Development of a Mass Spectrometry-Based Test Panel for the Molecular Genetic Analysis of CYP21A2 and Its Pseudogene for the Diagnostics of CAH

Vom Promotionsausschuss der Technischen Universität Hamburg-Harburg
zur Erlangung des akademischen Grades
Doktor-Ingenieur (Dr.-Ing.) vorgelegte Dissertation

von

Cumhur Cantürk

aus

Uzunköprü, Türkei

2013

Members of thesis committee:

1. Prof. Dr.-Ing. Ralf Pörtner
2. Prof. Dr. Wolfgang Höppner
3. Prof. Dr. rer. nat. Rudolf Müller

Date of oral examination:

13th December, 2012

Acknowledgements

I am truly indebted and thankful to Prof. Dr.-Ing. Ralf Pörtner, who kindly accepted to supervise me throughout my dissertation for three years. This work would have not been this successful unless you had given your time and effort.

I owe sincere and earnest thankfulness to Prof. Dr. Wolfgang Höppner for agreeing as CEO that Bioglobe GmbH take part in this dissertation. The supports you gave without hesitation as a boss, professor and sometimes even more made the time pass by more easily and enjoyably, and will always be remembered.

I would like to show my gratitude to Dr. Niels Storm, who was always there when I needed with no exceptions at all. I do not have the slightest doubt that this PhD thesis would have not come to an end if you had not been there with your expert knowledge and your friendly manner.

It is a great pleasure for me to thank the whole Bioglobe GmbH team, with whom I had such a nice time together. Among them are Ulrike Baade, Dr. Ramona Salazar, Kerstin Bartels and Dr. Boris Leuenberger. Thank you all wholeheartedly so much for all your support and kindness at all times.

Last but not the least, I am obliged to my family and many friends of me who did not hesitate to show their support and love whenever I needed them. This page is not enough to name you all, but surely, you know who you are.

Abstract

Congenital adrenal hyperplasia (CAH) with a worldwide incidence rate between 1:5000 – 1:10000 is among the most common autosomal recessive disorders. The classical form of the disease can cause death by salt crises in the first weeks of life. More than 95% of the cases are due to 21-hydroxylase deficiencies, which is the product of CYP21A2. CYP21A2 is located in a highly variable region on the short arm of chromosome 6. High variability and the presence of a highly homologous pseudogene approximately 30 kb upstream, CYP21A1P, are the main hindrances in molecular analysis of the gene and of this region.

Today, there are more than 170 rare (family or individual-specific) mutations and more than 50 SNPs (single nucleotide polymorphisms) reported. Besides mutations and SNPs, copy numbers of CYP21A2 and CYP21A1P play an important role in a complete genetic analysis. Most commonly applied methods are direct gene sequencing (Sanger sequencing) for mutation analysis; MLPA, Southern blotting and Real Time PCR for copy number variation (CNV) analyses. Each method requires a considerable amount of time and labor.

In scope of this dissertation, it was made possible to analyze CYP21A2 and CYP21A1P for mutations and SNPs together with CNVs on Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF MS) platform. Two experimental set-ups for variation detection analysis and one set-up for quantitative analysis were designed using MassARRAY kits from Sequenom GmbH. As a result, a total of 134 mutations and SNP-like variations in CYP21A2 could be detected automatically. A match rate of 98.61% was achieved in comparison to Sanger sequencing results. In all mismatches, discrepancies were found to be related to raw data interpretation in gene sequencing. 3 differentiating alleles were chosen for gene copy number ratio analysis; in Exon 1, Intron 2 and Exon 6. For absolute gene copy number experiments, a reference gene, SLC30A3, was chosen for normalization. Both gene copy number ratio and absolute gene copy number analyses delivered considerably more consistent and safer result compared to MLPA.

In conclusion, mutation detection and quantitative analysis were performed successfully on the same platform, which by other methods requires different platforms hence longer experiment and data interpretation times. Moreover, analyses covering a wide range of variations were made less operator-dependent and remarkably more automated. Finally, owing to the nature of mass spectrometry, the complete analysis became safer and more accurate.

Table of Contents

List of Figures

List of Tables

Abbreviations

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
A	Adenine
G	Guanine
C	Cytosine
T	Thymine
I (i)	Inosine
PCR	Polymerase chain reaction
SNP	Single nucleotide polymorphism
Mut (Mt)	Mutation
del	Deletion
ins	Insertion
CAH	Congenital adrenal hyperplasia
CFTR	Cystic fibrosis transmembrane conductance regulator
21-OHase	21-hydroxylase
17-OHP	17-hydroxyprogesterone
ACTH	Adrenocorticotropic hormone
ng	Nanogram
ml	Millilitre
mg	Milligram
µl	Microlitre
nl	Nanolitre
µM	Micromolar
U	Unit
min	Minutes
sec	Seconds
NC	Negative control
hom	Homozygous
het	Heterozygous
wt (WT)	Wild type
CVS	Chronic villous sampling
PRA	Plasma rennin activity
SW	Salt wasting

SV	Simple virilization
NC	Non-classical
HLA	Human leukocyte antigen
CNV	Copy number variation
Cd.	Codon
rev	Reverse
fwd	Forward
UEP	Unextended extension primer
MLPA	Multiplex Ligation-dependent Probe Amplification
MALDI-TOF MS	Matrix-assisted laser desorption/ionization mass spectrometry
SAR	SNP Allele Ratio
ACN	Absolute Copy Number
ASCN	Allele Specific Copy Number
Da	Dalton
3-HPA	3-hydroxypicolinic acid
hMC	Homogeneous MassCLEAVE
hME	Homogeneous MassEXTEND
SAP	Shrimp alkaline phosphatase
BSA	Bovine Serum Albumin
DMSO	Dimethyl sulfoxide
TBE	Tris/Borate/EDTA
EDTA	Ethylenediaminetetraacetic acid
UTR	Untranslated region
E6	Exon 6
bp	Base pair
kb	Kilobase pair
Amp	Amplicon
NCBI	National Center for Biotechnology Information
SLC	Solute-carrier

1 Objectives

Congenital adrenal hyperplasia (CAH) is one of the most common autosomal recessive disorders. The disease manifests due to 21-hydroxylase deficiency. 21-hydroxylase has an important function in keeping the balance of adrenal steroid hormones. It is the product of CYP21A2 gene. Due to mutations in CYP21A2, the enzyme activity of its final product might be reduced in different extents.

CYP21A2 is located on the short arm of the sixth chromosome in human genome. This region is called Human Leukocyte Antigen (HLA) major histocompatibility complex. A pseudogene, CYP21A1P, is present about 30 kb upstream of CYP21A2, which is inactive due to the mutations it possesses. Two genes being highly homologous in all exons and introns represents a big difficulty in molecular biological analysis. CYP21A2 becomes fully or partly inactive via the transfer of inactivating mutations from CYP21A1P by different intergenic mechanisms.

In most genetic laboratories mutation detection in CYP21A2 is performed usually by direct gene sequencing applying “Sanger” method. This is a long experimental process. Results can be delivered in days and the evaluation phase is highly operator-dependent. The level of automation is very low. In addition, relatively long size (3.4 kb) and high number of exons increase the risk of failing to notice a base exchange especially in a homozygous state.

CYP21A2 is located in a genomic region where copy number variations (CNVs) take place with a high frequency. Deletion of an approximately 30 kb portion, starting from the third exon of CYP21A1P and spanning to the third or eighth exon of CYP21A2, results in “hybrid genes”. Independent gene duplications or deletions with or without hybrid gene formation are observed both in CYP21A2 and CYP21A1P. To conclude on protein expression levels, it is necessary to perform gene copy number experiments as well as the mutation detection analysis.

The most up-to-date and common method to analyze CYP21A2 and CYP21A1P quantitatively is Multiplex Ligation-dependent Probe Amplification (MLPA), which

utilizes gene-specific probes, universal PCR primers for probe amplification and a special software to interpret the final output. Due to the long ligation time, results cannot be obtained within one day. Moreover, due to the high homology between both genes some probes have been shown not to be specific enough for a safe quantitative analysis in a diagnosis environment.

Taking into account the drawbacks of the present methods explained above, it is aimed with this dissertation work, to transfer the process of detecting inactivating mutations and other relevant variations to mass spectrometry platform. The accuracy of Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight (MALDI-TOF) MS would minimize the rate of false data interpretation and automize the process. To achieve this goal, it is planned to adjust homogenous mass cleavage and homogenous mass extension approaches from Sequenom GmbH specifically to CYP21A2 and its pseudogene CYP21A1P, and utilize them together. This novel idea of combining two approaches would double the amount of final information obtained without increasing the workload two-fold. This goal would be achieved by designing gene-specific primers for long-range PCR and nested PCR primers for sub-amplification. Primers which would be included in mass extension reactions and assay group design would be designed specifically by introducing base modifications where necessary. Each PCR cycling condition and recipe would be modified in order to make the combination of both approaches possible and increase the final experimental yield. The new set-up would have the advantage of being highly accurate and comparably more automated as compared to conventional Sanger sequencing. The benefit of this is expected to be a remarkable decrease in operator errors, the number of repeat experiments due to dubious cases and data evaluation time.

The second goal of this work is to design a reliable quantification analysis of both CYP21 genes on mass spectrometry platform. This would eliminate the need for an extra experimental kit which would reduce operation costs drastically. This goal would be achieved by determining gene-specific genotypes at certain positions, which are conserved, thus differentiate between CYP21A2 and CYP21A1P at all times. Mass extension with single base stop (iPLEX) would be performed to obtain peak area data, and a gene on a different chromosome would be chosen within the

sample itself as internal control purposes. This would make the normalization of ratio information and the assignment of absolute copy number possible. Since the positions chosen for analysis would be completely reliable, the set-up would not have the problem of suffering from not being specific enough. Moreover, without the necessity of a long ligation step, results would be obtainable within the same working day.

Both set-ups were designed within the frame of this work. They were run in parallel with the routine analysis which included real samples, and the results were compared with each other. Parallel blind studies were continued until the improvement and optimization studies were finalized, which formed the final configuration of the designed set-ups. Finally, the traditional method and the newly designed mass spectrometry diagnostic kits were compared to each other with respect to their advantages and disadvantages.

2 Introduction

2.1 Basics

DNA (deoxyribonucleic acid) is the hereditary material in living organisms where genetic information is stored and passed on to next generations. It has a double helix structure and is made up of a sugar-phosphate backbone to which nitrogenous bases are bonded. Genetic information is stored in the form of a code which is determined by four nitrogenous bases; Adenine-(A), Guanine-(G), Cytosine-(C), Thymine-(T), and how they are ordered; that is, their sequence. With the simplest definition, a gene is a meaningful portion of human genome which encodes for a functional protein via transcription into RNA and translation into amino acid chains.

The introduction of Polymerase Chain Reaction (PCR) in 1983 was revolutionary to analyzing several diseases from molecular biological aspects. *In vitro* replication of DNA templates using primers enabled scientists to read base sequences of regions throughout the genome and detect base changes, sequence repeats or haplotype

motives which have unique meanings. Only seven years later, in 1990, the rapid improvement in the field resulted in the commencement of Human Genome Project (HGP) initially headed by U.S. National Institutes of Health. The project was completed in June 2000 and the genetic map of man was sketched [1, 2]. This brought in an important step in medicine and health care development. Many PCR-based molecular techniques were developed shortly afterwards.

Less than 0.5% difference in over 3 billions of nucleotide bases in human genome is what makes every human being unique. If a certain variation is observed with a frequency of equal to or greater than 1% in a certain population, it is named a Single Nucleotide Polymorphism (SNP). SNPs are known to occur on average once in every 100 – 300 base pairs and they are considered to have no negative effects on the metabolic level although some might have an influence in increasing genetic predisposition for certain diseases. Variations with a frequency of lower than 1% are termed mutations. Mutations might affect a single base or a large chromosomal segment. They are either inherited (germline mutations) or gained in life (somatic mutations) due to different environmental factors. They can be beneficial, neutral or harmful. Large-scale mutations are mutations which cause duplications and deletions of large chromosomal regions, chromosomal translocations, inversions and loss of heterozygosity. Small-scale mutations are mutations which affect smaller segments and a few nucleotides. They can be classified into three main groups: point mutations, small insertions and deletions. Insertions and deletions add or remove one or more nucleotides to or from the DNA, respectively, causing frame shifts in the coding sequence. Point mutations can be silent mutations, which do not result in a different amino acid to be coded, missense mutations, which result in a different amino acid to be coded, and nonsense mutations, which result in the terminating amino acid to be coded hence premature termination of the chain.

Some variations in the human genome do not have any clinical impact on individuals but they rather determine features like final body height, eye color, hairiness, smell, etc. Some; however, do cause severe diseases which are called genetic disorders. Genetic disorders might be triggered by a variation (or combination of more than one variation) alone or in combination with environmental factors. If the presence of a variation on one allele alone (heterozygosity) is sufficient for the genetic disease to

be observed in the phenotype, that disease is called a dominant disorder. Huntington's disorder, hypercholesterolemia, Marfan's syndrome are among such diseases. If both alleles of a certain position are required to be mutated (homozygosity or compound heterozygosity) for the disease to manifest, the disease is called a recessive disorder. A few examples are cystic fibrosis, Werner's syndrome, sickle-cell anemia, galactosemia, growth hormone deficiency, autism, and congenital adrenal hyperplasia.

2.2 Congenital Adrenal Hyperplasia (CAH)

2.2.1 Definition and Biochemistry

The term CAH (OMIM +201910) embraces a number of autosomal recessive disorders which stem from deficiencies in enzymes involved in the steroidogenic pathways of cortisol biosynthesis. In approximately 95% of the cases, 21-hydroxylation is disturbed because of deficiencies in 21-hydroxylase activity (CYP21, 21-OHase), where 17-hydroxyprogesterone is prevented from being converted to 11-deoxycortisol, and progesterone to deoxycorticosterone. The complete pathway is presented in Figure 1 together with catalyzing enzymes [3]. The enzymes involved belong to the cytochrome P450 family, which contains diverse enzymes, and control steroidogenesis. The process is stimulated by adrenocorticotropic hormone (ACTH) which is then reverse-blocked by one of the final products, cortisol. When cortisol production is hindered due to 21-OHase deficiency, ACTH is not blocked and becomes present in excess, which keeps the adrenal cortex continuously stimulated. This is called hyperplasia of the adrenal cortex. Another result of 21-OHase deficiency is the increase in the amount of 17-hydroxyprogesterone (17-OHP) and progesterone. Unlike these steroid precursors, aldosterone and cortisol cannot be produced sufficiently, whereas being the product of the uninterrupted pathway adrenal androgen is produced in excess.

