*	-0.54	No

Array number 509388

Subarray A01 Sample number 2P

Position	Start position	End position	probes	*Genes*	Mean	DGV
8p23.1	6824999	7799999	9	*AK0307331, SPAG11B, DEFB10, HE2, LOC728358, AF355799, DEFA3, DEFA5, BC030294, AK307331, DEFB109, DEFB4, HE2, DEFB103B, DEFB104A, DEFB106A, DEFB105A, DEFB107B, SPAG11A, DEFB103A*	-0.42	No
15 q11.1-11.2	18394999	20214999	16	*HERC2P3, AX748135, DQ58226, DQ582073, GOLGA8E, DQ592322, DQ578838, POTEB, M84131, BC047459, OR4N4, IGH, OR4M2, CR619482, DQ786202*	-0.31	No
16p11.2	28274999	28599999	6	*EIF3C, CLN3, APOB48R, CCDC101, SULT1A1, SULT1A2, NPIPL1, NUPR1,*	-0.42	No
16p11.2	32564999	33474999	8	*TP53TG3, TP53TG3b, DQ574674, BC038215, LOC440366, SLC6A10P, IGHV, IGH, BC041879, DQ571479*	-0.28	No
17q21.31	41599999	42119999	10	*KIAA1267, ARL17, LOC51326, LRRC37A, L RRC37A2, NSF*	0.38	No

Subarray A02 Sample number 4P

Position	Start position	End position	Probes	*Genes*	Mean	DGV
4q12-31.1	58304999	64024999	76	*LEC3, LPHN3, BC039452,*	0.21	No
7q11.22-q11.23	71629999	76439999	73	*SBDSP, NSUN5, POM121, STAG3L3, TRIM74, PMS2L14, SPDYE8P, STX1A,MLXIPL, WBSCR22-27-28, RFC2, PMS2L5, STAG3L2, GTF2IPI,DTX2, STYXL1, around 100 genes in total.*	-0.25	No
7q.22.1	99969999	102244999	33	*AGFG2, PP1483, FLJ00248, LRCH4, MOSPD3, TFR2, GNB2,EMID2, ZAN, PLOD3, CUX1,PRKRIP1, KIAA, +++++ around 80 genes*	-0.29	No
9p13.1-q13	39129999	70134999	125	*CTNAP3, KIAA1714, FAM75A2, ZNF658B, KGFLP, ANKRD20A3, AK126080, CR605783, FAM75A7, LOC554249, FOXD4L5+++++ around 130 genes*	-0.25	No
19q13.41-q13.42	58174999	60124999	26	*ZNF,VNIR2-4, KIAA, NALP12, PRKCG, MYADM, CACNG7, PRPF31, FLJ,LAIR1, KIR2DL3, LILRA4, +++ around 80 genes*	-0.23	No

Subarray A03 Sample number 7P

Position	Start position	End position	Probes	*Genes*	Mean	DGV
8p23.1	6824999	7864999	9	*DEFB10, SPAG11B, SPAG11A, ZNF705D, HE2, AK307331, LOC728358, AF35579, DEFA3, BC030294, AK307331, DEFB109, DEFB4, DEFB103B, SPAG11B, DEFB104A, DEFB105A, DEFB106A, DEFB107B, SPAG11A, HE2,ZNF705D*		No
9p13.1-p12	38999999	41859999	41	*CNTNAP3, KIAA1714, DKFZ572C163, ZNF658B, FAM75A2, LOC401507, DQ586551, DKFZp686I15204, CTNAP3B, FAM75A7,*	-0.3	No
9p11.2-q12	43224999	67339999	62	*AK054645, DQ586551, DQFZp686115204, DQ579969, KGFLP1, BC134347, FAM27E3, CR615666, LOC554249, AK126080, LOC401507, FAM75A7,DQ570938, BC011779, BC068605, AQP7P1, MGC21881, DQ597175, AK131029, DQ594428, AL953854.2-002, LOC643648, AK308561 ,BC01179, CR627148, BC068605, LOC442421, AK309896, AK000451, MGC2881, BC019880, BC065763, AK310876, AK074198, BC070322, BC071803, AQP71, AK074198, BC064331*	-0.21	No

Position	Start position	End position	Probes	*Genes*	Mean	DGV
14q11.1-11.2 Centromer	18394999	19109999	9		-0.27	No
15q11.1-11.2	18394999	19954999	15	*DQ576041, CR622584, DQ592463, HERC2P3, AX748135, DQ582073, GOLGA8E,D Q578838, DQ582025, A26B1,M84131, BC047459, POTEB, ORAM2, AK058056*	0.26	No
15q14	32499999	32694999	5	*DQ570464, DQ585295, GOLGA8B, GOLGA8A, DQ582688,*	-0.42	No
17q21.31-21.32	41729999	42119999	8	*LOC51326, LRRC37A-A2, ARL17 ,NSF*	0.54	No

Subarray A04 Sample number 8P

Position	Start position	End position	Probes	*Genes*	Mean	DGV
3p26.2	4159999	4354999	5	*SUMF1, SETMAR*	-0.65	No
6q27	168089999	168284999	5	*MLLT4, HGC6.1.1, KIF25, FRMD1, AK289488, FLJ00181*	0.37	No
9p21.2	25609999	25869999	5	*TUSC1, BC043546,*	-0.27	No
17q21.31-21.32	41729999	42119999	8	*ARL17P1, LOC51326, LRRC37A, LRRC37A2, NSF, ARL17*	-0.51	No

Subarray A05 Sample number 9P

Position	Start position	End position	probes	Genes	Mean	DGV
10q26.3	132014999	132404999	8		0.5	No
14q11.1-q11.2 Centromer	183949999	19304999	10		0.34	No
21q11.2-22.3	1345499	46853544	465	The whole q arm is duplicated.	0.44	No
17q21.31-21.32	41794999	42119999	7	L*RRC37A2, ARL17P1, ARL17 ,NSF*	-0.42	No

Subarray A06 Sample number 10P

Position	Start position	End position	Probes	*Genes*	Mean	DGV
8p23.1	7214999	7864999	8	*DEFB4, SPAG11B,SPAG11A, DEFB10103-104-105-106A, DEFB107B ,ZNF705D, HE2,*	0.38	No
10q11.23	50894999	51284999	6	*PARG, AGAP8, CTGLF1, DQ588223, TIMM23, AK098044, AGAP7, DQ588224, MSMB, NCOA4,*	0.24	No
13q21.31	62009999	62594999	8		0.25	No

15q11.1-q11.2	18394999	19954999	15	*DQ576041, DQ592463, HERC2P3, AX748135, DQ582073, DQ595648, GOLGA8E, DQ592322, POTEB, DQ582073, AL832227, A26B1, M84131, BC047459, DQ786202, OR4N4, OR4M2, AJ004954, LOC650137, LOC642311, DQ576041, DQ592463, DQ578838, DQ582025,D Q582260, CR619482, AK058056*	-0.42	No

Subarray A07 Sample number 12P

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
1p36.21	12869999	13324999	5	*PRAMEF9, PRAMEF10, PRAMEF8, PRAMEF5, PRAMEF22, PRAMEF6, PRAMEF13, PRAMEF14*	-0.26	No
1p36.13	16769999	17159999	8	*NBPF10, FLJ00313, MST1, AL035288, MSTP9, KIAA0445, CROCC*	-0.24	No
9q12	65454999	65779999	6		-0.23	No
10q11.22	46149999	46799999	12	*BMS1P5, CR604707, SYT15, GPRIN2PPYR1, ANXA8L1, LOC643650, BC075841*	0.33	No

Subarray A08 Sample number 13P

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
16p11.2	31914999	33474999	14	*TP53TG3B, HERC2P4, IG VH, IGHV3-30, IGH, X69637, BC042588, BC041879, DQ571479, TP53TG3, LOC440366, DQ574674, BC038215, SLC6A10P, BC038215*	-0.27	No
22q11.21	18654999	19077499	9	*DGCR6L, AK129567, GGT2, AX747971 ,AK094309, LOC375133, USP41, KIAA1653, RIMBP3*	0.21	No

Subarray A09 Sample number 15P

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
14q11.1-11.2 centromer	18394999	19304999	10	*POTEG, A26C2, AY458019, AK056135, DQ595091, DQ591735 ,DQ595091, DQ591735, AY338954, BX248778, AK022914,Or4Q3*	0.21	No

15q11.1-q11.2	18394999	21059999	25	*HERC2P3, DQ592463, POTEB, A26B1, M84131, OR4N4, GOLGA8E, TUBGCP5, AJ004954, NIPA2, WHAMML1, KIAA0393, DQ576041, CR622584,BC012949, AX748135, DQ582073, DQ595648, DQ592322, DQ578838, DQ582025 ,AL832227, DQ582260, BC047459, DQ786202, AK058056, BC06,OR4N4, ORAM2, LOC650137, CR619482, AJ004954, DQ582025, DQ582940, DQ599733, LOC283767, DQ582071, TUBGCP5, KIAA1899, CYFIP1, NIPA2, BC044583, WHAMML1, DQ583692, DQ572986, AX747926, DQ600343, KIAA0393, BC012949, FLJ36131, DQ582023, DQ588972,*	-0.24	No
17q21.31-21.32	41729999	42119999	8	*LOC51326 ,LRRC37A, LRRC37A2, ARL17P1, ARL17, NSF,*	-0.61	No
22.q11.23 3	23919999	24179999	5	*KIAA1671, CRYBB3,CRYBB2 BC040576, LRP55L , BC047380, BC037884, IGLL3*	0.27	No

Subarray A10 Sample number 16P

No deletions or duplications were detected for this subarray.