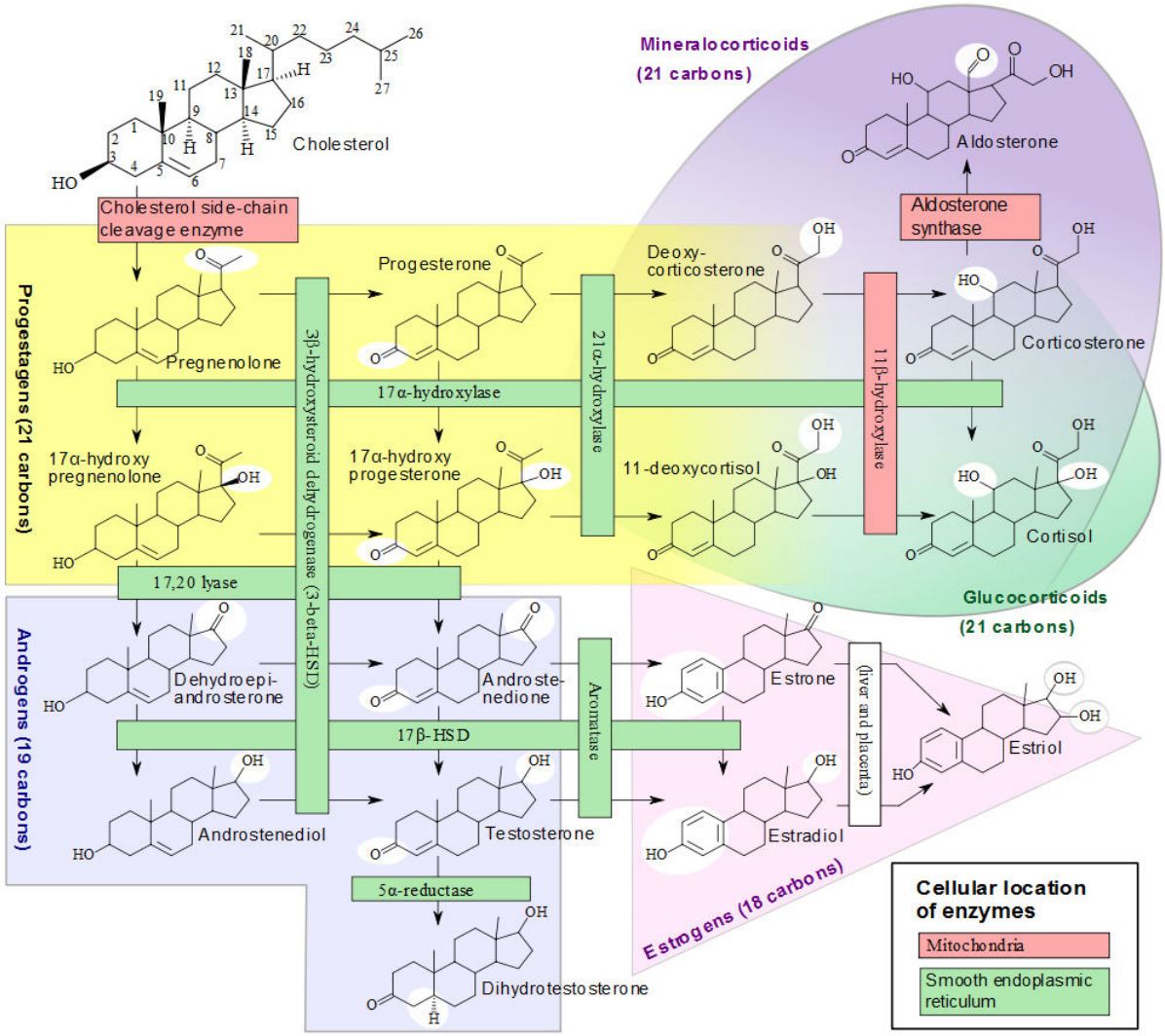


Figure 1. Schematic representation of pathways in biosynthesis of steroid hormones [3].

2.2.2 Types and Incidence Rates

CAH due to 21-OHase deficiency is classified into two groups: Classical and non-classical (mild or late-onset). Classical form of the disease can manifest in simple virilizing and/or salt wasting forms. Depending on the severity of the enzyme deficiency, masculinization of external genitalia to various extents is observed in females with simple virilizing form. Internal female organs are not affected and the karyotype is 46XX. In salt wasting form, due to insufficient production of aldosterone, sodium cannot be retained in the kidney, colon and sweat glands. This is observed in both sexes. If not diagnosed and treated in time, it usually causes acute adrenal crisis, which might result in sudden death in the early weeks of life. While ambiguous

genitalia usually make early diagnosis and treatment in girls possible, this might be missed in boys [4]. Individuals with non-classical CAH do not present ambiguous genitalia at birth and they do not suffer from salt-wasting crisis. They usually present premature pubarche, cystic acne, hirsutism and oligomenorrhea [5]. This form can manifest anytime between later childhood and young adult life.

Although the incidence rate is highly dependent on ethnicity and geographical area, the worldwide average rate for classical CAH ranges between 1:8000 – 1:16000 [6, 7, 8]. The Yupik Eskimos (Alaska) and the people of La Réunion (France) have been reported to show higher rates; 1:282 and 1:2141, respectively [9]. New Zealand population shows the lowest rates; 1:23344 [10]. The frequency of the non-classical form is the greatest for Ashkenazi Jews; 1:27. It is followed by Hispanics; 1:53, Yugoslavs; 1:63, Italians; 1:333, and other Caucasians; approximately 1:1000 [6].

2.2.3 Diagnosis and Treatment

The diagnosis is usually performed by measuring the elevated serum levels of 17-OHP, which is one of the substrates of 21-OHase, in blood spots on filter paper. Depending on the severity of the disease, 17-OHP levels can go up to 1000 ng/ml or even higher in salt-wasting type, whereas it ranges between 10-20 ng/ml in healthy individuals. Lower levels of aldosterone in serum or urine and higher levels of ACTH are other confirmatory measurements [9]. Since patients with the non-classical form of the disease show mildly high (>200 ng/ml) or normal levels of 17-OHP, it is believed that ACTH stimulation test must be done to support diagnosis [9, 11, 12]. Another helpful tool in confirming or discarding the diagnosis in cases of slightly elevated 17-OHP levels is neonatal genotyping. This also reduces the risk of over-treatment of mildly affected children [13]. Prenatal diagnosis is possible with DNA analysis from the amniotic fluid obtained by chorionic villous sampling (CVS). The risk of contamination of the fetus DNA with maternal DNA should be considered. Therefore parents are usually included in the genetic analysis as well. If the fetus is female and the risk is considerable, treatment with dexamethasone is started as soon as the pregnancy is confirmed. This must be done before genetic testing because CVS can be performed in the 10^{th} week of gestation but virilization starts already around the 7^{th} week. The treatment is stopped in case of a male or an unaffected

female fetus, it is continued throughout pregnancy in case of an affected female fetus [14]. In about 85% of the cases females are born with normal to slightly virilized genitalia after prenatal treatment [15]. Since the probability of having an affected female is one in eight when both parents are known carriers, this would mean seven fetuses out of eight are exposed to unnecessary dexamethasone treatment until the treatment is stopped. This is the reason why prenatal therapy still remains to be controversial.

Treatment involves glucocorticoid and mineralocorticoid replacement in classical forms. In cases of extreme masculinization of genitalia surgical correction might be necessary. Clinical and biochemical parameters should be monitored continuously along treatment. Typical glucocorticoid replacement used in younger children is hydrocortisone although in some countries where it is not available, cortisone acetate is also used. As cortisone must be converted to cortisol for biological activity and differences in conversion rate influence drug efficacy, cortisone acetate is mostly not recommended. Excessive glucocorticoid treatment in newborns should be prevented since it leads to early growth suppression. An initial dosage of 20-25 mg/m^2/day is recommended and may be reduced to 10 mg/m^2/day in the next two years as body surface area increases [16]. Longer-lasting glucocorticoids, such as prednisolone and dexamethasone, are often used in adult CAH patients due to better compliance [17]. However, once-a-day dosage of dexamethasone showed to provide normal growth with no considerable side-effects for children with a bone age of or younger than 3.5 years [18]. Patients with salt-wasting form of the disease require life-long mineralocorticoid treatment. This is mostly compensated by oral fludrocortisone in a dose that suppresses plasma rennin activity (PRA) to normal levels without inducing hypertension. In addition, these individuals need additional salt supplements to maintain plasma sodium concentration and PRA in the normal range. This supplement treatment should be carefully controlled, especially in infancy [19]. Since 21-OHase deficiency is a genetic disorder caused by a single gene defect, gene therapy might be feasible. A single intra-adrenal injection of an adenoviral vector encoding 21-OHase was shown to compensate for the biochemical, endocrine and histological alterations in 21-OHase deficient mice. However, due to the possibility of immune reactions and tissue damage which can be induced by the large doses of such vectors, the topic remains to be controversial [20].

2.2.4 Genetics of CAH

2.2.4.1 Chromosome 6 and RCCX Module

21-OHase is the product of 21-hydroxylase gene (CYP21A2, GeneID 1589). CYP21A2 is located within the third class of major Human Leukocyte Antigen (HLA) histocompatibility complex on the short arm of chromosome 6 (6p21.3). A highly homologous pseudogene, CYP21A1P, is mapped approximately 30 kb upstream in the direction of telomere. Both genes are arranged adjacent to and alternating with the genes which encode the fourth component of the serum complement system, C4A and C4B [21].

Figure 2 shows a schematic representation of the structure of the chromosomal region. This genetic unit is designated as the RCCX module (RP-C4-CYP-TNX). Most chromosomes (71.6%) have a bimodular RCCX unit with two RP, C4, CYP and TNX genes; however, monomodular (16.2%) and trimodular (12.2%) haplotypes are reported in Caucasian populations as well [22].

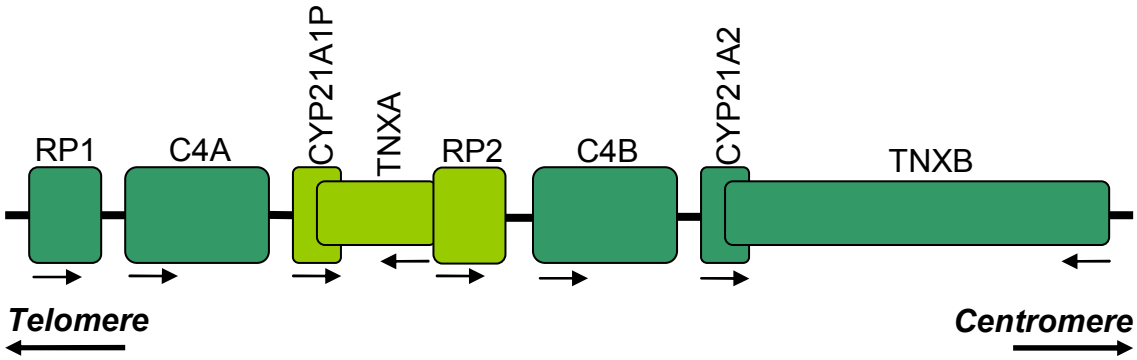


Figure 2. Schematic representation of HLA Class III on chromosome 6p21.3: Genes encoding for a protein are shown with dark green boxes, pseudogones which do not encode for a protein are shown with light green boxes. Names are indicated above the boxes. The orientation is shown with **bold** arrows at the bottom of the figure. Arrows below the boxes show the direction of transcription.

Although CYP21A2 and its pseudogene CYP21A1P are 98% and 96% homologous in their exons and introns, respectively, certain vital differences, which have accumulated in the evolutionary time in CYP21A1P make it an inactive gene [23, 24].

The product of CYP21A1P is a transcribed, but a truncated, nonfunctional amino acid chain. In about 95% of CAH cases mutant alleles are the result of two types of recombination between CYP21A2 and CYP21A1P: Unequal crossing-over during meiosis, which results in a 30 kb deletion spanning from the 3'-end of CYP21A1P, through all of C4B to the 5'-end of CYP21A2, and the transfer of deleterious mutations normally present in CYP21A1P to CYP21A2, which is termed as gene conversion [25]. These recombination events occur when a loop forms in the linear DNA during crossing-over. This lets two genes face each other and exchange, lose or duplicate genetic material (see Figure 3). This is a normal phenomenon in the HLA locus which increases immunological diversity. About 5% of affected alleles possess spontaneous mutations (rare mutations) which are family or individual-specific and are not found in CYP21A1P. The number of such mutations is greater than 100 and continues to increase every day.

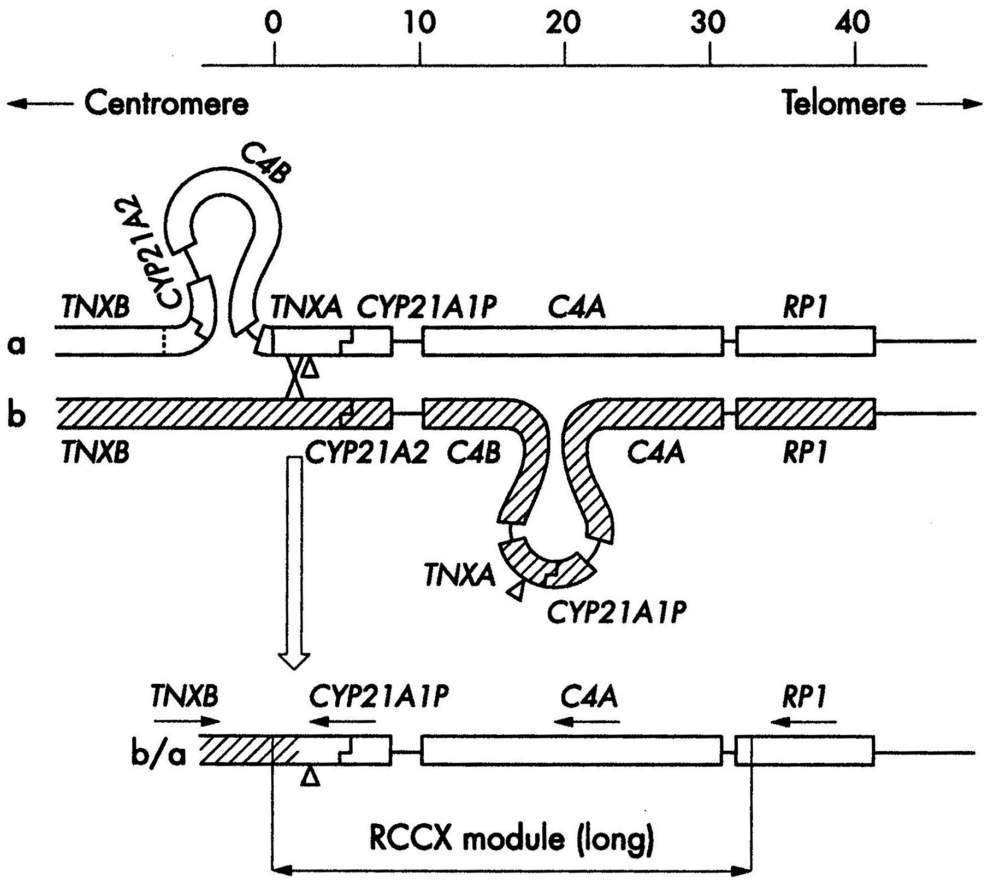


Figure 3. Schematic representation of unequal crossing-over in the HLA region resulting in a 30 kb deletion with a chimeric CYP21 gene, having a CYP21A1P-like 5'-end and a CYP21A2-like 3'-end [26].

2.2.4.2 Inactivating Mutations

The frequency of unequal crossing-over events that cause 21-OHase deficiency is about 20-25%, whereas 70-75% is caused by gene conversions, which transfer mutations from CYP21A1P to CYP21A2 mainly during mitosis [27]. The ten most common pseudogene-derived mutations and the type of CAH they cause are summarized in Figure 4. These mutations are found with similar frequencies in most populations although some are detected with higher frequencies in certain populations. For example, p.Gln318X is found in Tunisian CAH population with a higher frequency (35.3%) compared with a maximum of 13.8% in other studies [28]. A study showed that p.Val281Leu was prevalent in Ashkenazi Jews, p.Arg356Trp in the Croatians, IVS2AS,A/C-G,-13 (Intron2-splice site) in the Iranians and Yupik-speaking Eskimos of Western Alaska [29].