Position	**Start position**	**End position**	**Number of probes**	**Genes**	**Mean**

Subarray A11 Sample number 17P

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
1q31.3 middle	19486999	195129999	6	*CFH, CFHR3, CFHR1, CFHR3,CFHR4*	-0.41	No
5q13.2	68834999	70654999	22	*OCLN, LOC730394, GTF2H2D, AX748379, DQ570149, SMA4-5,SMN2, NAIP, AK311046, S74729, BC045789, SERF1B, DQ570081, AK311627, DQ571461, DQ575504, DQ571494, AK124130, DQ587764, SERF1, DQ575504, DQ570150, DQ596233, DQ591061, DQ786306, AF073520, BC045789, DQ786263, DQ587763, GTF2H2B, GUSBP1, LOC730394, AX748379, SMA3, DQ596042, DQ573166, DQ587763, DQ570081, DQ599922, DQ577094, DQ570150, AF073520, BC045789, OCLN*	-0.3	No

7q11.23	72019999	72409999	8	*POM121, NSUN5C ,TRIM74, BC073780, LOC541473,STAG3L2, PMS2L14, SPDYE8P, FKBP36, GTF2IP1, hPMS7, WBSCR20C BC073780, BC022013, LOC541473, STAG3L2, STAG3L3, PMS2L14, hPMS7, BC110795, BC068610, NCF1C, GTF2IRD2B, BC018166, NSUN5, FKBP6, TRIM50*	-0.45	No
7q11.23	73774999	74879999	16	*GTF2I GTF2IRD2, PMS2L5, STAG3L2, WBSCR16, GTF2IP1,PMS2L14, SPDYE8P ,NSUN5, TRIM73, BC068610, BC070376,TFII-I,NCF1, BC047594, BC110795, hPMS7, DQ579684, AK095583, AL832384, WBSCR16, NCF1C, BC068610, BC047594, BC110795, AL831977, DQ571357, DQ579684,BC068610, BC047594, BC110795, LOC441259, STAG3L1, LOC541473, BC073780*	-0.36	No
14q11.1-q11.2 centromer	18394999	19304999	10		0.21	No
15q13.1 2	26194999	26714999	5	*HERC2, FLJ36131, GOLGA8G, FLJ32679, LOC283767, DQ592322, DQ57883B, DQ583954, FLJ32679, DQ579249, DQ599733, FLJ36131, DQ582071, AK309255, AK307870, AK095014, BC047911*	-0.31	No
16p12.1	22359999	22554999	5	*DQFZp547E087, LOC641298, LOC641298, LOC23117, AK054709, LOC23117, DQ595631, DQ576951, AK57335*	-0.5	No
17q21.31-21.32	41599999	42119999	9	*KIAA1267, LOC644246, LOC51326, LRRC37A, LRRC37A2, ARL17P1, ARL17, NSF*	0.33	No

Subarray A12 Sample number 20P

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
5q35.3	178619999	178814999	5	*ADAMTS2*	0.44	No
14q11.2	21644999	21969999	7	*TCRA, TCRAV8.1a, AV2S1A1, TRD, AV30S1, TCRAVN1, TCR-alpha, hADV29S1, hDV102S1, av27SI,AV4S1*	-0.26	No

Position	Start position	End position	Probes	*Genes*	Mean	DGV
15q11.1-q11.2	18394999	19954999	15	*HERC2P3,P OTEB ,OR4N4, DQ576041, LOC642311, DQ592463, CR622584, BC012949, AX748135, DQ582073, DQ595648, GOLGA8E, DQ592322, DQ578838, DQ582025, DQ582260, A26B1, M84131, BC047459, DQ786202, OR4N4,AK058056, BC062994, OR4M2,LOC650137, CR619482, AJ004954*	-0.48	No

Array number 514596

Subarray A01 Sample number 4P

Position	Start position	End position	Probes	*Genes*	Mean	DGV
5q13.2	69322499	70622499	15	*SMA4, SERF1, SERF1B, SMN2, AF073520, BC045789, NAIP GTF2H2B, GTF2H2, OCLN, AK311046, GUSBP1, SMA5, DQ786263, LOC730394, S74729, DQ596042, DQ571461, DQ599922, AX748379*	0.247	No
8p23.1	7182499	7897499	8	*DEFB4, SPAG11B ,DEFB104A ,DEFB105A, SPAG11A, HE2, DEFB103A, DEFB109, AK307331, ZNF705D*	-0.295	No
17q21.31-21.32	41827499	42152499	6	*ARL17, LRRC37A2 ,ARL17P1, ARL17,NSF*	-0.469	No

Subarray A02 Sample number 9P

Position	Start position	End position	Probes	*Genes*	Mean	DGV
9p11.2	44167499	45207499	12	*KGFLP1 , AK126080, CR615666, DQ594366, BC134347, AK131029, FAM27E3*	0.21	No
9p11.2-11.1 centromer	46182499	46702499	7	*LOC554249, KGFLP1, AK126080, LOC643648*	0.21	No
10q26.3	132047499	132437499	7	*No genes found*	0.401	No
14q11.1-11.2 centromer	18752499	19142499	6	*BC016035, AK022914, BC041856, BX248778, BC017398, AY338954, DQ583164,DQ590589, DQ583610, A26C2,POTEG, AY338954*	0.516	No
17q21.31-q21.32	41827499	42152499	6	*ARL17, LRRC37A2, ARL17PI,NSF*	-0, 515	No

21q11.2-22.3	13487499	46886044	464	*The whole q arm is duplicated*	0.48	No

Subarray A03 **Sample number 12P**

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
7q11.23	74327499	74847499	6	*BC068610, BC047594, BC110795, PMS2L14, hPMS7, AL831977, DQ571357, BC068610, LOC441259, SPDYE8P, STAG3L1 WBSCR16,*	0.214	No
8q11.21	48717499	48977499	5	*KIAA0146, CEBPD, PRKDC, AK095778,*	0.284	No
9p13.1-q13 (ncluding several bands and the centromer	39292499	70167499	122	*FAM75A2, DKFZp572C163 ZNF658B, KGFLP2, FAM74A3, CTNAP3B, DKFZp686115204, LOC401507,CR933660,DQ588 135, DQ587539, CBWD5, CCDC29, AK054645, LOC554249, AK126080, CBWD3, FOXD4L4,++++++*	0.281	No
10q11.22	46117499	46832499	12	*BMS1P5, CR604707, SYT15 GPRIN2, ANXA8 PPYR1, ANXA8L1, LOC644054, AK309109, BC075841, AF130068, LOC728643, BC127745, AK3092109, AK309024*	0.518	No
15q11.2	18427499	20247499	15	*DQ576041, CR622584, HERC2P3,AX748135,GOLGA8 E), POTEB, DQ582073, DQ595648, DQ578838, DQ582025, A26B1, M84131, BC047459, DQ786202, IGH,OR4N4, DQ786202, AK058056, BC062994, ORAM2*	0.231	No
15q11.2	20637499	21092499	5	*NIPA1, BC044583, WHAMML1, AX747926, DQ600343, DQ582940, DQ578838, DQ588972, DQ583692, KIAA0393, BC012949, FLJ36131, DQ582073, GOLGA8E, DQ592322, DQ578838, DQ582025*	0.226	No

Subarray A04 **Sample number 12P**

Position	**Start position**	**End position**	**Probes**	**Genes**	**Mean**	**DGV**
10q11.22	46247499	47092499	12	SYT15, PPYR1 ,ANX8L1, ANXA8, LOC642826, LOC643650, GPRIN2, LOC728643, AK309109, AK309024, BC075841, AF130068, BC127745	0.369	No