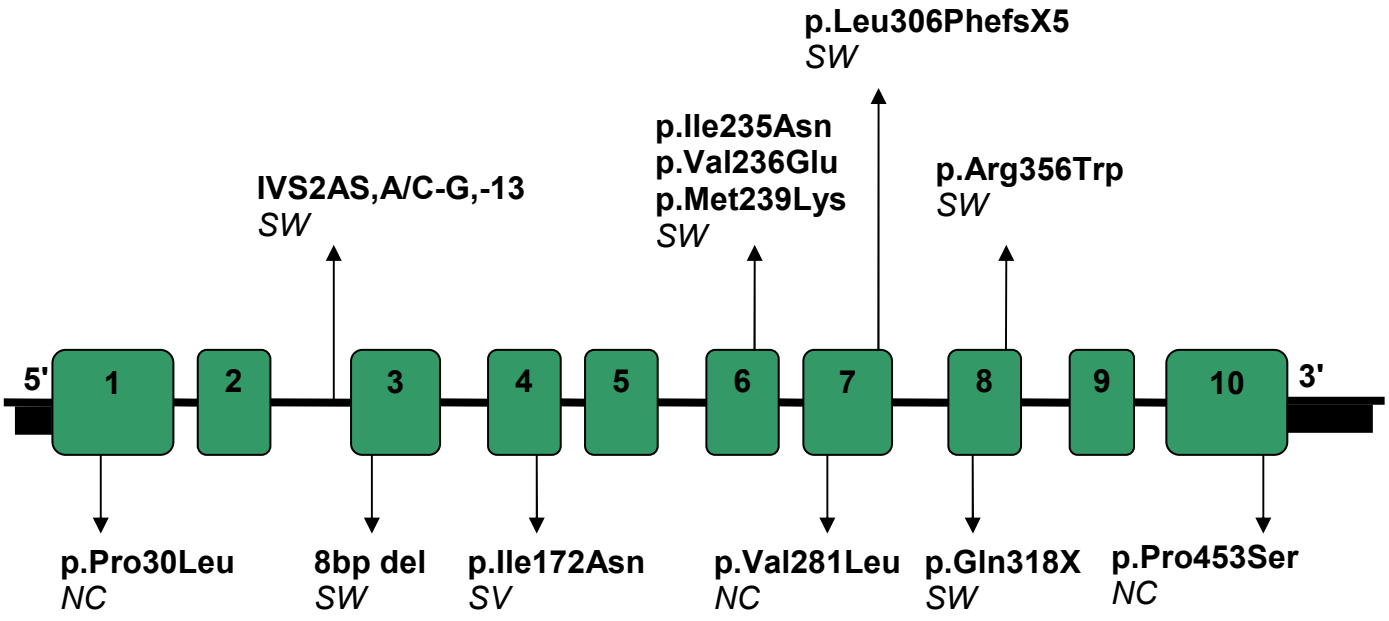


Figure 4. Schematic representation of CYP21A2: Exons are numbered and shown in dark green color. Introns are shown in bars. The ten most common mutations, their positions and resulting phenotypes are designated. 5'- and 3'-untranslated regions are indicated by rectangles under introns. NC: Non-classical, SV: Simple virilizing, SW: Salt-wasting.

2.2.4.3 Genotype-Phenotype Correlation

About 65-75% of CAH patients are compound heterozygous. Since this genetic disorder is recessive, the clinical expression is determined by the less severely

affected, “mild” allele. In most cases the correlation between genotype and phenotype is reported to be high [30, 31, 32]. 8-bp deletion in exon 3 is associated with salt-wasting type of 21-OHase deficiency. p.Ile172Asn in exon 4 is the most common cause of simple virilizing form of the disease. Intron2-splice site mutation (IVS2AS,A/C-G,-13) results in classical CAH (depending on whether or not accompanied by other mutations, simple virilizing or salt wasting). Mutations p.Ile235Asn, p.Val236Glu, and p.Met239Lys are named altogether as exon 6 cluster. They abolish the enzyme activity and cause salt wasting form of the disease to manifest. Mutations p.Gln318X and p.Arg356Trp in exon 8 together with the mutation in exon 7, p.Leu306PhefsX5, lead also to the classical form of CAH. Mutations p.Pro30Leu, p.Val281Leu and p.Pro453Ser are associated with the non-classical form of the disease [5].

2.2.5 Copy Number Variations at CYP21A2 Locus

It was until the year 2004 that sequence polymorphisms were thought to be the major source of individual variability and account for normal phenotypic variation. In this year, two study groups, Iafrate et al. and Sebat et al., announced high occurrence rate of copy number variations (CNVs) in genomes of normal individuals [33, 34]. Major research followed this announcement in the next two years which added invaluably to our knowledge of CNVs.

In 2006, researchers working on the International Genome Structural Variation Consortium’s Copy Number Variation Project generated the first CNV map of human genome. Analyzing 270 subjects with European, African and Asian origin, they ended up in 1447 CNVs which possess many genes, disease loci, indels, duplicons, segmental duplication, etc. This constitutes roughly 12% of the human genome [35].

Haploinsufficiency is a condition that results in delay or impairment in development when one copy of a dosage-sensitive gene is deleted [36]. According to Qian et al., the number of haploinsufficient genes is at least 51, and probably more [37]. These genes are usually genes that take part in the immune system and in brain development and functionality. Among CNV related diseases one can name Down’s

syndrome (Trisomy 21), breast cancer (HER2), AIDS (CCL3L1), Charcot-Marie-Tooth disease (PMP22).

It is predicted that haplotypes with duplicated CYP21A2 genes could have a predisposing role for *de novo* aberrations for heterozygous CYP21A1P deletions in combination with standard RCCX units. Therefore, detection of CNV haplotypes for CYP21A2 and CYP21A1P is of importance in terms of strategies for prenatal CYP21A2 genotyping and genetic counseling. Moreover, duplications of CYP21A2 have been linked with being a risk factor for *de novo* mutations in the offspring [38].

Most chromosomes possess two RCCX modules; that is, a bimodular RCCX unit, which carries one copy of CYP21A1P and CYP21A2. Despite a bimodular unit being the standard, monomodular, trimodular units and in some very seldom cases even four RCCX modules (a quadrimodular unit) are observed [22, 39, 40]. A monomodular RCCX unit might have a 30 kb deletion starting from between exon 3 or exon 8 of CYP21A1P, through the complete TNXA, RP1 and C4B, ending at the corresponding position in CYP21A2, which results in a chimeric CYP21 gene. Trimodular RCCX units have most of the time duplicated CYP21A1P genes; however, haplotypes with duplicated CYP21A2 genes and a single copy of CYP21A1P were also reported [41, 42]. Figure 5 shows the three most common RCCX module structures, mono-, bi- and trimodular units.

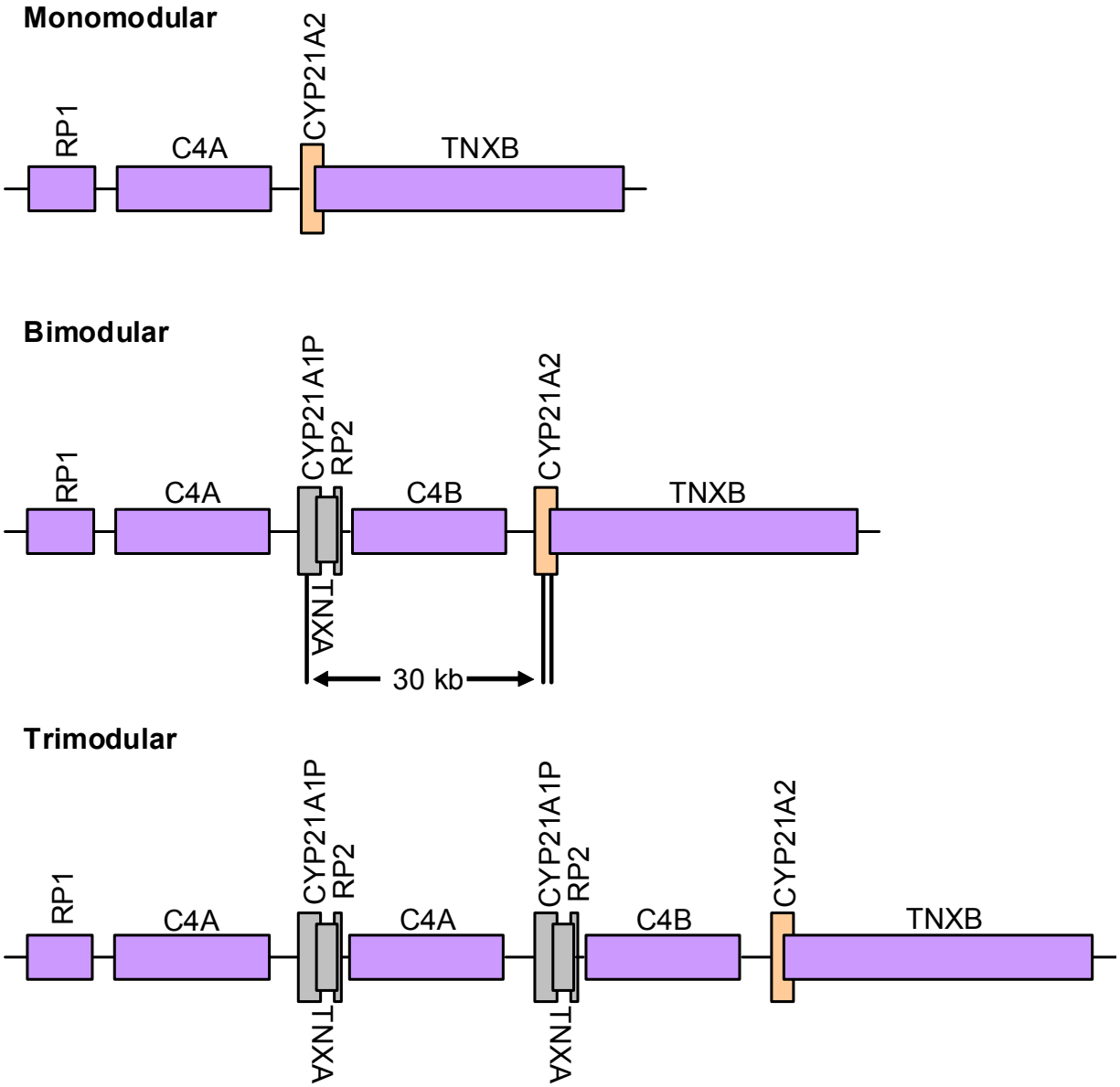


Figure 5. Mono-, bi- and trimodular RCCX units: Horizontal bars in bimodular form indicate the approximate span of 30 kb deletion. Duplication of CYP21A1P is presented in trimodular form.

As far as only CYP21 genes are concerned, three different chimeric gene (fusion/hybrid gene) structures are possible to be observed as the result of unequal crossing-over during meiosis and gene conversion events during mitosis. These are summarized in Figure 6 [43]. The only difference between the first and the second haplotype in the figure is the presence and absence (30 kb deletion) of TNXA, RP2 and C4B in haplotype (1) and (2), respectively, which is not shown in the illustration. 30 kb deletion continues until exon 8 in haplotype (3), and until exon 3 in haplotype (2).

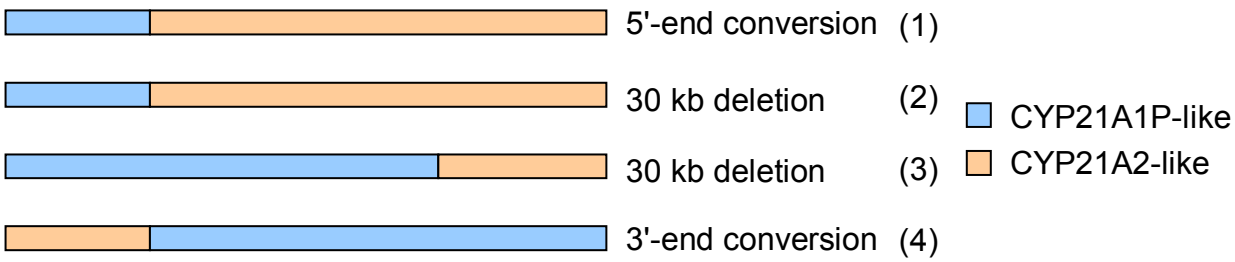


Figure 6. Four different haplotypes which might result after recombination events between CYP21A2 and CYP21A1P: Blue and pink colors indicate CYP21A1P- and CYP21A2-like regions, respectively.

One of the most commonly used methods to assign copy numbers to CYP21A2 and CYP21A1P is the commercially available SALSA MLPA KIT P050-B2 CAH from MRC-Holland. MLPA, Multiplex Ligation-dependent Probe Amplification, is a technique which is based on the amplification of specific probes that hybridize to their specific target sequences. Once all probes are hybridized, just one universal primer pair amplifies them all since all probes have the same 5'-end and 3'-end tag sequence. Two probes bind immediately to adjacent target sequences. The determining step here is the ligation step. When both probes are hybridized, they can be ligated by a thermostable ligase and amplified exponentially in the next step by universal primers. So, only ligation products are amplified in the PCR step. Finally, the intensity of the resulting peaks after capillary electrophoresis separation is a measure of the amount of ligated product, thus the copy number (Figure 7) [44]. The specific MLPA kit for CAH consists of 5 probes for CYP21A2 (exons 1, 3, 4, 6 and 8), 3 for CYP21A1P, 3 for TNXB, 1 for C4A, 1 for C4B, and 1 for CREBL1. A total of 19 control probes are included, 2 of which are located on chromosome 6p21.3, 1 on Y-chromosome and 16 elsewhere on human genomic DNA.