15q11.2	18427499	19987499	14	LOC6446096, HERC2P3, GOLGA6L6,GOLGA8E,NBEAP1, POTEB, LOC646214, CXADRP2, LOC348120, NFIP2, LOC727924, ORAM2, OR4N4, OR4N3P, AX748135, A26B1, DQ595648, M84131, LOC650137,AJ004954 AL832227	0.277	No

Subarray A05 **Sample number 21P**

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
9p13.1-p12	39877499	40332499	7	*FAM75A2, DQ586551, DKFZp686115204*	0.226	No
9p12-11.2	41697499	42997499	13	*MGC221881, KGFLP2, ANKRD20A2, ANKRD20A3, FOXD4L2, FOXD4L4, AQP7P3, KGFLP2, DQ588135, DQ587539, AK125850, BC064148, CBWD5, AK128231, AK310876, BC070322, BC052332, LOC554249*	0.234	No
9p11.2	43192499	43582499	7		0.207	No
9p11.2-q12 (including several bands and the centromer)	44102499	65812499	39	*KGFLP1, AK126080, DQ594366, BC134347, AK131029, FAM27E3, CR615666, DQ594428, LOC100132167, BC134347, AK131029, LOC554249, KGFLP1, AK126080, LOC401507*	0.24	No
9q12	66787499	68412499	12	*BC070322, BC071803, AQP7P1, BC064331, FAM27E3, DQ786189, BC110369, AK308561, ANKRD20A3, ANKRD20A1 BC110369, CR626459, AK311167, PGM5P2*	0.231	No
14q11.1 centromer	18427499	19077499	7		0.331	No
17q21.31-q21.32	41762499	42152499	7	*LOC51326, LRRC37A, ARL17P1, ARL17, LRRC37A2, NSF, LOC51326*	-0.444	No

Subarray A06 **Sample number 21P**

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
9p11	44297499	46182499	17	*2KGFLP1, AK126080, DQ594366, BC134347, AK131029, FAM27E3, DQ594426, CR615666, LOC100132167,*	0.266	No
17q21.31-21.32	41827499	42152499	6	*ARL17, LRRC37A2, ARL17P1, NSF, LRRC37A,*	-0.323	No

Subarray A07 **Sample number 3P**

Position	Start position	End position	Probes	*Genes*	Mean	DGV
1q21.1	143292499	143812499	6	*NBPF1, NBPF10(NBPF family), NBPF20, NBPF14, CR603237, KIAA1245, NBPF8, COAS3, PDE4DIP, SEC22B, BX647792, AX747132*	0. 211	No
1q21.1	147517499	147777499	5	*FCGR1C, BC023516,*	0. 232	No
1q21.1	146542499	147192499	7	*KIAA, NBPF1, NBPF12, KIAA1693, BC110832, NBPF14, NBPF10, NBPF20, NBPF16 NBPF15, CR617196, AK310441*	0.226	No
5q13.2	69062499	70687499	19	*SMA4, SERF1, DQ575504, DQ570150, SERF1B, SMN2, GUSBP1, GTF2H2D, GTF2H2B, SMA5, DQ596233, DQ591061, AF073520, BC045789, DQ786306,S74729, NAIP, DQ786306, AX748379, OCLN (LOC730394, DQ599922, AX748379,GTF2H2, AK311627,*	0.342	No
6p11.2-p11.1 centromer	57557499	58402499	13		0.266	No
7q11.21	62302499	62757499	6	*LOC643955, BC040831, BC089392,*	0.209	No
7q11.22-q11.23	71662499	72442499	13	*TYW1B,SBDSP, POM121, NSUN5C,TRIM74, STAG3L2, LOC441259, PMS2L14, NCF1C GTF2IRD2B, TRIM50, BC047594, SPDYE8P, WBSCR20C, GTF2IP1 hPMS7 ,SPDYE7P, BC047594, BC068610, STAG3L3, LOC441259,FKBP6*	0.279	No
7q11.23	73807499	74977499	16	*GTF2I, STAG3L2, TFII-1, NCF,GTF2IRD2, PMS2L5, BC047594, hPMS7, BC110795, GTF2IRD2B, BC068610, PMS2L14, WBSCR16, AL831977, DQ571357, DQ579684, BC110795, TRIM73, NSUN5B, POM121C, SPDYE5, PMS2L,POM121-2, DQ601342, LOC441259, SPDYE8P, BC073780*	0.334	No
9p13.1	39227499	39617499	7	*CTNAP3,KIAA1714, FAM75A2, ZNF658B, DKFZp572C163*	0.217	No
9p12-q13 (including several bands and the centromer)	40852499	70167499	100		0.307	No
16p13.11 middle	16152499	16672499	8	*MRP6, ABCC6 , NOMO3, FLJ00322, Nb1a00537, LOC339047, AK310228*	0.211	No

Position	Start position	End position	Probes	*Genes*	Mean	DGV
16p11.2	28827499	29477499	9	*RABEP2, CD19, NFATC2IP , SPNS1, LAT,BX647358, LOC653390, AK075019, RUNDC2B, DQ576952, LOC606724, AK096982, MLAS BOLA2B ,BOLA2, SULT1A3, GIYD, CR627362,IMAA, DQFZp547E087, LOC23117, DQ576952,*	0.237	No
21q11.2	13487499	13812499	6		0.29	No

Subarray A08 **Sample number 24P**

Position	Start position	End position	Probes	*Genes*	Mean	DGV
8q24.3	143552499	144137499	7	*BAI1, ARC, JRK, C8orf55, LY6D LY6K, CR590598, PSCA, ,SLURP1, LY6uPAR GML ,LYNX1, CYP11B1, CYP11B2, H8*	0. 205	No
9p13.1	39357499	40072499	12	*ZNF658B , FAM75A2,*	0.225	No
9p12	40917499	41242499	5	*No genes found*	0.223	No
9p11.2	42867499	43712499	13	*BC052332,BC071803, ,DQ587119, BC064148, AK125850, ANKRD20A2, ANKRD20A3 DQ786189, AK054645,CCDC29,*	0. 229	No
9p11.2-p11.1 centromer	46182499	46702499	7		0.213	No
10q11.2	46247499	47027499	11	*SYT15 GPRIN2, ANXA8, PPYR, ANXA8L1,LOC644054, AK309109, LOC643650,AK309024, BC075841, AF130068, BC127745*	0.341	No
16p11.2	31687499	33377499	15	*ZNF720, FKSG41, ZNF267, IGHV, IGHV3-30, HERC2P4, TP53TG3b, BC041879, LOC440366, DQ574674, BC038215, SLC6A10P, IGHV, IGH, BC038215, LOC440366, BC041879, DQ571479*	0.212	No

Subarray A09 **Sample number 25P**

Position	Start position	End position	Probes	*Genes*	Mean	DGV
2p25.3	747499	1072499	6	*BC043553, BC128532, CR592733, LOC391343,SNTG2*	0.518	No
5q13.2	69062499	70817499	21	*SMA4, SERF1, DQ575504, DQ570150, DQ596233, DQ591061, ERF1B,SMN2, GUSBP1, GTF2H2D SMA5 ,BDP1, PMCHL2, OCLN,NAIP, LOC730394, S74729, AF073520, DQ571494, DQ571461, DQ596042, LOC730394, DQ599922,*	0.386	No

				DQ786306, AK311627		
9p13.1-p12	38772499	40397499	25	*CNTNAP3, KIAA1714, ,ZNF658B FAM75A2,*	0.262	No
9p12-q13 (including several bands and the centromere)	40722499	70167499	102	*Many genes*	0.326	No

Subarray A10 **Sample number 26P**

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
1q21.1	146867499	147647499	8	*KIAA1693, NBPF20,NBPF16, CR17196, AK310441, AK311729, BC023516, FCGR1C, FCR1B receptor*	0.228	No
5p14.2	23692499	24342499	7	*AX747383*	-0.21	No
7q11.23	73937499	74912499	13	*STAG3L2, PMS2L5, BC047594, BC110795, BC068610, hPMS7, GTF2IRD2B NCF1C, GTF2IP1, PMS2L14, WBSCR16, AL831977, DQ571357, DQ579684, LOC541473, TRIM73, NSUN5B,POM-121-2 AL831977 POM121-C, SPDYE8P, LOC441259, STAG3L1, AL831977*	0.399	No
9p12-p11.2	41892499	43192499	12	*BC019880, MGC21881, KGFLP2, AK126080, LOC554249, ANKRD20A2, DQ588135, AK125850, CR605783, AK094644, CBWD5, FOXD4L4, BC070322, AK128231, BC052332.BC071803, CCDC29, DQ587119, DQ786189, ANKRD20A3*	0.275	No
14q11-q11.2 Centromere	18427499	19337499	9	*POTEG, A26C2, AY458019, AK056135 BX248778, AK022914,DQ786293, AY338954, OR4q3,++++*	0.636	No
15q11.2	18427499	20377499	17	*HERC2P3, DQ582260,POTEB, A26B1, M84131,GOLGA8E ,DQ582025, DQ582025, DQ592322, OR4N4, DQ578838, M84131, AK058056, BC062994,OR4M2, AJ004954 , CR622584, AX748135, DQ786202, DQ599733*	0.391	No
16p11.2	32207499	32792499	5	*BC041879, DQ571479, TP53TG3, TP53TG3b, LOC440366, DQ574674, BC038215*	-0.302	No
17p11.2	18297499	18622499	5	*LGALS9C, tI132 CCDC144B, CCDC144B, TBC1D28, TRIMI6L, FBXW10, DKFZp43401826*	0.242	No

Position	Start position	End position	Probes	*Genes*	Mean	DGV
17p11.2	18752499	19077499	5	*PRPSAP2, SLC5A10, FAM83G, BC112347, GRAP, BC058012, LOC400581*	0.476	No
17q21.31-q21.32	41827499	42152499	6	*ARL17,LRRC37A2 , ARL17P1,*	-0.36	No
21q11.2	13487499	13812499	6	*No genes found*	0.218	No

Subarray A11 **Sample number 30P**

Position	Start position	End position	Probes	*Genes*	Mean	DGV
1p36.13	16802499	17127499	6	*NBPF1 ,FLJ00313, MST1, AL035288, KIAA0445, MSTP9, BC070363, CROCC*	0.266	No
5q13.2	69127499	70687499	18	*SMA4, SERF1, SMA3, GTF2H2D, GTF2H2, NAIP, SERF1B, GUSBP1, SMN2, DQ570150, BC045789, LOC730394, AX748379, DQ571494, AK12386, AK31001, LOC730394, DQ591060, DQ786306, DQ575504, AK12386, S74729, DQ596233, DQ786306, AK12386, DQ599922*	0.388	No
7p22.3	487499	877499	7	*PDGFA, PRKAR1, BDKFZp762F1415,HEATR2,UNC84A,KIAA0810*	0.22	No
7q11.23	71857499	75107499	48	*TYWIB, SBDSP, SPDYE7P, POM121, WBSCR20C, NSUN5C, TRIM74, TRIM50, FKBP6, BAZ1B, BCL7B, TBL2, WBSCR14 , MLXIPL, DNAJC30, WBSCR22, STX1A, ABHD11,ELN, LIMK1, EIF4H WBSCR5, LAT2, RFC2, CLIP2 ,GTF2IRD1, GTF2I, BC047594, BC110795, PMS2L14, SPDYE8P, BC68610, hPMS7, STAG3L1, STAG3L2, TRIM73, NSUN5B, BAZ1B, FZD9, RBAP2*	0.276	No
7q11.23	75692499	75952499	5	*FLJ37078, CCDS43604, HSPB1, HSP27, AX747594,YWHAG, SRCRB4D, ZP3 ,DTX2, FDPSL2A*	0.224	No
7q22.1	99482499	100652499	18	*DKfZp43L134, ZSCAN21, ZNF3, PP838, COP56, MCM7, STAG30S, AP4MI, HSMU4, TAF6, PVRIG, PILRB, FDFACT2, hPMS8, CNPY4, C7orf59,PILRB, TSC22D4, FDFACT2, ZCWPW1,MOSPD3,PCOLCE,FBX24, TFR2, ACTL6B, BAF53b, AK055267, GIGYF1, GNB2, ZAN,F LJ0009, FLJ00010, SLC12A9, EPHB4, TRIP6, GNB2, ACHE, MUC12, ARS2, MUC17,*	0.243	No

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
				MUC3, TRIM56, AP1S1,VGF, MOGAT3, PLOD3, ZNHIT, UFSPI ,ZAN		
7q22.1	101822499	102082499	5	*PMS2L3, PRKRIP1, ORA12, ALKBH4, LRWD1, KIAA0538, RASA4, POLR2J, BC085014, UPLP,POLR2J3, SPDYE2, RASA4, HSPC047, AK311374, BC085014, BC041025*	0.302	No
9p13.1	39292499	40137499	14	*FAM75A2, ZNF658B, DKfZp572C163, ZNF658B, FAM75A2*	0.257	NO
9p12-p11.2	40787499	43582499	32	*ZNF658, FAM75A7, ZNF658B, FAM75A5, LOC401507, DKFzp572C163, BC019880, MGC21881, KGFLP2, AK126080, LOC554249, ANKRD20A3 ANKRD20A2 ,DQ558135, DQ587539, AK125850, BC061448, CR605783, AK094644, CBWD5, FOXD4L4, AK310876, AK074198, BC052332, BC071803, DQ587119,CCDC29*	0.258	No
9p11.2-q12 (including several bands and centromere)	44167499	65552499	34	*KGFLP1, AK126080, DQ594366, BC134347, AK131029,F AM27E3,CR615666, DQ594428, LOC100132167, BC134347, AK131029, LOC554249, KGFLP1, AK126080, LOC643648, LOC401507, FAM75A7, AL953854.2-002*	0.265	No
9q12-q13	65747499	70167499	28	*DQ570938, AK308561, BC011779, CR627148, BC068605, LOC442421, AK309896, AK000451, MGC21881, BC019880, BC065763, BC070322, BC070322, BC071803, AQPZPI, AK074198, FAM27E3, DQ594366, BC002886, DQ786189, ANKRD20A3, BC110369, AK308029, CR626459,AK31167,FOXD4L6, CCDC29, FOXD4L5, CBWD3,CBWD5, AY343894, LOC440896,+++++++++*	0.254	No
16p11.2	31492499	32532499	11	*BC021555, BC000552, AK054698, ZNF720, FKSG41, ZNF267, IGHV, HERC2P4, X69637, TP53TG3b ,BC042588, BC041879, DQ571479, IGHV3-30*	0.282	No
16p11.2	32922499	33507499	5	*BC038215, LOC440366, DQ574674, TP53TG3, TP53TG3b, DQ571479*	0.316	No

Position	Start position	End position	Probes	*Genes*	Mean	DGV
17q25.3	73027499	74002499	16	*FLJ45079,BC040189,AK130926, BC039479,TNRC6C,KIAA1582,A K123771,TMC6,TMC8,FLJ00400 ,AK125672,UNQ464PRO809,SY NGR2 (integral membrane protein,membrane trafficking?),AFMID (arylformamidase),SURVIVIN-3B (inhibiotor of apoptosis),TK1,BIRC5(apoptosis inhibitor),EPR-1,BC042066,BC036810,SOCS3(s uppressor of cytokine signaling),PGS1 (phosphatidylglycerophosphate synthase 1),DNEL2,DKFzp76M186*	0.228	No
17q25.3	75432499	75952499	9	*BC044939,AK097283,TBC1D16, CCDC40 (cilia formation),KIAA1640,GAA(alpha glucosidase),EIF4A3 (translation initiation factor),CARMA2 (scaffold protein, plasma membrane),AK002138,CARD14(s caffold protein, plasma membrane),SGSH,SLC26A11,KIA A1618,RNF213,pp7201*	0.213	No
17q25.3	76602499	77837499	14	*FLJ90757,BAIAP2,AATK,BAIAP 2,AZ11,AK098403,C17orf55,AL8 32593,BAHCC1,KIAA1118,SLC3 8A10,KIAA1499,TSPAN10,NPLO C4,PDE6G,ARL16,CCDC137,MR PL12,DYSFIP,P4HB,ARHGDIA, THOC4,DCXR,ASPCR1,FASN,C CDC57,CSNK1D,LRRC45,DYSFI P1,STRA13,ANAPC11,ARHGDIA ,CSNK1D,PYCR1++++++*	0.252	No
17q25.3	78292499	78611061	6	*FN3K (fructosamine 3 kinase)TBCD (beta tubulin cofaktor),ZNF750,KIAA0988,B3G NTL1 8(glycosyl transferase),*	0.231	No
20q13.33	62172499	62410787	7	*TCEA2,RGS19,OPRL1,C20orf20 1,NPBWR2,MYT1,KIAA0835,PC MTD2,DQ574507, C2orf69,*	0.242	No
21q11.2	13487499	14072499	7	*POTED (ankyrin domain, placenta),DQ590589,LOC441956 ,*	0.36	No

Subarray A12 **Sample number 33P**

Position	Start position	End position	Probes	*Genes*	Mean	DGV
1p36.33	32499	1722499	16	*OR4F5, AK125248, AK311604,*	0.39	No