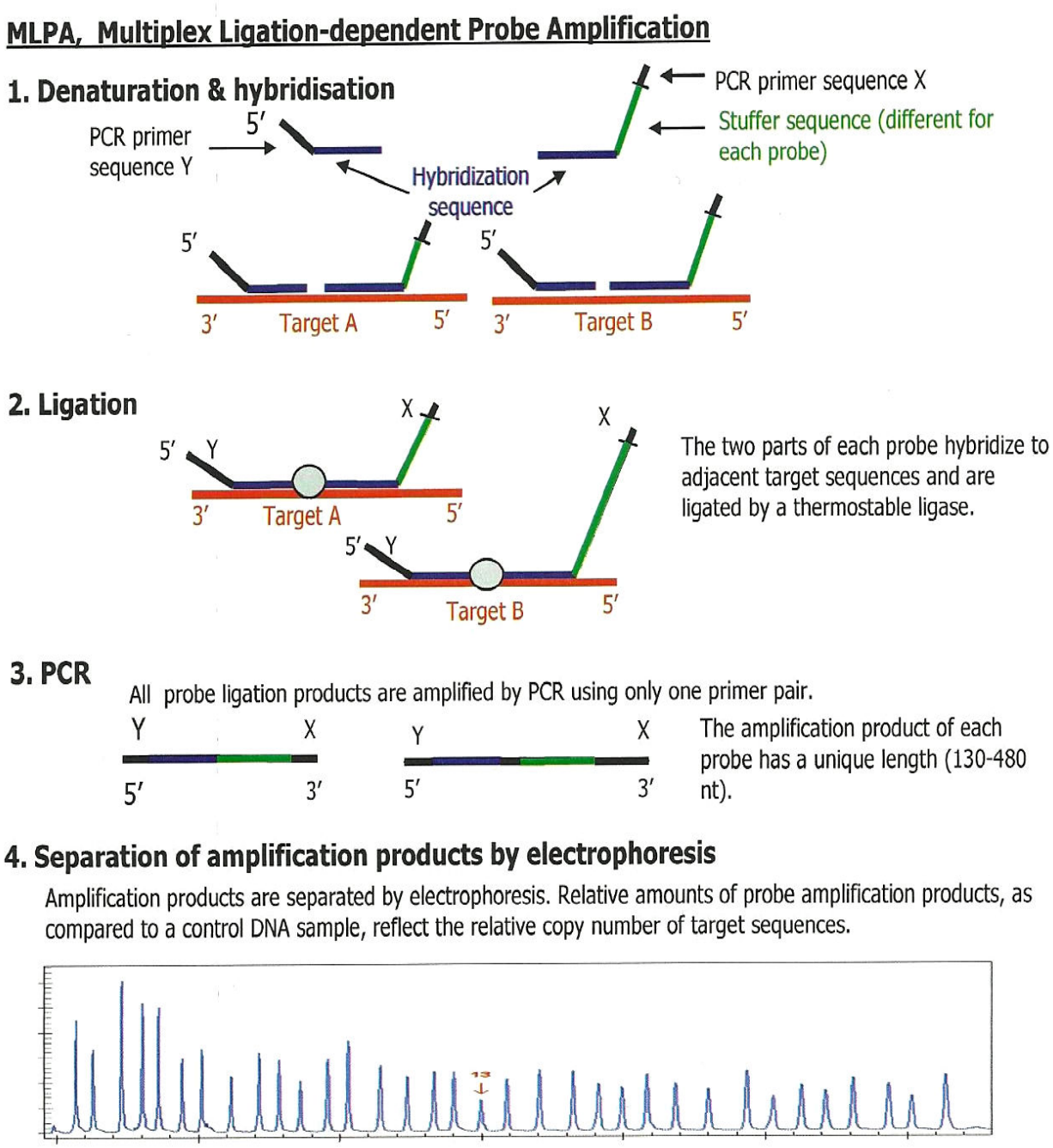


Figure 7. Schematic representation of MLPA protocol: Hybridization, ligation, amplification and peak acquire steps are shown in order of practice [44].

2.2.6 CNV analysis using MALDI-TOF MS

Sequenom, Inc., which is one of the leaders in the field of genetic diagnostics and which provides genetic analysis solutions by means of mass spectrometry, has developed three different approaches to enable CNV analysis using MassARRAY

platform: SNP Allele Ratio (SAR), Absolute Copy Number (ACN) and Allele Specific Copy Number (ASCN) (Figure 8) [45, 46].

- SAR: This method requires well-characterized SNPs for the CNV region. As the name implies, it provides relative copy number information (Figure 8.A). Distinguishing alleles are amplified in target amplicons, primer extension reactions are performed, resulting peaks for different alleles are analyzed with respect to their area under the peak.

- ACN: This approach does not rely on an SNP, can therefore make use of either a heterozygous or a homozygous site. It returns the absolute copy number in a highly quantitative manner (Figure 8.B). To enable the absolute copy number determination, a competitor template whose concentration is predetermined through serial titrations is used. Since the copy number of the control in the reaction is known, sample copy number can be calculated by normalizing the data.

- ASCN: This method combines the first two approaches and delivers highly informative data (Figure 8.C). Besides giving absolute copy number information, SNP allele ratios are compared. This way, wrong interpretations in inheritance pattern analyses can be prevented.

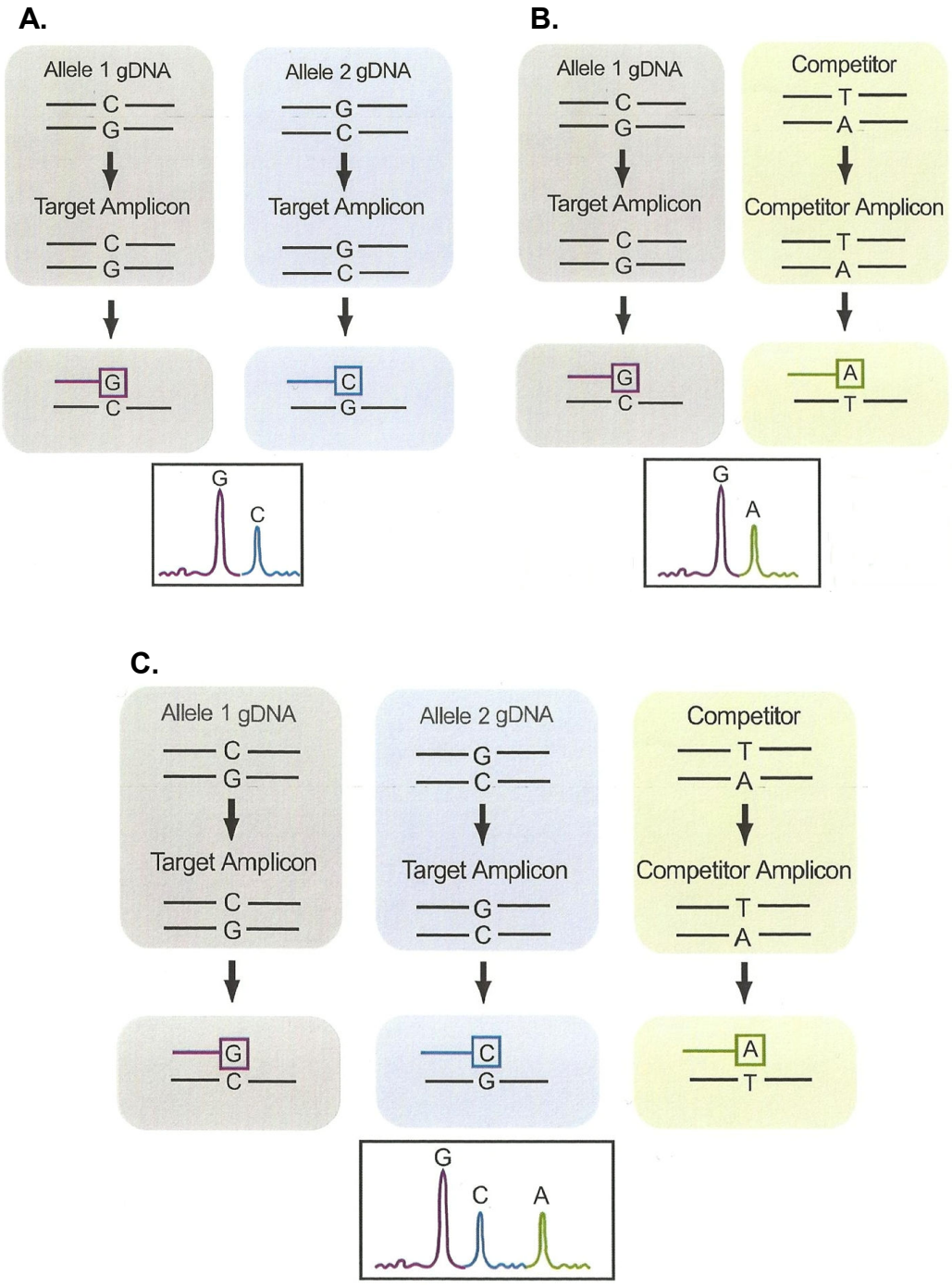


Figure 8. Schematic representation of three CNV analysis approaches on MassARRAY platform introduced by Sequenom: **A.** SNP Allele Ratio (SAR), **B.** Absolute Copy Number (ACN), **C.** Allele Specific Copy Number (ASCN).

2.3 Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight MS

2.3.1 History

MALDI technique was first introduced in 1988 to analyze protein molecules with a molecular mass greater than 10000 Da using mass spectrometry. Karas and Hillenkamp obtained spectra of lysozome (from chicken egg white), β-lactoglobulin A (from bovine milk), porcine trypsin and albumin (from bovine) successfully with low sample consumption (10^{-11} mol), short processing time and remarkable sensitivity [47]. Similar successful analyses were reported independently [48]. These authors combined MALDI with Time-Of-Flight mass spectrometers. By embedding the analyte in a matrix, irradiating the formed complex with an intense, pulsed laser beam, they were able to generate gas-phase ions of the analyte molecules, whose mass-to-charge ratios they detected by means of a detector at the end of the vacuumed flight path.

Since then, MALDI-TOF MS has been used in the analysis of nucleic acids. It was even considered a high-throughput alternative to Sanger sequencing; however, insufficient signal intensities in large sizes of DNA, adduct formation, multiple charging were major limitations to sequencing by MALDI-TOF MS. These difficulties limited the sequencing read length to <100 nucleotides [49, 50, 51]. In the first years of invention, negative ion mode was used to generate analyte ions; however, with the introduction of a more suitable matrix in 1993, 3-hydroxypicolinic acid (3-HPA) [52], than the matrices already in use, positive ion mode started to gain more attention, which has finally become the matrix of choice today [53].

Addition of di- and triammonium salts of organic or inorganic acids during matrix preparation minimized the adverse effect of sodium and potassium adducts [54]. To decrease the fragmentation of the phosphodiester backbone in large DNA sizes and obtain more stable matrix-analyte complexes, *in vitro* transcription into RNA was reported to be a considerable development [55]. Vestal applied the theory of time-lag energy focusing originally developed by Wiley and McLaren to MALDI-TOF MS and obtained significant improvement in mass resolution by delaying the extraction of

ions from the source. This improved the spectrum quality extensively by suppressing the matrix background, reducing chemical noise and minimizing the effect of laser intensity on performance [56].

Despite the developments achieved, the problem of size-dependent fragmentation and thus limited size range of detection, adduct formation and resolution still prevented MALDI-TOF MS from being a true alternative to DNA sequencing method. This made scientists start using MALDI-TOF MS for SNP analysis or genotyping from the second half of 90s. It proved to serve for this goal efficiently since it was highly accurate, fast, reproducible and most important of all, generated analytes with masses between 4500-8500 Da which were easy to detect. SNP identification of a polymorphic site in exon 13 of BRCA1 was accomplished using PCR products without purification or strand separation [57]. Soon after the same research group announced the multiplex genotyping of five SNP sites in the same PCR amplicon using MassTags [58]. Since the A to T change with a 9 Da difference was not resolved good enough in mid to high mass ranges, dideoxynucleotides were used together with extendable deoxynucleotides, which separated different alleles further apart from each other and provided a safer diagnostic product [59].

Sequenom initiated a campaign using this technology, which broadened the use of MALDI-TOF MS technology in the field of SNP genotyping. Using post-PCR *in vitro* transcription, the amount of final analytes for analysis was increased due to additional amplification during transcription and the stability of ribonucleotides was improved. The products were analyzed on MALDI-TOF MS for SNP discovery and pathogen identification after RNase T1 mediated base-specific cleavage (MassCLEAVE, hMC) [60]. The classical primer extension principle was introduced first under the name PROBE and then MassEXTEND [61], which was soon improved by using acyclic mass-modified base terminators for higher levels of multiplexing and named iPLEX. Today, hME is considered to be the cost-effective solution for plexes between 2 and 15, whereas iPLEX offers a lower cost per genotype for plexes higher than 25 [62].

2.3.2 Applications

Many genes were investigated successfully by researchers using MALDI-TOF combined with mass spectrometry. Analyses on MALDI-TOF MS were made for concurrent analysis of 40 SNPs in CYP2C9 and 50 SNPs in CYP2A13 by solid-phase capture-single-base extension by Misra et al. [63], for detection of *E2*, *E3* and *E4* alleles of *ApoE* using homogenous MassEXTEND by Ghebranious et al. [64], for identification of DNA sequence changes that occurred in *Escherichia coli* K-12 MG1655 during laboratory adaptive evolution to new optimal growth phenotypes by Honisch et al. [65], for simultaneous genotyping of Indels and SNPs in *ApoE* gene using $LuCl_3$ molecular scissors and oligonucleotides with RNA-activators by Sasayama et al. [66], for evaluating 15 relevant SNPs of CYP2B6 using primer extension approach by Blievernicht et al. [67].

3 Materials & Methods

3.1 Patients

All genomic DNA aliquots used within the frame of this work belong to patients who agreed to have the genetic test for CYP21-OHase deficiency offered by Bioglobe GmbH via the doctor's referral.

3.2 Chemicals

General

- Ethanol; 99% (AppliChem)
- HPLC water (AppliChem)

DNA extraction

- Proteinase K (Qiagen)
- QIAmp DNA Mini Kit (QIAGEN)
  - Buffer AL
  - Buffer AW1
  - Buffer AW2
  - 2 ml collection tubes

Enzymatic reactions

- PCR Hi-Fi Buffer; 10x (Invitrogen)
- 10x HotStar PCR buffer (Qiagen)
- SAP buffer (Sequenom)
- 5x T7 Polymerase buffer (Sequenom)
- 10x iPLEX buffer (Sequenom)
- dNTP Mix; 25 mM (Qiagen)
- T/C cleavage mix (Sequenom)
- iPLEX termination mix (Sequenom)

- $MgSO_4$; 50 mM (Invitrogen)
- $MgCl_2$ (25mM) (Qiagen)
- DMSO (Merck)
- Platinum DNA Taq Polymerase High-Fidelity (Invitrogen)
- HotStar Taq DNA Polymerase (Qiagen)
- Shrimp Alkaline Phosphatase (Sequenom)
- T7/SP6 DNA Polymerase (Sequenom)
- RNase A (Sequenom)
- Thermosequenase (Sequenom)
- iPLEX enzyme (Sequenom)
- Bovine Serum Albumin – BSA (BioLabs)

Gel electrophoresis

- Biozym LE Agarose (Biozyme)
- TBE Buffer (10 mM Tris-HCl, 10 mM boric acid, pH 7,5, 1 mM EDTA)
- Tris-hydroxymethyl-aminmethane (Merck)
- EDTA (Merck)
- Boric acid (Merck)
- LE Agarose (Biozyme)
- Ethidiumbromide; 10 mg/ml (Roth)
- MassRuler™ DNA Ladder; Mix, 80 – 10000 bp (Fermentas)
- FastRuler™ DNA Ladder; High Range, 500 – 10000 bp (Fermentas)
- DNA Loading buffer; 6x (5 ml Glycerin, 5 ml TE-Buffer, 0,25% BPB, 100 µl 0,5 M EDTA, pH 8,0) (AppliChem)

3.3 Devices

- Pipettes; 0,5-10/2-20/10-100/50-200/100-1000 µl (Eppendorf)
- Multi-channel Pipettes; 0,5-10/5-100/50-1250 µl (Eppendorf)
- Centrifuges; Rotanta 460R, Rotina 35 (Hettich), Centrifuges 5415C and 5415D (Eppendorf)
- Heating block; Thermomix 5436 (Eppendorf)

- Vortex (Scientific Industries)
- PeqLab Advanced Primus 96 cyclers (384-well block)
- Gel electrophoresis chamber (MWG-Biotech)
- Gel electrophoresis power supply (Biometra)
- Microwave oven (Siemens)
- Balance (Sartorius)
- Dark Hood DH-40/50 Bio-imaging (Biostep)
- MassARRAY Nanodispenser (Samsung)
- MALDI-TOF Mass Spectrometry Compact Platform (Bruker)

3.4 Plastics

- 384-Well Microtiter PCR Plates (Sarstedt)
- 384-Well Dimple Plate (3 mg, 6 mg) (Sequenom)
- Pipette tips; 0,5-20/10-200/100-1000 µl (Sarstedt)
- Reaction tubes; 0,5/1,5/2,0 ml (Sarstedt)
- Plastic films (Sarstedt)

3.5 Long-range Amplification of CYP21A2

Due to its relatively large size and high similarity to CYP21A1P, the common practice has usually been to amplify CYP21A2 in two overlapping fragments. This way, specificity is assured with either forward or reverse primer. For the first fragment, depending on the size of the region which is desired to be analyzed prior to exon 1, the forward primer binds in the promoter region and the reverse primer in exon 6 cluster (E6-cluster), where three thymine bases in CYP21A2 are located in close vicinity and three adenine bases instead of them are present in CYP21A1P. For the second fragment, the forward primer binds on the eight base pairs (8-bp), which are deleted in CYP21A1P, and the reverse primer binds in the 3'-UTR region universally. It is the reverse and the forward primer in the amplification of the first and the second fragment, respectively, which assures specificity.