Position	Start position	End position	Probes	*Genes*	Mean	DGV
				DQ580039, DQ600587, DQ599874, DQ599768, ORF416, LOC643837, SAMD11, HES4, ISG15, UNQ2998, AGRN, C1orf159, TNFRSF4, SDF4, SCNN1D, DVL1, PUSL1, TASIR3, UBE2J2, NOC2L, FAM132A, B3GALT6, KLHL17, PLEKHN1, CPSF3L, GLTPD1, ATAD3A-B and C, MIB2, NADK, GNB1,CDC2L2, MMP23B, MXRAS,+++++		
1p12-p11.2 centromer	120347499	120867499	9	*NOTCH2, N2 ,NOTCH2NL, FAM72B, FCGR1B, SRGAP2*	0.263	No
1q21.1	142642499	143942499	14	*NBF1, NBPF14, NBPF20, AF379632, AF161426, DQ786323, AL050141, AF379632, AF161426, COAS3, CR603237, PDE4DIP, SEC22B, NOTCH2NL, KIAA1245, AK747132, BX647792,NRG2*	0.265	No
1q21.1-q21.2	145892499	148037499	23	*KIAA1693, NBPF1, GPR89B, GPR89A, GPR89C, PDZK1P1, LOC200030, NBPF11, NBPF15, NBPF16, NBPF10, CR617196, FCGR1B, FCGR1A, AF161|426, AF380582, BC089397, BC110832, BC073170,AK309334, AX746564,+++++*	0.382	No
4p13-12	45142499	45792499	11	*GABRG1*	-0.226	No
4p15.1	29997499	30257499	5	*No genes found*	-0.234	No
4q13.1	63537499	63927499	5	*No genes found*	-0.214	No
7q11.22-q11.23	71597499	75042499	51	*TYW1B, SBDSP, SPDYE7P, POM121, WBSCR20C, NSUN5C, TRIM74, BC073780, STAG3L2, LOC441259, PMS2L14, SPDYE8P, hPMS7, BC047594, GTF2IP1, TRIM50, TBL2, FKBP36, MLXIPLVPS370, WBSCR14, WBSCR22, STX1A, ELN,LIMK1, LAT2, RFC2, CLIP2, GTF2IRD1,TFII-1,STAG3L1,+++*	0.244	No
7q11.23	76082499	76407499	6	*LOC554248, POMZP3, BC043544,*	0.234	No
9p23	12122499	12512499	6	*No genes found*	-0.245	No
9p13.1-q13 (including several bands and the centromer)	38967499	70167499	127	*Many genes*	0.371	No
14q11.1-q11.2 (including several bands and the centromer)	18427499	19337499	9		0.64	No
16p13.11	14982499	15372499	7	*PDXDC1, LOC728138, RRN3, NPIP, NTAN1, AK125313,*	0.215	No

Position	**Start position**	**End position**	**Probes**	***Genes***	**Mean**	**DGV**
				FLJ00285, DQ596229, AK125313, NBla00537		
16p11.2	28437499	30127499	20	*NUPR1, CCDC101, SULT1A2, SULT1A1, EIF3C, NPIPL1, AK125489, ATXN2L, SH2B1, RABEP2, ATP2A1, NFATC2IP, SPNS1, A2LG, BX647358, RUNDC2B, CD19, MLAS, GIYD2, BOLA2B, SULT1A3, KIF22,SPN, QPRT, TMEM219, PPX,TAOK2, TBX6, HIRIP3 ,MAZi, SEZgL2, MVP, KTCD13, FKSG86, YPEL3, MAPK3, ALDOA, CORO1A, GDPD3, IMAA, INO80E,++++++*	0.272	No
16p11.2	30647499	34157499	29	*SRCAP, PHKG2, LOC90835, RNF40, ZNF629, BCL7C, CTF1, MYST1,ORA13, STX4, FBXL19, ZNF668, ITGAM, VKORC1, BCKDK, FUS, PYCARD, TRIM72, ITGAM, ITGAX, ARMC5, ERAF,ZNF720, HERC2P4, TGFB1I1,SLC5A2, TP53TG3b, TP53TG3,X69637, COX6A2, IGHV, SLC6A10P, IGH,SLC6A8, Z14218,+++++++*	0,285	No
17p11.2	18752499	19077499	5	*PRPSAP2, FAM83G, BC112347, SLC5A10, GRAP, BC058012, LOC400581*	0.642	No
17p11.2	20182499	20767499	8	*CCDC144C, BC019672, AK310665, BC044655, LGALS9B, BC110641, AX748015, AK127974, DQ577610, DQ586256, DQ577296, DQ576593, DQ571776, DQ583069, DQ571391, DQ598579, DQ597254, DQ572047,DQ574249++++*	0.281	No
19p13.11	17062499	17452499	7	*MYO9B, USHBP1 ,AIEBP, ANKLE1, NR2F6, C19orf62, OCEL1, PCIA1, DDA1 ,ABHD8, MRPL34, DDAI, GTPBP3, ANO8, AK310794, PLVAP, BST2,AK311380, FAM125A,NXNL1,SLC27A1,++++*	0.203	No
19p13.11	17907499	18232499	6	*CCDC124, KCNN1, ARRDC2, IL12RB1, IL12RB1 MAST3,PIK3R2, IFI30,FKSG24,PDE4C, KIAA1683*	0.202	No
19p13.11	184227499	18687499	5	*ELL, FKBP8 ,C19orf50, UBA52,c19orf60,TEMEM59L, KLHL26, CRTC1,*	0.203	No
21q11.2	13487499	13812499	6	*No genes found*	0.313	No

Appendix B

In this appendix a summary table of all detected copy number variants on array numbers 461678, 509338 and 514596 are presented. The mean values are not documented. The green colour for the columns marked with mean represent duplications and the red colour deletions.

Summary table of all mutations found in array 461678, 509338 and 514596

Chromosome number	Start position	End position	Array 461678	Array 509338	Array 514596	Mean
1	2499	787499	A02	-	-	
1	32499	1722499	-	-	33P	
1	120347499	120867499	-	-	33P	
1	12869999	13324999	-	12P	-	
1	143292499 146542499 146867499 147517499 142642499 145892499	143812499 147192499 147647499 147777499 143942499 148037499	-		23P 23P 26P 23P 33P 33P	
1	16749999	16962499	A02	-	-	
1	16769999	17159999	-	12P	-	
1	16802499	17127499	-	-	30P	
1	19486999	195129999	-	17P	-	
2	747499	1072499	-	-	25P	
3	4159999	4354999	-	8P	-	
4	29997499	30257499	-	-	33P	
4	45142499	45792499	-	-	33P	
4	58304999	64024999	-	4P	-	
4	63537499	63927499	-	-	33P	
4	98549999	98762499	A01	-	-	
4	131337499	131449999	A02	-	-	
5	178619999	178814999	-	20P	-	
5	23692499	24342499	-	-	26P	
5	68834999	70654999	-	17P	-	
5	69322499 69062499 69062499 69127499	70622499 70687499 70817499 70687499	-	-	4P 23P 25P 30P	
6	168089999	168284999	-	8P	-	
6	57557499	58402499	-	-	23P	
7	487499	877499	-	-	30P	
7	62302499	62757499	-	-	23P	
7	71629999	76439999	-	4P	-	
7	71597499	75042499	-	-	33P	
7	72087499 72062499 72019999	72349999 72337499 72409999	A01 A02	 17P	-	
7	71662499	72442499	-	-	23P	
7	71857499	75107499	-	-	30P	
7	73807499 73937499	74977499 74912499	-	-	23P 26P	
7	73774999 73774999	74879999 74949999	 A02	17P	-	

Chromosome number	Start position	End position	Array 461678	Array 509338	Array 514596	Mean
7	74812499	74937499	A01	-	-	
7	74327499	74847499	-	-	12(1)	
7	75692499	75952499	-	-	30P	
7	76082499	76407499	-	-	33P	
7	99482499	100652499	-	-	30P	
7	99662499	99762499	A02	-		
7	99969999	102244999		4P		
7	101849999 101849999	102062499 102099999	A01,A03 A02	-		
7	101822499	102082499	-	-	30P	
8	0 0	7499 49999	A01,A02 A03	-		
8	48717499	48977499	-	-	12p (1)	
8	6824999	7799999	-	2P		
8	682499 7182499	7864999 7897499	-	7P	 4P	
8	7214999	7864999	-	10P		
8	143552499	144137499	-	-	24P	
9	12122499	12512499	-	-	33P	
9	25609999	25869999	-	8P	-	
9	31274999	31349999	A02	-	-	
9	38999999	41859999	-	7P	-	
9	38967499	70167499	-	-	33P	
9	39877499	40332499	-	-	21P(1)	
9	38772499 39292499	40397499 40137499	-	-	25P 30P	
9	39129999	70134999	-	4P	-	
9	39292499 39227499 39357499	70167499 39617499 40072499	-	-	12P(1) 23P 24P	
9	40917499	41242499	-	-	24P	
9	40787499 41697499 41892499	43582499 42997499 43192499	-	-	30P 21P(1) 26P	
9	40852499 40722499	70167499 70167499	-	-	23P 25P	
9	42867499	43712499	-	-	23P	
9	43224999	67339999	-	7P	-	
9	43192499	43582499	-	-	21P(1)	
9	44102499 44167499 44167499 44297499	65812499 45207499 65552499 46182499	-		21P(1) 9P 30P 21P(2)	
9	46182499	46702499	-	-	9P, 24P	
9	65454999	65779999	-	12P	-	
9	65747499	70167499	-	-	30P	
9	66787499	68412499	-	-	21P(1)	
10	46117499 46247499 46247499	46832499 47092499 47027499	-	-	12P(1) 12P(2) 24P	
10	46149999	46799999	-	12P	-	
10	50894999	51284999	-	10P	-	
10	132014999 13204799	132404999 132437499	-	9P	 9P	
11	62499	224999	A02	-	-	

Chromosome number	Start position	End position	Array 461678	Array 509338	Array 514596	Mean
13	62009999	62594999	-	10P	-	
14	18427499	19337499	-	-	33P	
14	183949999 18752499 18427499 18427499	19304999 19142499 19077499 19337499	-	9P,15P,17P	- 9P 21P(1) 23P	
14	21644999	21969999	-	20P	-	
15	18394999 18394999	21059999 19954999	-	15P 20P	-	
15	18412499	20274999	A03	-	-	
15	18394999 18394999 18394999 18427499	20214999 19954999 21059999 19987499		2P 10P	-	
15	18427499	20377499	-	-	26P	
15	18427499 18427499	20247499 19987499	-	-	12P(1) 12P(2)	
15	20637499	21092499	-		12P(1)	
15	26112499 26049999 26112499 26194999	26649999 26812499 26824999 26714999	A01 A02 A03	 17P	-	
15	32499999	32694999	-	7P	-	
16	14982499	15372499	-		33P	
16	16152499	16672499	-		23P	
16	22359999	22554999	-	17P	-	
16	2827499	28599999	-	2P	-	
16	28827499 28437499	29477499 30127499	-	-	23P 33P	
16	30647499	34157499	-	-	33P	
16	31812499 31914999 31862499	33712499 33474999 33712499	A01 A02	 13P	-	
16	31687499 31492499 32922499	33377499 32532499 33507499	-	-	24P 30P 30P	
16	32564999 32207499	33474999 32792499	-	2P	 26P	
17	20182499	20767499	-	-	33P	
17	4159999	42119999	-	2P,17P	-	
17	41729999 41794999 41762499	42119999 42119999 42152499	-	8P,15P 9P	 21P(1)	
17	41827499	42152499	-	-	4P,9P 21P(2), 26P	
17	18297499 18752499	18622499 19077499	-	-	26P 26P,33P	
17	73027499	74002499	-	-	30P	
17	75432499	75952499	-	-	30P	
17	76602499	77837499	-	-	30P	
17	78292499	78611061	-	-	30P	
19	48162499	48299999	A01	-	-	
19	17062499 17907499	17452499 18232499	-	-	33P 33P	
19	184227499	18687499	-	-	33P	
19	58174999	60124999	-	4P	-	

Chromosome number	Start position	End position	Array 461678	Array 509338	Array 514596	Mean
20	62172499	62410787	-	-	30P	
21	9887499	13512499	A02	-	-	
21	1345499 13487499	46853544 46886044	-	-	 9P	
21	13487499 13487499	13812499 14072499	-	-	23P,26P,33P, 30P	
21	1345499	46853544	-	9P	-	
22	17012499 17074999	17224999 17224999	A01 A02, A03	-	-	
22	18654999 18787499 18762499 18774999	19077499 19024999 19024999 19024999	A01 A01 A02 A03	-	-	
22	18654999	19077499	-	13P	-	
22	19824999 19824999	20237499 20024999	A02 A01	-	-	
22	23919999	24179999	-	15P	-	

Appendix C

Documentation of the laeverin gene sequence. Forward primers (blue) and reverse primers (pink) used in sequencing analysis are shown.

Laeverin gene sequence
NM_173800.4

```
.....................................................ccacgtttaa
ggtgtccggattcgtttggcttgaccacccctccccgcgccaaccctggaggcatcttc
cgtggggtctgtctcttcgaaccccaaaggggtacgcgtctgggtgagccaggtccgcctc Primer 1_1F
.........................
AGGAGGAAGAGGCACGATACAAGAGAGGAGGGGCAGGGGTCGCAGCACTGAACACCCTGG Exon 1
CCGGGGTTTTGACAGCTGCCACAGTCTCTGAGCTCCAGCCTCGCGCCTGAACCCGGTCCC
TGCCATGGGGCCCCCTTCCAGCTCAGGCTTCTATGTGAGCCGCGCAGTGGCCCTGCTGCT
GGCTGGGCTGGTAGCCGCCCTCCTGCTGGCGCTGGCCGTACTCGCCGCCTTGTACGGCCA
CTGCGAGCGCGTCCCACCGTCGGAGCTGCCTGGACTCAGGGACTTGGAAGCCGAGTCTTC Primer 1_2F
CCCTCCCCTCAGGCAGAAGCCGACGCCAACCCCGAAACCCAGCAGTGCACGCGAGCTAGC Primer 1_1R
GGTGACGACCACCCCGAGCAACTGGCGACCCCCGGGGCCCTGGGACCAGCTACGCCTGCC
GCCCTGGCTCGTGCCGCTGCACTACGATCTGGAGCTGTGGCCGCAGCTGAGGCCCGACGA
GCTTCCGGCCGGGTCTTTGCCCTTCACTGGCCGCGTGAACATCACGGTGCGCTGCACGGT
GGCCACCTCTCGACTGCTGCTGCATAGCCTCTTCCAGGACTGCGAGCGCGCCGAGGTGCG
GGGACCCCTTTCCCCGGGCACTGGGAACGCCACAGTGGGCCGCGTGCCCGTGGACGACGT
GTGGTTCGCGCTGGACACGGAATACATGGTGCTGGAGCTCAGTGAGCCCCTGAAACCTGG
TAGCAGCTACGAGCTGCAGCTTAGCTTCTCGGGCCTGGTGAAGGAAGACCTCAGGGAGGG
ACTCTTCCTCAACGTCTACACCGACCAGGGCGAGCGCAGgtaagggctgtacagcccggg
gcccctctcggcccccgcccctgcgtcccggtgcaggct
gcgggtccagctgactaccgtgtccaggtgcgcgtctgctgcccttccaaagaatcccc Primer 1_2R
..........................................................
gaattctaaggtgccatgataattggagtgtgttggtgtgatgggatgggcagtgaggaa
gatgagaggaaatttcctctaacacaccaggctaacttacaagccatcatctctcaatgt Primer 2F
ttttttattccactgggttgaaataaaaataaataaaaatttccccaaaatgtatttcag
GGCCCTGTTAGCGTCCCAGCTGGAACCAACATTTGCCAGGTATGTTTTCCCTTGTTTTGA
TGAGCCAGCTCTGAAGGCAACTTTTAATATTACAATGATTCATCATCCAAGTTATGTGGC  Exon 2
CCTTTCCAACATGCCAAAGCTAG
gtaagtaatgctttctgtctatatctagctgtctatctatataagctttgtaatcacatt
tataccagtagcttttctaggcagtgtattatggtgtattataattacagtctactataa Primer 2R
..........................................................
gctagactatgaggaatagctggaaaggagtagctgggtgacaattgacttgacaaatta Primer 3F
tttcaaagacatgtaattgcttagatttgctgtgatttgaactaaaataaaattctctag
GTCAGTCTGAAAAAGAAGATGTGAATGGAAGCAAATGGACTGTTACAACCTTTTCCACTA  Exon 3
CGCCCCACATGCCAACTTACTTAGTCGCATTTGTTATATGTGACTATGACCACGTCAACA
GAACAGAAAGGGGCAAGGAG
gtgagtgaggaagattctgtaggtaagggagattgtactggctgcagattattcatctat
tctttctgcatccagtggttatggctgcatatccccaaatcgcttcctagttctctctcc Primer 3R
..........................................................
acttagtgaacagccttcaggggttttctcaggacatgtaagcatgcaagcattttatta
tgcagcttcttctggaaaggtaagacttggttcagcagcagggaattggagcaagcaga Primer 4F
ctacctctttccaaaatcactactggttttcctaatcactgcttaactgttttgatttag
ATACGCATCTGGGCCCGGAAAGATGCAATTGCAAATGGAAGTGCAGACTTTGCTTTGAAC
ATCACAGGTCCCATCTTCTCTTTTCTGGAGGATTTGTTTAATATCAGTTACTCTCTTCCA  Exon 4
AAAACAG
gtgaggtaatcttttcctttcagtgcatttggtttcctgttatttgggcccttggttgag Primer 4R
..........................................................
aataatttatttgctggcttttgtctttggcattagtctttaaaaagccttctctccgtt Primer 5F
caggttataaatattcccttgaatttttttggtaattttatgatttcattttgagatattt
aaaaaatcccatttttttggaaaagattatttttattttctcctccatttaaatccacttag
ATATAATTGCCTTGCCTAGTTTTGACAACCATGCAATGGAAAACTGGGGACTAATGATAT
TTGATGAATCAGGATTGTTGTTGGAACCAAAAGATCAACTGACAGAAAAAAAGACTCTGA  Exon 5
TCTCCTATGTTGTCTCCCACGAGATTGGACACCAG
```