In rare cases 8-bp is known to be present in CYP21A1P. More importantly, in hybrid genes, which result after deletions and large conversions, E6-cluster is present in either heterozygous or homozygous form. This was proven in a previous study, which included 200 individuals where all subjects were analyzed for their CYP21A2 and CYP21A1P genes separately [68]. Results show that amplification of CYP21A2 in two overlapping fragments is prone to deliver false and/or incomplete genotype results. Therefore, the long-range amplification of CYP21A2 in one fragment was adopted in this work.

In this approach, the forward primer, which binds in the promoter, enables specific amplification of CYP21A2. This 22-bp-long primer binding site has two nucleotides which differentiate between the two CYP21 genes at all times. One of them is the outermost 3' base and the other is the 12th base from the 3'-end. Two reverse primers are used in combination with the forward primer. These primers are designed within the scope of this work since the 3'-end of the gene is highly variable and contains no constant distinguishing regions between the two genes. The last five bases of each primer possess thio-phosphate bonds to increase specificity by preventing any miscorrection of the primer sequence by proof-reading activity of the applied DNA polymerase from annealing it to CYP21A1P (Figure 9).

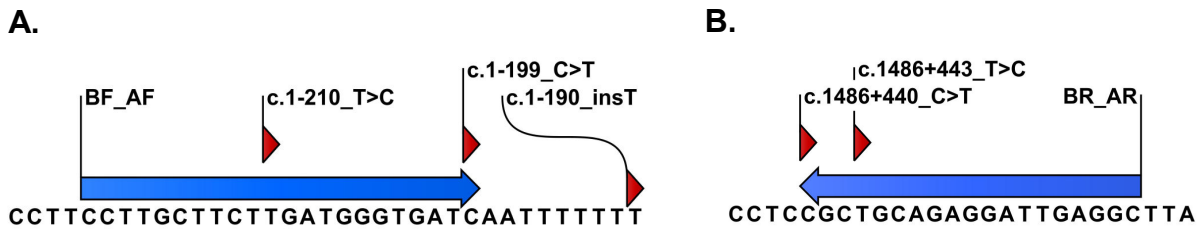


Figure 9. Schematic representation of the primer binding sites for long-range amplification of CYP21A2: Distinguishing bases are indicated on CYP21A2 reference sequence together with their positions. Corresponding genotypes in CYP21A1P are shown in the annotation. **A.** Specific forward primer for CYP21A2 and CYP21A1P, BF and AF, respectively. **B.** Target sequence of reverse primers, BR and AR.

Using BF in combination with BR and AR enabled specific amplification of CYP21A2 provided that 5'-end region is not deleted. Using AF in combination with BF and AR enabled the amplification of possible hybrid genes. After having optimized the PCR

reagent concentrations and cycling conditions, products of expected size were obtained (Table 1).

Table 1. **A.** PCR cocktail recipe designed and optimized for the long-range amplification of CYP21A2. A final volume of 15 µl was used. Final concentrations are given. **B.** Designed long-range CYP21-PCR program.

A.

Buffer for Hi-Fi Taq 1x	
dNTPs	0.2 mM
BF	0.8 µM
BR	0.5 µM
AR	0.5 µM
$MgSO_4$	2 mM
DMSO	v/v 1%
BSA	0.8 µg/µl
DNA Polymerase Hi-Fi	0.3 U
Genomic DNA	80-130 ng
dd H_2O	NA

B.

94ºC	2 min	
94ºC	35 sec	35x
62ºC	45 sec	
68ºC	4 min	
68ºC	10 min	
4ºC	∞	

3.6 Nested PCR

To enable CYP21A2 analysis with mass spectrometry, the long-range PCR product has to be re-amplified in shorter sub-amplicons. The 3 kb amplicon is divided into nine sub-amplicons, whose lengths range from 284 bp to 475 bp, using primers which are tagged with T7 and SP6 RNA polymerase promoters (Table 2). A much shorter amplicons with a size of 141 bp and named Amp356, had to be designed additionally for the analysis of a vital mutation which will be explained in the following sections in more detail.

Nested PCR functions as the first step of the hMC protocol in terms of reagents and cycling conditions, except that the total volume in one well is 7.5 µl instead of the standard 5.0 µl due to the experiments that follow. Detailed conditions are described in Table 3.

Table 2. Primers designed and optimized for the nested PCR for CYP21A2 after long-range PCR: Forward primers are tagged with T7-tag and reverse primers with SP6-tag. T7-tag= cagtaatacgactcactatagggagaaggct, SP6-tag= cgatttaggtgacactatagaagagaggct, I: Universal binding base inosine. Position numbers are assigned relative to the ENSEMBL transcript ENST0000448314 (a merged manually-curated transcript from Havana/Vega matching to the genebuild ENSG00000198457).

Name		Sequence (5' → 3')	Position
Ex1F	Amp1	T7-GGGATGGCTGGGGCTCTTG	c.1-60 - c.1-42
Ex1R	327 bp	SP6-GAGGACCCTCTCCGTCACC	c.200-33 - c.199+47
Ex2F	Amp2	T7-GCTGCAAGGTGAGAGGCTGAT	c.192 - c.199+13
Ex2R	318 bp	SP6-CTTGAGGCTGAGGTGGGAG	c.289+124 - c.289+106
Ex3F	Amp3	T7-GCCCAGGCTGGTCTTAAATTC	c.289+69 c.289+89
Ex3R	381 bp	SP6-AGCCCAGCCTTACCTCAC	c.444+13 - c.440
Ex4F	Amp4	T7-AAGCCCACAAGAAGCTCACC	c.350 - c.369
Ex4R	361 bp	SP6-CAGGACAAGGAGAGGCTCAG	c.547-31 - c.546+39
Ex5/6F	Amp5/6	T7-GATCAAGGTGCCTCACAGCC	c.540 - c.546+13
Ex5/6R	451 bp	SP6-GCAATGCTGAGGCCGGTAGC	c.735+67 - c.735+48
Ex7F	Amp7	T7-AGGCAGCACAAGGTGGGGAC	c.724 - c.735+8
Ex7R	426 bp	SP6-GCCAGGTTGCTGGGAAGGAG	c.936+45 - c.936+26
Ex8F	Amp8	T7-TTTTTTTGCTTCACCACCCTG	c.914 - c.934
Ex8R	475 bp	SP6-GCTGGAGTTAGAGGCTGGC	c.1116-10 - c.1116-28
Ex9F	Amp9	T7-CACCACACGGCCCAGCAGG	c.1098 - c.1115+1
Ex9R	284 bp	SP6-GGTGGGTGGGGAGGCGTTC	c.1120-18 - c.1120-36
Ex10F	Amp10	T7-CCTGCCGTGAAAATGTGGTGG	c.1219+20 - c.1219+40
Ex10R	443 bp	SP6-GCGATCTCGCAGCACTGTGT	c.1486+100 - c.1486+81
356F	Amp356	T7-CCIAGGTGCTGIICCTG	c.1049 - c.1065
Ex8R	141 bp	SP6-GCTGGAGTTAGAGGCTGGC	c.1116-10 - c.1116-28

Table 3. **A.** Recipe of nested PCR designed for CYP21A2. A final volume of 7.5 µl was used. Final concentrations are given. **B.** Designed nested PCR (hMC-PCR) program.

A.

HotStar PCR buffer	1x
dNTPs	0.2 mM
PCR primer forward	0.2 µM
PCR primer reverse	0.2 µM
HotStar Taq	0.15 U
Long-range PCR product	NA (1 µl)
dd H_2O	NA

B.

94ºC	15 min	
94ºC	20 sec	45x
62ºC	30 sec	
72ºC	1 min	
72ºC	3 min	
4ºC	5 min	
15 ºC	∞	

3.7 Shrimp Alkaline Phosphatase (SAP) Digestion

Crude PCR products are treated with Shrimp Alkaline Phosphatase (SAP) to dephosphorylate any remaining triphosphates and prevent their incorporation in the following experiments (Table 4). 3 µl of SAP mix is added directly to each reaction well.

Table 4. **A.** Recipe of SAP digestion cocktail applied after nested PCR. Final concentrations are given. **B.** SAP treatment program.

A.	SAP buffer	1x	**B.**	37°C	20 min
	dd H_2O	NA		85°C	5 min
	SAP enzyme	0.6U		15°C	∞

3.8 Mass Spectrometry Methods – MassCLEAVE and iPLEX

Commercial kits from Sequenom GmbH, MassCLEAVE and iPLEX, were used during the project. The working principle of both methods is explained below in detail and outlined in Figure 10 [69, 70].

Base-specific cleavage is the basic biochemical method for SNP discovery using mass spectrometry. Reference sequences for amplicons are loaded into the software database, experiments are run as described in the manufacturer's manuals and signals are obtained for all four reactions; T/C forward and T/C reverse. This procedure results in fragments from an amplicon, which are truncated at every base. Generated signals are retrieved from SNP discovery workstation and analyzed. Depending on the settings, detected masses are listed together with new and missing masses, intensities and sequences of fragments from which they are generated. Once this information is at hand, the operator can locate the position of the base change, deletion or insertion easily.

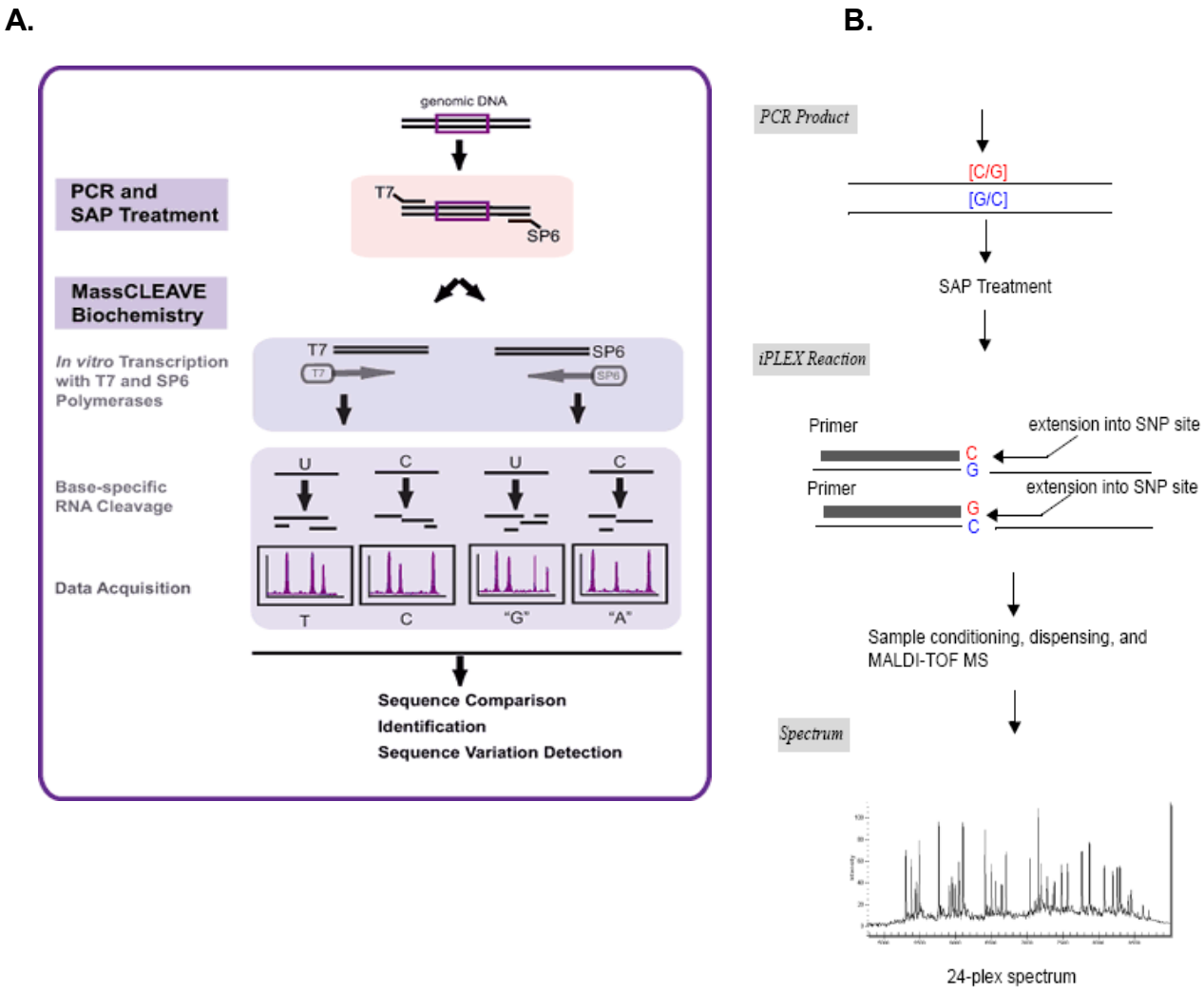


Figure 10. **A.** Schematic representation of MassCLEAVE reactions. **B.** Schematic representation of iPLEX reactions.

iPLEX is the second mass spectrometry method which is used for genotyping. This approach has a major difference to MassCLEAVE; that is, the position where an SNP is expected must be known in advance. It is not possible to discover any variants located in a position different than the SNP site itself. Unlike MassCLEAVE, the method does not include a transcription step. It exploits the change of mass of the fragment after terminating the primer elongation after one single specific base.

3.8.1 Transcription and Base-specific Cleavage

Once the SAP digestion step is completed, the final step before the measurement is the transcription into RNA and cleavage (Table 5). Base-specific cleavage takes place after transcription into RNA by incubation at 37°C for 3 hours.

Table 5. **A.** Recipe of hMC-transcription cocktail. Final concentrations are given. **B.** hMC-transcription and cleavage program.

A.	5x T7 Polymerase buffer	1x	**B.**	37°C	180 min
	T/C cleavage mix	NA		15°C	∞
	DTT	5mM			
	T7/SP6 DNA Polymerase	20U			
	RNase A	0.08mg/ml			
	SAP treated nested PCR product	1 µl			
	dd H_2O	NA			

3.8.2 Single Base-extension Reactions

Mass extension reactions are performed using the amplicons created already in the nested PCR step. Both homozygous mass extend (hME) and iPLEX reactions were tested during the evaluation period (Table 6, Table 7).