gcatgtggtaaaatgttctttttatttcacttgaagttattctccaggtgcgctaccaaa
atagcataaaaaccccagtaagaccctgccatataccatattttaaaaagaaaaacatct
caaagtatgacagtgagta**aaaatggggaagactcagattca**gggtggatttgagctact **Primer 5R**
.....
gatggcaaataaatttaattatgtcttttgaccttacat**ttacttaacatgcaattaca** **Primer 6F**
acgaatataacttaaaaattaaatttaagaaagctctttttttcttctcatttccaaaag
TGGTTTGGAAACTTGGTTACCATGAATTGGTGGAACAATATCTGGCTCAACGAGGGTTTT Exon 6
GCATCTTATTTTGAGTTTGAAGTAATTAACTACTTTAATCCTAAACTCCCAAGA
gtaagtatgtttactaagtttatttaacatttttctcctaattataaaagtaatgcatat
tcattcttgaaaatgtataaaatacaagaaaatacaagaaataataaattacctataata
ccaaaaataatttgtgttagctttttgatatttt**tttttaattccaggctttttatcta**t **Primer 6R**
.....
tatagttttgttattcaaggaatcctgactttat**ttggcatatttcaactcaaagaca**ct **Primer 7F**
tctcaattactgaattttctccctcttcaatatcagtttttgaaagtattttagggccat
agaattttggtcttctgtgacacctcaggagttaatgtgcctccatcttttccttttatag
AATGAGATCTTTTTTTCTAACATTTTACATAATATCCTCAGAGAAGATCACGCCCTGGTG
ACTAGAGCTGTGGCCATGAAGGTGGAAAATTTCAAAACAAGTGAAATACAGGAACTCTTT Exon 7
GACATATTTACTTACAGCAAG
gtaaaagcagttagaaatttcctttggttttgtactctggtagaaagttgcataaaatgg
attctaagggtcctatatcatcatacaacttaaaacattttatgaaaataaccttatttt
aattcaatacataggtatcagaaaacatgattaaat**taattctcaccttgcaaatatcca** **Primer 7R**
tcaagccaatattaatgcttgggtccatcaaccatagactttaacaacttgctactttta
.....
ttagtgacagcctactattttgttagcacattccaaaataattcaaacattcaactaaaa
t**gcaaaatgcaatcaggagca**cagctcagaatattgcttataaatatggaatgtgtcagt **Primer 8_9F**
atgtggatattgaattttgttcaaataatggacagcttctttttgtcctag
GGAGCGTCTATGGCCCGGATGCTTTCTTGTTTCTTGAATGAGCATTTATTTGTCAGTGCA Exon 8
CTCAAG
gtgagtttgcaaaatagtcgttacctggatggcagtaaacataaattccattggcatgct
atttattttaaccttaacttttctatttttacag
TCATATTTGAAGACATTTTCCTACTCAAACGCTGAGCAAGATGATCTATGGAGGCATTTT Exon 9
CAAATG
gtaattgtcctactttctgacacattcttgctgagttgttttgtatgatacaggaaaaga
ctgggaggccatgggtgaatggca**gttccacaggaaaacagctttg**tcttgttgtagaac **Primer 8_9R**
.....
ctacataagtgaggaagagaatagctac**cgagtttctttttggagattgc**tacttcaaaa **Primer 10F**
tgtttcacggataaccttacctgcctttgtggtgatttttttaaaacag
GCCATAGATGACCAGAGTACAGTTATTTTGCCAGCAACAATAAAAAACATAATGGACAGT
TGGACACACCAGAGTGGTTTTCCAGTGATCACTTTAAATGTGTCTACTGGCGTCATGAAA Exon 10
CAGGAGCCATTTTATCTTGAAAACATTAAAAATCGGACTCTTCTAACCAGCAA
gtaggtagctttgctcctctttgtcttcaccttccttggggttgagctctgacccaatgg
aatccagcctgtgggtga**agaagcgttacccccgct**gggcatacacacttctgcaatgtg **Primer 10R**
.....
agagagaatgaccaaagaaac**tgagttctgtgtgctatttcatca**tctctttaatagtgg **Primer 11F/11F1**
ctcagagtgtatgtttttatgtgtgaaaagtaactaatttttttatttcttccaataag
TGACACATGGATTGTCCCTATTCTTTGGATAAAAAATGGAACTACACAACCTTTAGTCTG Exon 11
GCTAGATCAAAGCAGCA
gtaagtaacaaattttaaaaacatctttatatatatgcatatgtaaaggcagggaataaa
atattgaaaagtctagcttttgaaaagccatt**tcatttcaaactaaaatatcattcttgt** **Primer 11R**
agagcatagagttttaaagaaaagtgtcaaatcaaaagccattatatgtgaaaaatatat
aattacctggtaagagcaatatatattttagaattcctttaatttttttgttttgttggtt
tattgtttagaatcgagtgtgtagtattaataggtattaagatctttttttaaatgtacca
tttcaaattattcccatattccttttatag
AAGTATTCCCAGAAATGCAAGTTTCAGATTCTGACCATGACTGGGTGATTTTGAATTTGA
ATATGACTGGATATTATAGAGTTAATTATGATAAATTAGGTTGGAAGAAACTAAATCAAC Exon 12
AACTTGAAAAGGATCCTAAG
gtaaggttacttttgatacttttaattaaatataatttataatgccttctctagtgtttt
agaagttgaacattctggtgggaag**agattccaccttcattcccg**cccctcacctccagg **Primer 12R2**
catgcaaaagatttttttttcttattcttaacaagtggagattttgattaaatgtgttgt
ctgtgtggttttttttttgttttttttgttttttttttttgagacg**gagtcttgctctgtc** **Primer 12R**
gcccaggctggagtgcagtgacgcgatctcggctcactgcaagctccgcctcccgggttc
acgccattctcctgcctcagcctccggagtagctgggactacaggcgcctgccaccacgc
ccggctaatttttttgtattttttagtagagacggggtttcaccgtgttagccaggatgatc
tcaatctcctgacctggtgatccacctgcctcggcctctcaaagtgctgggattacaggc

gtgagccaccgcgccggcca.....atgtaaagggatctgttggt
ggcattctgtcacgttagtaaaatcttaagataaataccgtgtgaaaaaatatggtacct
caatgttaccaaaatagcgctaaataatatttatttaagggcttctatgccaaaaacatc
cctgaaaataatttggatgttgtaatacaacttgatgtatacttgttttggggaggcctt
ttaaatttggttgcctgttttatttttaacatggtatgtttcaggagtattttggcaaa
actttctgccatatttaaaccttatgattgatgtccagagggttcatttatttcaaaaac
atatgtatgataaataaattatggggactgtcctcagtaattgttgtttctctgcagatg
agataagtcacgt**gaaggtgctttataggcagcag**taatctttaaattgttgttattttg **Primer 13F**
aaaactttgcggaaaaatgttctcatcttcttctgggcatatttgggtctaattctgcag
GCGATTCCTGTTATTCACAGACTGCAGTTGATTGATGATGCCTTTTCCTTGTCTAA Exon 13
gtgagtatattttcttctctcatggtttcagaatataccttgagacctttataaaaataa
tagcttcatcactatgaa**tttgattatttagtctatacaggccat**aactgattaagatta **Primer 13R**
.....
cttgacttattcacatgccttattgaag**tcatgcattgaaacatggaa**caaatatcgaac **Primer 14F**
gtatgcaaattacttagcaagtgaaggtttttgagtgtgtgtgtttttaaatcaaacag
AAACAATTATATTGAGATTGAAACAGCACTTGAGTTAACCAAGTACCTTGCTGAAGAAGA
TGAAATTATAGTATGGCATACAGTCTTGGTAAACTTGGTAACCAGGGATCTTGTTTCTGA Exon 14
GGTGAACATCTATGATATATACTCATTATTAAAG
gtaatttcattctttcttatgtagtttttaaataaatcctctcttcttcttttcatattt
tagccagcatctgtgagacaggtctt**ttccaaagtaatagtcttagttccat**atatatat **Primer 14R**
atatatatatatggaagtatatacatttccttaacaagggctgattgatcttagaagaaa
.....
ttctgaagaaata**tctgccacttgtctatcaatgtc**ttttcagataaataaatagaaaca **Primer 15F**
cagtgaaaattcagtgtttgcaagccagaactaacagtgtattttctttattgtttatag
AGGTACCTATTAAAGAGACTTAATTTAATATGGAATATTTATTCAACTATAATTCGTGAA
AATGTGTTGGCATTACAAGATGACTACTTAGCTCT Exon 15
gtaagtatgttttcagaagtgataattgaataaaatgcattcatattaataggcttccag
aagtaataaaataaaatcttcatatc**tgttattcattgagagattttatatgc**tattatt **Primer 15R**
.....
gatgtcattgaaaggatatttaaaaata**cccatctttttatgaaacacatattc**ttggag **Primer 16F**
gcagagaaaatgaaggtggtaatgcaaaataaactgtttttctttgacttttttcttcaag
AATATCACTGGAAAAACTTTTTGTAACTGCGTGTTGGTTGGGCCTTGAAGACTGCCTTCA
GCTGTCAAAAGAACTTTTCGCAAAATGGGTGGATCATCCAGAAAATGA Exon 16
gtaagagtaatatcataattcctcttgtttttgtcctattttagcaccagcaatttccct
ttttgatgtactcac**tgtgggaatgagtgttttgc**tttcacttggttaggcgctctctcc **Primer 16R**
.....
t**tcctacttttgacctttatatctgtga**ctataaatatttttttctagattcatacttcta **Primer 17F**
aaatatgataacctgggtaaacataactactatgaaaaatatcattttatgttatatttt
acag
AATACCTTATCCAATTAAAGATGTGGTTTTATGTTATGGCATTGCCTTGGGAAGTGATAA
AGAGTGGGACATCTTGTTAAATACTTACACTAATACAACAAACAAAGAAGAAAAGATTCA Exon 17
ACTTGCTTATGCAATGAGCTGCAGCAAAGACCCATGGATACTTAACAG
gtgattatggtcaacttaccttgaaagtttctgttataggaattaaattaataaaggaaa
aaaaatagaaatgt**gctaatgtaagtattaaaacgttgcg**tatttgtgcttcacta**ccttt** **Primer 17R**
agaaccatgggattttgattttgtttatttaaattaactctttatgttggatgtttaaaa **Primer 18F**
tgtatttttgaaaatgcacttttgcag
ATATATGGAGTATGCCATCAGCACATCTCCATTCACTTCTAATGAAACAAATATAATTGA
GGTTGTGGCTTCATCTGAAGTTGGCCGGTATGTCGCAAAAGACTTCTTAGTCAACAACTG Exon 18
GCAAGCTGTGAGTAAAAG
gtaagaaggaaagtgagacctttctttcatttaggccactg**gtttggcactggaagctca** **Primer 18R**
gctttagtctagcttggaagctcagctttagtctagctaggccacaaacgtcctttgcat
.....
ttagatttgtgacataggaattaata**ggggccatacacttgacct**tcatctgctatataa **Primer 19F**
aatatagctttatttttgcaaaatgctttccacccttgattaacatcttattgccttgtag
GTATGGAACACAATCATTGATTAATCTAATATATACAATAGGGAGAACCGTAACTACAGA Exon 19
TTTACAGATTGTGGAG
gtaagtactttaaatattatgaaatacaatgataatttggaaaccacttttatgcaatt
tggacttgctaaaattaaaaataatatgcttatctgaattattacattctatgcttctat
tggtgctgtatcactgtgaaagaagtgtttttaattgtatcctcctctgaaatagagatc
acctgttcatggagggtggcacttcatcc**tctctcacctgaatgcaggg**gactctggccc **Primer 19R**
agggttcccgtcccagaccgggcttccctccatcccaccttgccatcactgccagtcat
.....
tactcattccaaccatcatcaccaattta**caaactcatgttgctacttagcattta**gaca **Primer 20_1F**
tttactttgtccaaatgggaagttggattgcactgatcatctgtctctctgtccttccag

CTGCAGCAGTTTTTCAGTAACATGTTGGAGGAACACCAGAGGATCAGAGTTCATGCCAAC Exon 20
TTACAGACAATAAAGAATGAAAATCTGAAAAACAAGAAGCTAAGTGCCAGGATAGCTGCG
TGGCTAAGGAGAAACACATAGCTTGTGGCTATCTTTCAGCACTCCTCTTGCATATTATAA
TGTAGTTTGTTCACAGTTTTGTCTTCCAATACTTTGTGAGTCTGGAAAACCACACATTTT
ATTTGTATTTCAGTCACATTTATTACTCAGAGTGCCATTCTTCTCATATTGTCATGTTTG
GCCCTGAGGGTGGGTGATTGCTGACAATTTTGCCAATGCTGCTGTATTTCTGGGAAAGAT
GTCACTTCATGTTGGGTTATAATCCCACAGAATTTACTTTAAATGTCACGTAAAAACAAA Primer 20_2F
TTCACCTAAGATAGTCTTGCTTATTTTGTTGCGAAGGCCAGTGGAATATAAAAATCAATG
GCATTAATGAGGAGCACATTCTTTGCTGAGGGAAATAACAGTTTTTCCAGGCCCTAGGGT Primer 20_1R
TTATTTAGTTCAACATTGAAGATTGAAAAGACTAATGAGAAGATAACATGACAATATAAA
AAGACAGAAACTCAAAAAGTATATTCACTAGAAGACAGTGGAGCCAAAGAAAAGCAAGTT
GAAAAAGAAAGAGGAGACAATTTAACGAGCCCTAAAGGCAAGAAGTGTGTCTTCTCCTGC
CCATTTCTCTCTCCCACAGTACATCCTGCATTGAACCACGTGTTGCTGGTTACTAAACTT
TATGGTGTCACAAAAAGACAAATGATTGTAGGATCCATGGAAGTGCACTTAGGCTTGCTG
TGGTGGAGGCCATAGTGGAGATGAAGGCTCTGCTCTGAAGCTCTTAATTGTGTATTCACC
AGTGGGATGGACACCCTTTCAGCCTAAGCGTACGAGTGAAAACAATGATGTCAAAAGACA Primer 20_3F
ATAGCAAGTATAGGCTACTATTTAGGGCAAAAGGAATCTTGAAATGAAAGAAACAGAACT Primer 20_2R
AAAAGGCTAACTGGAGCATAATCCTTACAGTTTTGCACTCAGGGGATTAAGTTTGAAACA
TAGTGTCTTTTGGCCAAATTGAACCTACTTGTAATGTGGACCAACTTACTTTACACATTT
TCATCAGTCCCAATGTATGGAATTAGCTTAGTTTTCTCATGCTTGATTAAAACAAGACAA
CTAAAGCTAATCATTCCTCTTCTAGACTTTGTGTATATGATTTAATCTTTTATTTTTATT
TTTGTAGAATGGGGGTAAAACTACAAATCTCCTCCTCAGGGTGATATTTTTAAAGATTAA
GGAGTATGACATCAATTAACAGTCTTTAAAGATGTTACATACGCTCAAAGCAATGTGATT Primer 20_4F
AAGACTTTTTGCCTATTTCTTGTATAAGTTAAGGCAGCCATAAATGAAAATACAAAACAA
ACGTTTGCTTTTGTATTTGATTCACAAGAGGATTCTCCCTGAGTGTCAGGGGAGGCTGTG Primer 20_3R
GGACTTCACCACGTGGACAAACGAAAGCACAGGCTACTCAGACTTGGCCACAGGTCATAG
GTCAGAGAATTTTTTCTTCAAATGTAGAAATGTAAATTTTCAGCAATAAAATATCTGCAT
GCCTA

caaagtgttgtagctgagataccttcctgtttttttttcttaaacactcattctatatttc 3'Downstream seq.
ccatatttcctctgataatgaatctcatctttcaataactttgtcttcacttgtcctctc
aactgcttttctcttcctcttcttcttctgtgtcttag**tccattttgtgctgctgtca**ta **Primer 20_4R**
gaataccttagactgggtaatttatgaagaacagagatgagtttcttaccattctgaagg
ctggagattc..