Table 6. **A.** Recipe of MassEXTEND reaction cocktail adjusted to iPLEX (buffer-enzyme pair) specifically for CYP21A2 analysis. Final concentrations are given. **B.** MassEXTEND reaction program.

A.	hME buffer	0.22x	**B.**	94ºC	2 min	
	iPLEX termination mix	0.5x		94ºC	5 sec	75x
	Primer mix (10µM-20µM)	1.3µM-2.7µM		52ºC	5 sec	
	hME enzyme	0.5x		72 ºC	5 sec	
	dd H_2O	NA		15ºC	∞	

Table 7. **A**. Recipe of iPLEX-Gold reaction cocktail for low multiplexing. Final concentrations are given. **B.** iPLEX reaction program.

A.			**B.**			
	iPLEX buffer	0.22x		94ºC	30 sec	
	iPLEX termination mix	0.5x		94ºC	5 sec	55x
	Primer mix (10µM-20µM)	1.3µM-2.7µM		52ºC	5 sec	5x
	iPLEX enzyme	0.5x		80ºC	5 sec	
	SAP treated product	6 µl		72ºC	3 min	
	dd H_2O	NA		15 ºC	∞	

A total number of 31 variations, 17.9% of all, is analyzed successfully by iPLEX approach in four multiplex groups. These multiplex groups are named plex 1, 2, 3 and Poly since the last group contains polymorphisms only. The first and the second multiplex groups are 9-plex, the third is 6-plex and Poly is 7-plex (Table 8). Detecting polymorphisms is not as vital as detecting mutations as they do not alter the phenotype; however, this multiplex group serves as a proper internal quality control to assure no allele drop-out has taken place during amplification. If the genotyping data do not correlate with clinical manifestations and all polymorphism positions result in homozygous state, the experimenter might suspect a possible allele drop-out and perform further inspection.

Table 8. Recipes of extension primer mixes for each multiplex group consisting of specifically designed primers: Primer mixes are prepared in a volume of 30 µl. Final concentrations of extension primers in the mixes and their sequences are shown under assay groups 1, 2, 3, and Poly. The adopted nomenclature is given in parenthesis, but the assay names used during the test phase are kept. **i**: Universal binding base Inosine.

Assay name	**Sequence**	**Concentration**
1.		
8bp (c.329_336delGAGACTAC)	TTGTGGGCTTTCCAGAGCAG	10µM
E351K (c.1051G>A)	TTGCTCAATGCCACCATCiCC	12µM
G291C (c.871G>T)	GCAGTGGACCTCCTGATC	12µM
P453S (p.Pro453Ser)	GGGCAGGGCGTCCCCGGAGG	12µM
R356P/Q (c.1067G>C/A)	CGCCAGCGGCTCGCCCAGGCACACG	10µM
R426C (c.1276C>T)	TCTGGCCTTCGGCTGCGGTiCC	15µM
R426Hlong (c.1277G>A)	CGCCAGCGGCTCGCCCAGGCACACG	15µM
V281G (c.842T>G)	GGAGGTCCACTGCAGCCAiGTGC	12µM
V281L(p.Val281Leu)	CTCCTGGAAGGGCAC	10µM
2.		
A362V (c.1085C>T)	GCCCGTTGTGCCCTTAG	10µM
E6 (c.707, c.710, c.716T>A)	AGGCCATAGAGAAGAGGGA	12µM
G178R (c.532G>A)	TGAGGCACCTTGATCTTGTCTC	12µM

Table 8. cont'd

I172N (p.Ile172Asn)	CGAAGGTGAGGTAACAG	10μM
L317M/V (c.949C>A/G)	CGTGGTCTAGCTCCTCCTiCA	12μM
Mt.Cd.241 (c.721C>G)	GTCCCCACCTTGTGCTGCCTCA	12μM
Mt.Cd.450 (c.1349C>T)	GTCCCCGGAGGiCAGCAGC	12μM
R483Q/P (c.1448G>A/C)	GTGGGCCCCCATCCCC	12μM
Mt.Cd.491 (c.1471G>A)	CCCATCACTGGiTCTGGC	10μM
3.		
1789_G/C (rs6442)	CCACCCTGAGiiGCGTCCTG	20μM
P105L (c.314C>T)	CTGGTGTCTAiGAACTACC	10μM
W302R (c.904T>C)	GACCCCAGCAAACACCiTCTCA	15μM
R339H (c.1016G>A)	GGGTCCCCTACAAGGACC	10μM
V304M (c.910G>A)	GGTGGTGAAiCAAAAAAAACCA	12μM
Q318X (p.Gln318X)	TGGTCTAGCTCCTCCT	10μM
Poly.		
-4_C/T (rs6470)	CCCTGACGGGCGTCT	10μM
P.Cd.493 (rs61732563)	TCCTGCCCCATCACTGG	10μM
P.560 (c.290-109G>C)	GTCiCTGGAAGCTCTTG	12μM
P.601 (rs6451)	GAGGGCCTGACCTTCTTC	10μM
delCTG (rs28381641)	CAGGGGCAGCAGCAGCAG	10μM
P.629-30 (rs35147842)	GCTGGAGGGTGGGAACTGA	12μM
P.490 (rs61732562)	GGGATGGGGGCCCACAGCCC	12μM

The protocols which have been explained so far are related to the mutation detection section of the project. The CNV analysis is performed using the same protocols only with the exception of the absence of a gene specific long-range PCR as the beginning step. In CNV analysis, a 3-plex multiplex PCR replaces this step, by which the necessary portions of CYP21A2 and CYP21A1P are amplified simultaneously. This procedure is explained in more detail in section 3.2. Multiplex PCR recipe and the cycling program used in quantitative analysis are summarized in Table 9.

Table 9. **A.** Recipe of multiplex PCR designed for SAR analysis of CYP21A2 and CYP21A1P: A total standard volume of 5.0 µl was used. Final concentrations are given. **B.** Multiplex PCR program designed specifically for quantitative analysis of CYP21A2 and CYP21A1P. **C.** Extension primer mix designed for the quantitative analysis.

A.	Buffer for Hi-Fi Taq	1.25x	**B.**	94ºC	15 min	
	dNTPs	200µM each		94ºC	20 sec	25x
	PCR primer mix	0.1µM of each primer		62ºC	30 sec	
	$MgSO_4$	3.5mM		68ºC	1 min	
	Hi-Fi Taq	0.5U		68ºC	3 min	
	Genomic DNA	20-45 ng		4ºC	5 min	
	dd H_2O	NA		15 ºC	∞	
C.	c.1-126C/T	18µM				
	c.289+138A/G	15µM				
	c.702T/C	10µM				
	H_2O	NA				

3.9 Sample Conditioning

After mass extension and cleavage reactions are completed, reaction products are diluted with 20 µl deionized H_2O. Remaining metal ions are removed using Clean Resin, which is included in the reaction kits from Sequenom. hMC and iPLEX reaction products are treated with 6 mg Clean Resin per well, hME reaction products with 3 mg. The appropriate amount of Clean Resin is transferred to the well by the help of a dimple plate. The reaction plate is rotated at room temperature for approximately 15 min.

3.10 Sample Dispensing and Measurement

The resin is spun down at maximum centrifugation speed for 5 min. Approximately 15 nl of the reaction product is dispensed on a SpectroCHIP using the MassARRAY Nanodispenser. Measurements are performed and spectra are acquired using the default settings for hMC and iPLEX on the MassARRAY Analyzer Compact.

3.11 Target Mutation List

Due to technical principles of mass spectrometry methods, it is impossible to acquire a direct gene sequencing-like result. Variations that result in fragments after base-specific cleavage which do not generate differentiating signals between the mutant and the wild type in the detectable mass range are not possible to identify. As for SNP-site extension reactions, the researcher must know beforehand where the variation is expected to occur in order to design an extension primer adjacently. Therefore, a mutation set had to be defined for variations detectable by the protocol.

The approach was to list all annotated mutations and polymorphisms up-to-date together with their positions, the type of CAH they cause and *in vitro* enzyme activity they result in provided that the data was available. This information was obtained from the online database of The Human Cytochrome P450 Allele Nomenclature Committee for CYP21A2 and from NCBI's SNP database (dbSNP, mapping to genome build 37.2) for CYP21A2. Polymorphisms and rare mutations found in routine analyses in Bioglobe GmbH laboratories were included as well (Table 10) [71, 72].

Due to the fact that this region of human genome is highly rich in variations, there is not enough number of sites for extension primers to anneal for every single submission in the database. This brings one to the obligation to build a target list. The priority set in doing so was as follows: the ten most important pseudogene-derived mutations, rare mutations whose effect on enzyme activity is known, rare mutations found at Bioglobe GmbH laboratories throughout routine analyses until present date, and finally the most common polymorphisms found in German population. The variation with the ID 114 in indicates the end of the target list regarding rare mutations. The detectable variants after this point were still included in the study; however, no primer extension assays were designed for them if they were undetectable by hMC.

3.12 Software

To obtain wild type (reference) and mutation/SNP masses of fragments after cleaving the amplicons of nested PCR in all four reactions, MS-Analyzer is used for simulation purposes, which is an in-house developed prototype tool. With the aid of the MS-Analyzer tool, it was determined *in silico* whether or not a certain variation results in a differentiating signal by cleaving amplicons.

MassARRAY AssayDesign 3.1 is the basic tool to design assays for iPLEX. The software has the following default settings for single base extension assay design: Primer length between 17 bp and 28 bp, mass range between 4500 Da and 8500 Da, minimum peak separation for analytes and extension primers 20 Da and 10 Da, respectively, analyte by-product masses ± 30 Da. The software was used with default settings except the cases where certain assays needed to be manually designed when the software reported a design failure due to secondary SNPs in close vicinity or any other risk factor such as hair-pin formation, etc.

Table 10. **1-10**: The ten most frequent pseudogene-derived mutations. **11-59**: Polymorphisms submitted to dbSNP of NCBI and found in Bioglobe GmbH laboratories. **60-173**: Rare mutations submitted to the database of The Human Cytochrome P450 Allele Nomenclature Committee and found in Bioglobe GmbH laboratories. **>**: Nucleotide change, **del**: Deletion, **ins**: Insertion, **SW**: Salt wasting, **NC**: Non-classical, **SV**: Simple virilization, **N**: Normal, **hMC**: MassCLEAVE, **Mut**: Mutation, **Poly**: Polymorphism, **N.A.**: Not analyzed. Amino acids are shown in one letter code. Polymorphisms are shown with their rs-numbers, corresponding positions and frequencies if provided by the submitter. Numbering is done relative to ENSEMBL transcript ENST0000448314 (matching to the genebuild ENSG00000198457). [a]Among 115 individuals analyzed at Bioglobe GmbH. Frequencies without the symbol "[a]" are taken from the frequency data on NCBI's webpage as long as it was provided by the submitter.

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
1	p.Pro30Leu	Mut	c.89C>T	Exon 1	NC	60/30	-	[73]
2	IVS2AS,A/C-G,-13	Mut	c.290-13A/C>G	Intron 2	SW	Weak activity	-	[74] [75]
3	8bp-del	Mut	c.329_336delGAGACTAC	Exon 3	SW	-	-	[76]
4	p.Ile172Asn	Mut	c.515T>A	Exon 4	SV	<2	-	[77] [78]
5	E6*	Mut	c.707, c.710, c.716T>A	Exon 6	SW	1/2,4 – 0/0,1 – 95/98	-	[76] [79]
6	p.Val281Leu	Mut	c.841G>T	Exon 7	NC	50/20	-	[78] [80]
7	p.Leu306PhefsX5	Mut	c.920dupT	Exon 7	SW	-	-	[76]
8	p.Gln318X	Mut	c.952C>T	Exon 8	SW	-	-	[81]
9	p.Arg356Trp	Mut	c.1066C>T	Exon 8	SW	0/0	-	[82]
10	p.Pro453Ser	Mut	c.1357C>T	Exon 10	NC	50-68/20-46	-	[83] [84] [85]
11	rs6470	Poly	c.1-4C>T	5'-UTR	-	-	0.439	-
12	rs28381641	Poly	c.27insCTG	Exon 1	-	-	-	-
13	rs6468	Poly	c.115T>C	Exon 1	-	-	0.450	-
14	rs6464	Poly	c.135A>C	Exon 1	-	-	0.353	-
15	rs6462	Poly	c.289+9T>C	Intron 2	-	-	0.439	-
16	rs6448	Poly	c.289+15C>A	Intron 2	-	-	0.180	-
17	rs6463	Poly	c.289+33C>A	Intron 2	-	-	0.439	-
18	rs6449	Poly	c.289+67T>C	Intron 2	-	-	0.439	-

Table 10. cont'd

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
19	rs11757034	Poly	c.289+134C>T	Intron 2	-	-	0.024	-
20	c.290-109G>C	Poly	c.290-109G>C	Intron 2	-	-	0.152[a]	-
21	rs41315224	Poly	c.209-105delG	Intron 2	-	-	-	-
22	rs6450	Poly	c.209-74G>A	Intron 2	-	-	0.180	-
23	rs6451	Poly	c.290-67C>A/G	Intron 2	-	-	0.439	-
24	rs59064806	Poly	c.209-48A>G	Intron 2	-	-	-	-
25	rs6453	Poly	c.290-44G>T	Intron 2	-	-	0.049	-
26	rs35147842	Poly	c.290-38_39CA>GG	Intron 2	-	-	-	-
27	rs6467	Poly	c.290-13A>C	Intron 2	-	-	0.439	-
28	rs6454	Poly	c.290-12C>T	Intron 2	-	-	0.049	-
29	c.290-4G>A	Poly	c.290-4G>A	Intron 2	-	-	0.004	-
30	rs6474	Poly	c.305A>G	Exon 3	-	-	0.378	-
31	c.315G>C	Poly	c.315G>C	Exon3	-	-	-	-
32	rs6456	Poly	c.321G>T	Exon 3	-	-	0.135	-
33	rs6466	Poly	c.444+38C>T	Exon 3	-	-	-	-
34	rs58693631	Poly	c.444+39G>A	Intron 3	-	-	0.009	-
35	rs72552751	Poly	c.533G>C	Exon 4	-	-	-	-
36	rs1040312	Poly	c.547-15C>A	Intron 4	-	-	0.029	-
37	rs1040311	Poly	c.547-8T>C	Intron 4	-	-	-	-
38	rs1040310	Poly	c.549C>G	Exon 5	-	-	-	-
39	c.648+30G>A	Poly	c.648+30G>A	Intron 5	-	-	0.03[a]	-
40	rs12525076	Poly	c.648+35A>G	Intron 5	-	-	0.486	-
41	rs6457	Poly	c.676C>T	Exon 6	-	-	-	-
42	rs10947229	Poly	c.702T>C	Exon 6	-	-	-	-

Table 10. cont'd

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
43	c.708C>G	Poly	c.708C>G	Exon6	-	-	-	-
44	rs71552100	Poly	c.735+12_13AC>GT	Intron 6	-	-	-	-
45	c.736-74G>A	Poly	c.736-74G>A	Intron 6	-	-	0.035[a]	-
46	rs6465	Poly	c.736-21C>T	Intron 6	-	-	0.301	-
47	rs6477	Poly	c.744C>G	Exon 6	-	-	0.439	-
48	rs6472	Poly	c.803G>C	Exon 6	-	-	0.196	-
49	rs55695018	Poly	c.819T>C	Exon 7	-	-	-	-
50	rs72552753	Poly	c.936+2T>G	Intron 7	-	-	-	-
51	rs6442	Poly	c.936+11G>C	Exon7	-	-	0.049	-
52	rs6461	Poly	c.1116-34G>A	Intron 8	-	-	-	-
53	rs6469	Poly	c.1122C>T	Exon 9	-	-	0.309	-
54	rs2242571	Poly	c.1219+26G>A	Intron 9	-	-	0.013[a]	-
55	rs61732562	Poly	c.1470A>G	Exon 10	-	-	-	-
56	rs61732563	Poly	c.1478A>G	Exon 10	-	-	-	-
57	rs6447	Poly	c.1486+13G>A	3'-UTR	-	-	0.041	-
58	rs56148109	Poly	c.1486+222C>A	3'-UTR	-	-	-	-
59	rs415620	Poly	c.1486+368T>C	3'-UTR	-	-	0.5	-
60	c.289+45insTGT	Mut	c.289+45insTGT	Intron2	?	?	-	Bioglobe GmbH
61	c.290-129insTCC	Mut	c.290-129insTCC	Intron2	?	?	-	Bioglobe GmbH
62	c.290-85_G>A	Mut	c.290-85_G>A	Intron2	?	?	-	Bioglobe GmbH
63	c.290-84_G>T	Mut	c.290-84_G>T	Intron2	?	?	-	Bioglobe GmbH
64	c.324C>G	Mut	c.324C>G	Exon3	?	?	-	Bioglobe GmbH

Table 10. cont'd

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
65	c.339C>T	Mut	c.339C>T	Exon3	?	?	-	Bioglobe GmbH
66	Mt.Cd.124	Mut	c.370C>T	Exon 3	?	?	-	Bioglobe GmbH
67	Mt.Cd.198	Mut	c.594A>T	Exon 5	?	?	-	Bioglobe GmbH
68	Mt.Cd.241	Mut	c.721C>G	Exon6	?	?	-	Bioglobe GmbH
69	c.936+1G>C	Mut	c.936+1G>C	Intron7	?	?	-	Bioglobe GmbH
70	c.936+2T>G	Mut	c.936+2T>G	Intron7	?	?	-	Bioglobe GmbH
71	c.936+12G>A	Mut	c.936+12G>A	Intron7	?	?	-	Bioglobe GmbH
72	Mt. Cd. 350	Mut	c.1048G>A	Exon8	?	?	-	Bioglobe GmbH
73	Mt. Cd. 450	Mut	c.1349C>T	Exon10	?	?	-	Bioglobe GmbH
74	Mt. Cd. 459	Mut	c.1375C>T	Exon 10	?	?	-	Bioglobe GmbH
75	Mt. Cd. 463	Mut	c.1388C>T	Exon 10	?	?	-	Bioglobe GmbH
76	Mt. Cd. 484	Mut	c.1452delG	Exon 10	?	?	-	Bioglobe GmbH
77	Mt. Cd. 491	Mut	c.1471G>A	Exon 10	?	?	-	Bioglobe GmbH
78	c.1486+12C>T	Mut	c.1486+12C>T	3'-UTR	?	?	-	Bioglobe GmbH
79	R354H	Mut	c.1061G>A	Exon 8	SW	0	-	[86] [87]
80	G64E	Mut	c.191G>A	Exon 1	SW	0	-	[88]
81	G90V	Mut	c.269G>T	Exon 2	SW	0	-	[86] [87]
82	A362V	Mut	c.1085C>T	Exon 8	SW	0	-	[88]
83	G291C	Mut	c.871G>T	Exon 7	SW	0	-	[86] [87]
84	R426C	Mut	c.1276C>T	Exon 10	-	0/0	-	[89]
85	C169R	Mut	c.505T>C	Exon 4	SV	0.1/0	-	[89]
86	W302R	Mut	c.904T>C	Exon 7	SW	0.1/0	-	[89]
87	R356P	Mut	c.1067G>C	Exon 8	SW	0.15/0.15	-	[90]

Table 10. cont'd

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
88	P30Q	Mut	c.89C>A	Exon 1	SW	0.2/0	-	[91]
89	L166P	Mut	c.497T>C	Exon 4	-	0.3/0.4	-	[92]
90	G178R	Mut	c.532G>A	Exon 4	-	0.4/0	-	[89]
91	L446P	Mut	c.1337T>C	Exon 10	SV/SW	0.5/0.0	-	[93]
92	R426H	Mut	c.1277G>A	Exon 10	SV/SW	0.5/0.4	-	[94] [93]
93	R356Q	Mut	c.1067G>A	Exon 8	SV	0.65/1.1		[90]
94	I171N	Mut	c.512T>A	Exon 4	-	0.7/0.6	-	[95] [93]
95	R341P	Mut	c.1022G>C	Exon 8	NC/SV	0.7/0.7	-	[96] [93]
96	G291S	Mut	c.871G>A	Exon 7	SW	0.8/0.8	-	[97] [98]
97	G375S	Mut	c.1123G>A	Exon 9	SW	<1/<1	-	[99]
98	R483P	Mut	c.1448G>C	Exon 10	-	1.0/2.2	-	[100] [98]
99	E351K	Mut	c.1051G>A	Exon 8	SV	1.1/1.2	-	[101]
100	R483Q	Mut	c.1448G>A	Exon 10	NC	1.1/3.8	-	[102] [103]
101	V281G	Mut	c.842T>G	Exon 7	SV	3.9/3.9	-	[104] [105]
102	I77T	Mut	c.230T>C	Exon 2	SV	3/5	-	[106]
103	E196del	Mut	c.587_589delAGG	?		6/23	-	[85]
104	L300F	Mut	c.898_C/T	Exon 7	SV	9.5/4.4	-	[107] [105]
105	P105L + P453S	Mut	c.314C>T; c.1357C>T	Exon 3 Exon 10	NC	10/7	-	[108] [85]
106	A434V	Mut	c.1303C>T	Exon 10	SV	14/12	-	[106]
107	G178A	Mut	c.532G>C	Exon 4	-	19/0	-	[86] [87]
108	E380D	Mut	c.1140G>C	Exon 9	SW?	30	-	[109] [110]
109	A391T	Mut	c.1171G>A	Exon 9	-	38/23	-	[103]
110	V304M	Mut	c.910G>A	Exon 7	NC	46/26	-	[99]

Table 10. cont'd

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
111	R339H	Mut	c.1016G>A	Exon 8	NC	54/46	-	[111]
112	P105L	Mut	c.314C>T	Exon 3	NC	62/64	-	[108]
113	P482S	Mut	c.1444C>T	Exon 10	NC	70/72	-	[95] [112]
114	R479L	Mut	c.1436G>T	Exon 10	?	76/80	-	[113] [103]
115	L317M/V	Mut	c.949C>A/G	Exon 8	NC/?	?	-	[114] [115]
116	R435C	Mut	c.1303C>T	Exon 10	NC	?	-	[114]
117	R483Frameshift	Mut	c.1448_49delGGinsC	Exon 10	SW	?	-	[116]
118	c.936+1G>C	Mut	c.936+1G>C	Intron 7	SW	?	-	[117]
119	W405X	Mut	c.1214G>A	Exon 9	SW	?	-	[117]
120	R341W	Mut	c.1021C>T	Exon 8	?	?	-	-
121	W22X	Mut	c.66G>A	Exon 1	SW	?	-	[118]
122	c.200-2A>G	Mut	c.200-2A>G	Intron 1	SW	?	-	[118]
123	W302X	Mut	c.906G>A	Exon 7	SW	?	-	[119]
124	Y97X	Mut	c.291C>A	Exon 3	SW	?	-	[120]
125	C169Frameshift	Mut	c.505_506, delTGinsA	Exon 4	SW	?	-	[121]
126	c.289+1G>A	Mut	c.289+1G>A	Intron 2	SW	?	-	[122]
127	R316X	Mut	c.946C>T	Exon 8	SW	?	-	[122]
128	S330Frameshift	Mut	c.988_997delTCCAGCTCCC	Exon 8	SW	?	-	[122]
129	V397Frameshift	Mut	c.1176_1191, dupCCTGGATGAGACGGTC	Exon 9	SW	?	-	[122]
130	1780T>G	Mut	c.936+2T>G	Intron 7	SW	?	-	[123]

Table 10. cont'd

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
131	P475Frameshift	Mut	c.1426delT	Exon 10	?	?	-	123
132	L48Frameshift	Mut	c.141delT	Exon 1	SW	?	-	[124]
133	Q262X	Mut	c.784C>T	Exon 7	SW	?	-	[88]
134	K74X	Mut	c.220A>T	Exon 2	SW	?	-	[87]
135	G424S	Mut	c.1273G>A	Exon 10	SV	?	-	[125]
136	R354C	Mut	c.1060C>T	Exon 8	SW	?	-	[107]
137	H28Frameshift	Mut	c.82insC	Exon 1	SW	?	-	[126]
138	62_172dup111	Mut	c.62_172dup111	Exon 1	SV	?	-	[127]
139	L261P	Mut	c.782T>C	Exon 7	SW	?	-	[127]
140	L363W	Mut	c.1088T>G	Exon 8	SV	?	-	[128]
141	R435C	Mut	c.1303C>T	Exon 10	NC	?	-	[114]
142	c.290-2A>G	Mut	c.290-2A>G	Intron 2	SW	?	-	[129]
143	S170Frameshift	Mut	c.508insA	Exon 4	SW	?	-	[129]
144	E246Frameshift	Mut	c.737delA	Exon 7	SW	?	-	[130]
145	A15T	Mut	c.43G>A	Exon 1	SV	?	-	[131]
146	H62L	Mut	c.185A>T	Exon 1	NC	?	-	[96]
147	D183Frameshift	Mut	c.549delC	Exon 5	SW	?	-	[102]
148	G291R	Mut	c.871G>C	Exon 7	?	?	-	[102]
149	S301Y	Mut	c.903C>A	Exon 7	NC	?	-	[102]
150	Y376X	Mut	c.1128C>A	Exon 9	SW	?	-	[102]
151	R124H	Mut	c.371G>A	Exon 3	NC	?	-	[132]
152	W19X	Mut	c.56G>A	Exon 1	SW	?	-	[116]
153	R483Frameshift	Mut	c.1147insC	Exon 10	SW	?	-	[116]

Table 10. cont'd

ID	Variant	Type	Position, nucleotide change	Region	Phenotype	Activity in vitro 17OHP/prog. (%)	Frequency in CYP21A2	Reference
154	R483W	Mut	c.1147C>T	Exon 10	SW	?	-	[116]
155	P213Frameshift	Mut	c.636insT	Exon 5	SW	?	-	[132]
156	W22Frameshift	Mut	c.64insT	Exon 1	SW	?	-	[133]
157	Q228X	Mut	c.682C>T	Exon 6	SW	?	-	[134]
158	M283L	Mut	c.847A>C	Exon 7	NC	?	-	[135]
159	P42Frameshift	Mut	c.126delC	Exon 1	?	?	-	[113]
160	H365Y	Mut	c.1093C>T	Exon 8	?	?	-	[113]
161	H28Frameshift	Mut	c.81_93delCACCTCCCG CCTC	Exon 1	SW	?	-	[136]
162	Q481P + P482S	Mut	c.1442C>T + c.1444C>T	Exon 10	SW	?	-	[137]
163	Q315X	Mut	c.943C>T	Exon 8	?	?	-	[131]
164	R444X +E6	Mut	c.1330C>T + c.707/c.710/c.716T>A	Exon 10 Exon 6	SW	?	-	[138]
165	V139E	Mut	c.416T>A	Exon 3	SW	?	-	[139]
166	C147R	Mut	c.439T>C	Exon 3	NC/SV	?	-	[139]
167	T295N	Mut	c.884C>A	Exon 7	SV/SW	?	-	[139]
168	L308F	Mut	c.922C>T	Exon 7	SV	?	-	[139]
169	R366C	Mut	c.1096C>T	Exon 8	?	?	-	[139]
170	M473I	Mut	c.1419G>T	Exon 10	N/NC	?	-	[139]
171	M1I	Mut	c.3G>A	Exon 1	SW	?		[132]
172	R233G	Mut	c.697A>G	Exon 6	NC	?	-	[139]
173	R483W	Mut	c.1447_C>T	Exon10	SW	?	-	[116]

3.13 Experimental Layout

All experiments are performed on 384-well micro plates. Long-range PCR is the first step of experiments, where CYP21A2 is amplified from the genomic DNA in one fragment. After the PCR is done, the master mix for nested PCR is prepared for each amplicon using amplicon-specific primers. Forward primers for nested PCR have the T7 and reverse primers the SP6 promoter tag. To 6.5 µl of master mix for each amplicon, 1 µl of long-range PCR product is added, giving a total volume of 7.5 µl. After nested PCR is complete, SAP treatment is performed by adding 3 µl SAP mix to each well, giving a total volume of 10.5 µl. Transcription master mix is prepared for four reactions and distributed in 2.25 µl volumes to every well on a new 384-well plate. 1 µl of SAP treated PCR product from each amplicon was added to the transcription mix and the reaction is performed. 1 µl from the remaining SAP treated amplicons 3, 4, 5/6, 7, 8 and 10 is taken for iPLEX assay groups 1, 2 and 3 and mixed with 2 µl iPLEX mix. 2 µl of remaining SAP treated product from amplicons 1, 3 and 10 is taken and mixed with 2 µl iPLEX mix for iPLEX assay group Poly. Extension reaction is performed with 75 cycles (Figure 11a and 11b).

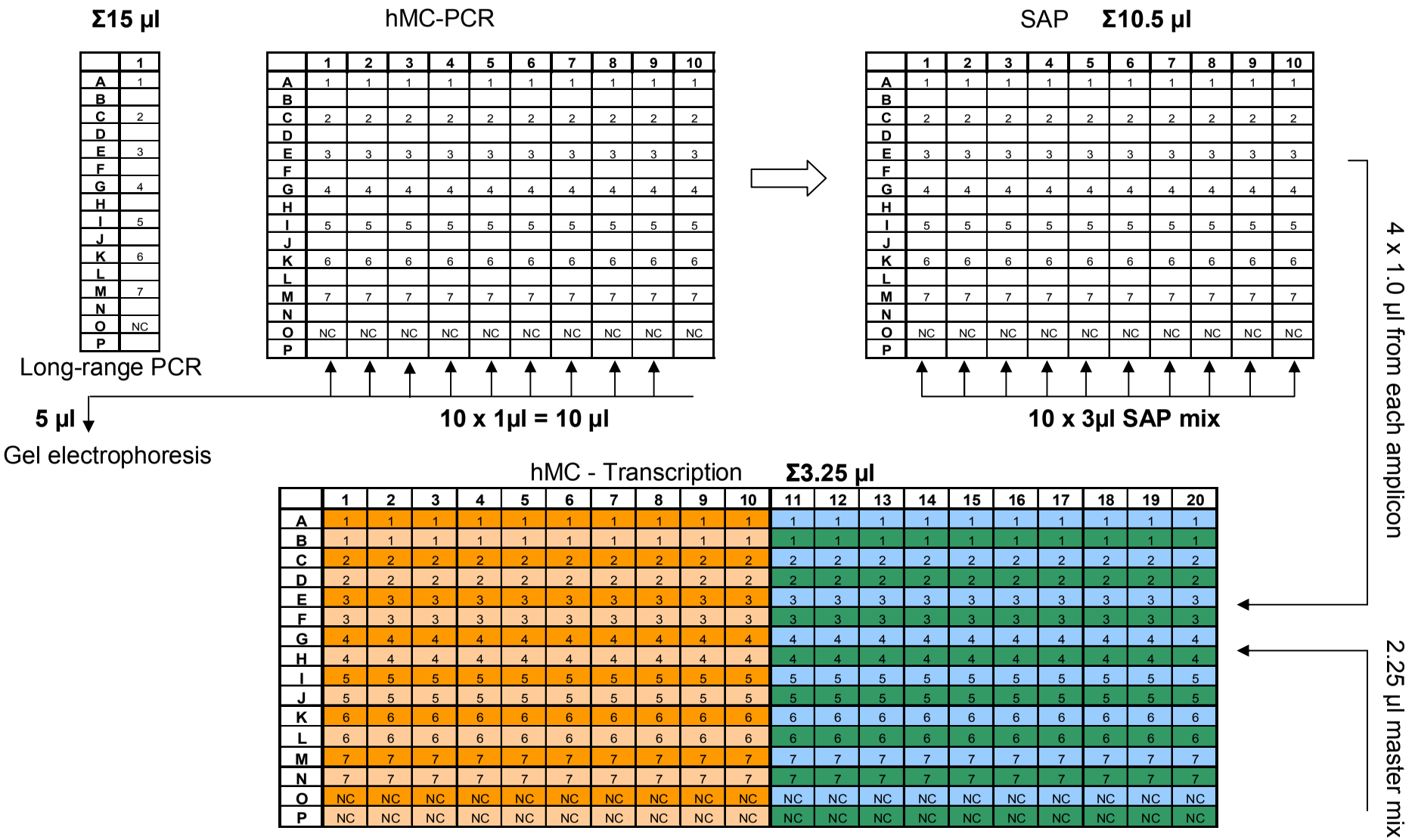
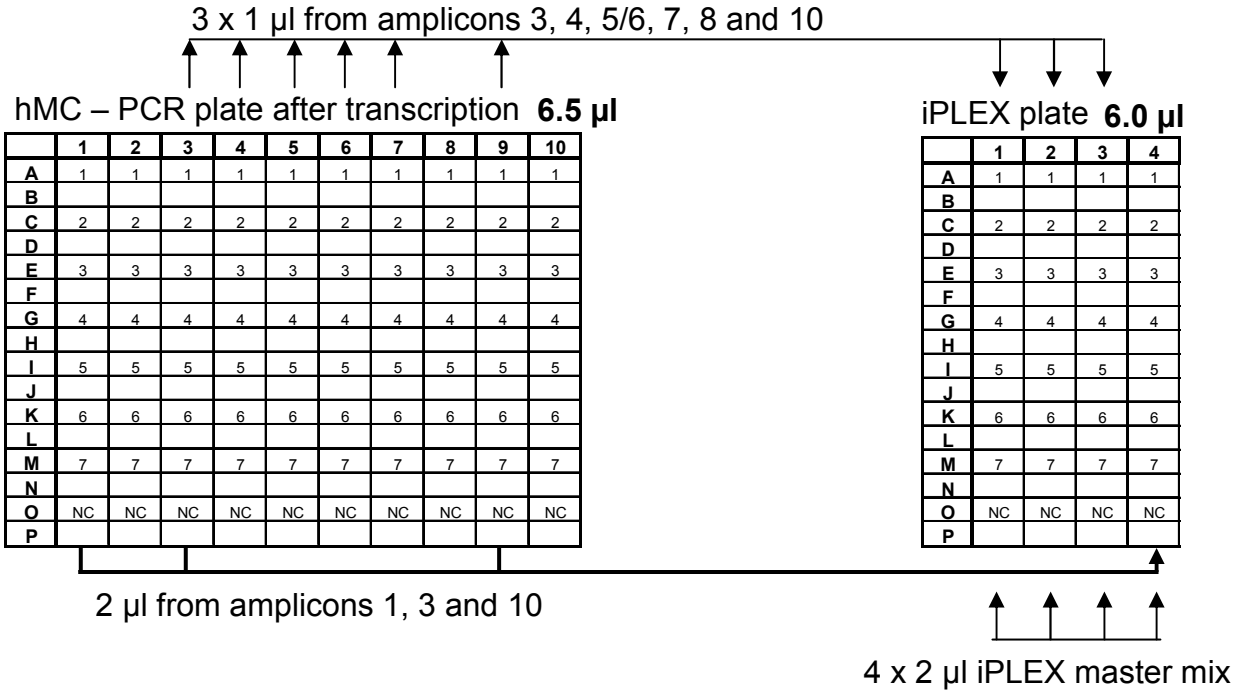


Figure 11a. Schematic representation of CAH mass spectrometry analyses experiments: The plate layout is presented from the first PCR to the transcription step together with volumes per well for seven samples and one negative control (NC).



	1	2	3	4	5	6	7	8	9	10
A	1	1	1	1	1	1	1	1	1	1
B										
C	2	2	2	2	2	2	2	2	2	2
D										
E	3	3	3	3	3	3	3	3	3	3
F										
G	4	4	4	4	4	4	4	4	4	4
H										
I	5	5	5	5	5	5	5	5	5	5
J										
K	6	6	6	6	6	6	6	6	6	6
L										
M	7	7	7	7	7	7	7	7	7	7
N										
O	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
P										

	1	2	3	4
A	1	1	1	1
B				
C	2	2	2	2
D				
E	3	3	3	3
F				
G	4	4	4	4
H				
I	5	5	5	5
J				
K	6	6	6	6
L				
M	7	7	7	7
N				
O	NC	NC	NC	NC
P				

Figure 11b. Schematic representation of CAH mass spectrometry analyses experiments: The plate layout is presented from the first PCR to the primer extension step together with volumes per well for seven samples and one negative control (NC). Assay groups 1, 2, 3 and Poly are represented by 1, 2, 3, 4, respectively.

3.14 Data Analysis

The in-house developed tool MS-Analyzer is used for post-measurement MassCLEAVE data analysis. Using MS-Analyzer in the analysis saves a considerable amount of time. Moreover, calls are grouped into high and low score lists according to allowed uncertainty levels, which can be edited by the user. This enables the analysis to be more flexible.

MS-Analyzer is inputted with the exported peak list from the system software DiscoveryRT. If at a certain position a certain change is expected, this can be loaded into the tool database beforehand with a desired name. It is then compared to the experimental data and the software displays all variations that are found in that specific amplicon per sample automatically together with their zygosity. The user does not have to locate the position and the base exchange, deletion or insertion manually. The results are shown in amplicon reports and they are grouped into high-

score and low-score calls. Variations found are classified into these two groups with respect to the number of expected and found differentiating signals. By changing the error factor in the software settings, one can adjust the accuracy and allowed uncertainty; thus, optimize the process for the best call rate. This automation step with additional settings such as, the analyzed mass range, check for salt adducts, doubly charged fragments, etc., saves a significant amount of time in data analysis. The default settings used are; error factor: 2.0, minimum mass cut-off: 1300, maximum mass cut-off: 7000, check for salt adducts: no, check for acy: no, check for doubly charged fragments: yes, check for matrix adducts: yes.

MassARRAY TyperAnalyzer 3.4 was used to interpret the spectrum outcome of iPLEX reactions. The program calibrates and analyzes the acquired spectra, calls genotypes automatically and displays them in a traffic light plot. This enables the experimenter to have a good overview of the assay efficiencies and overall quality. In case of poorly functioning assays or samples, visual investigation of the spectra to make operator calls or repeat experiments might be needed.

4 Results and Discussion

Congenital Adrenal Hyperplasia (CAH) due to 21-hydroxylase deficiency is one of the most common autosomal recessive disorders, which is analyzed genetically for mutations and polymorphisms with modern methods. Multiplex minisequencing [140], ligase-mediated mutation detection [141], fragment analysis with restriction digestion [142, 143], amplification-created restriction site [ACRS] [144], direct gene sequencing after amplification in one fragment [145] or two or more overlapping fragments using allele-specific primers [146] and linked microsatellites [5] are among the most common. To study large conversions, deletions and hybrid gene formations, real-time quantitative PCR-based methods [147] and Southern blotting [142, 143] and Multiplex Ligation-dependent Probe Amplification (MLPA) methods are reported [148].

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was first introduced in 1988 by Karas and Hillenkamp to analyze protein molecules which have molecular masses greater than 10000 Daltons [47]. Using MALDI-TOF MS to analyze nucleic acids has certain advantages over traditional methods. First of all, the method is more rapid. Genotype can be obtained in the scale of seconds while conventional methods usually need hours. Second, results are highly accurate and absolute since they are based on the intrinsic property of mass-to-charge ratio of the matter, which is not affected by secondary-structure formations in nucleic acids. Lastly, the automation of experimental steps and data acquisition/interpretation are more feasible than in conventional approaches [149].

Here, a new approach to detect mutations and polymorphisms in CYP21A2 is introduced as well as to quantitative analysis of CYP21A2 and CYP21A1P. Mutation and SNP detection begins with the long-range amplification of CYP21A2 gene. A nested PCR is performed using the crude first PCR products which produces shorter fragments for the mass spectrometry experiments. Homogenous mass-cleavage and extension reactions are run after SAP digestion. Finally mass spectrometry measurements are carried out, data are acquired and interpreted automatically. This approach combines mutation detection and CNV analysis on the same platform.

Throughout the entire study, the polymorphisms, pseudogene-derived mutations and rare mutations for CYP21A2 available in online database of The Human Cytochrome P450 Allele Nomenclature Committee, NCBI dbSNP database and molecular genetic laboratories of Bioglobe GmbH are taken into consideration. Position numbering is done relative to ENSEMBL transcript ENST0000448314 which matches to the gene build ENSG00000198457.

The results are grouped under two main headings: a) Mutation detection, b) CNV analysis.

4.1 Mutation Detection

4.1.1 Long-range Amplification of CYP21A2

CYP21A2-gene was specifically amplified by means of a primer mix containing a gene-specific forward primer and two universal reverse primers. The requirement of using a primer mix with a pair of universal reverse primers is due to the fact that the 3'-end of both genes lack conserved base pairs where gene-specific primers could be placed [68].

Bands of expected size were observed after gel electrophoresis after the long-range PCR. Specific amplification was confirmed by the following mass spectrometry experiments. The total volume of long range PCR mix should not be less than 15 µl in order to have enough material for all successive experiments. 15 µl and 25 µl of PCR protocols were tested for their efficiencies. 15 µl set-up delivered more intense bands and less by-products on the gel electrophoresis (Figure 12). This is likely to be due to the change of heat transfer dynamics because of the conical shape of the liquid inside the well as the level becomes higher. A final volume of 15 µl long-range PCR set-up was adopted as the initial step for the following experiments.

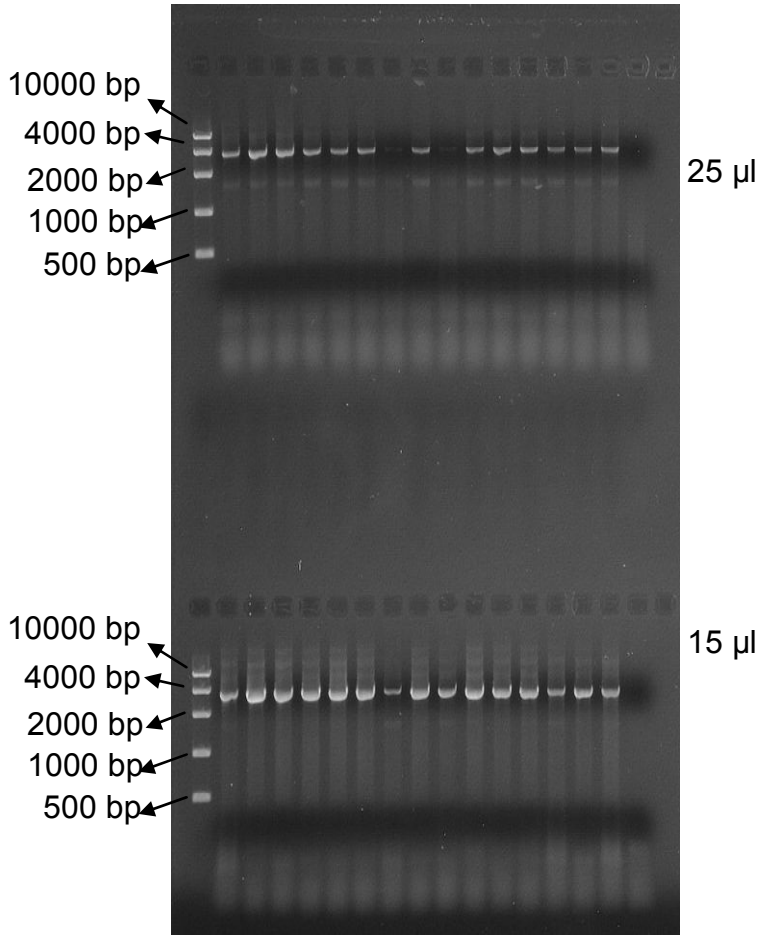


Figure 12. Gel electrophoresis of 15 and 25 µl long range PCR set-ups: The first and the last lanes in both parts belong to the DNA marker and the negative control, respectively.

4.1.2 Target Mutation List and Simulations

Each pseudogene-derived mutation, polymorphism and rare mutation was annotated on the wild type sequence of the relevant amplicon (reference sequence) and put into MS-Analyzer for simulation. For variations in close proximity, which affect the same wild type cleavage fragment in a certain reaction, haplotype combinations (the case of two or more variations being located on the same allele) were included in simulations. Simulation analysis showed that 89 variations in Table 10 are detectable with MassCLEAVE through differentiating signals resulting from the same cleavage reaction. 14 variations gave differentiating signals that are not generated in the same cleavage reaction but in two or more different reactions which cause a new signal to appear and a signal to disappear. In both cases the correct genotype could be

assigned by the automatic data analysis software successfully. Figure 13 displays an example of the simulation result for the mutation C147R. Differentiating signals in two reactions for wild type and mutation cases are highlighted in color. Complete simulation results are summarized in section 6.1.

		t-forward	t-reverse	c-forward	c-reverse
Mutation Signal	c.439_T>C(C147R)		2870,8	3991,4	
		t-forward	**t-reverse**	**c-forward**	**c-reverse**
Wild Type Signal	c.439_T>C(C147R)		2854,8	4295,6	1657

Figure 13. An example simulation of the mutation C147R: Two differentiating reactions are T-reverse and C-forward. Wild type and mutation signals coming from cleavage products of these reactions are highlighted in green.

For variations, which are not detectable with hMC, SNP-site extension assays were designed. Some positions were included in extension assays although they are analyzable with MassCLEAVE either because the signal resolution was not good enough for safe peak picking, or one or the other signal was not presented well enough due to the particular sequence at the region, which might hinder the enzyme performance. In total, these variations add up to 31.

The practical proof of every variation using real samples could not be tested within the allowed resources of this work due to very low frequencies. Although designing enough number of artificial model samples *in-vitro*, which would contain all rare mutations, was possible, it was economically not feasible. However; the proof of concept was shown in a reproducible and accurate manner for all variations present in real samples. Since the working principle of MALDI-TOF MS is straightforward, an extrapolation on variations, which were not available in real samples, can be done with confidence.

4.1.3 Base-specific Cleavage Analysis

CYP21A2 analysis using hMC is not a classical SNP discovery or species typing analysis because of the fact that the region harbors so many SNPs and mutations in such a short gene span. This represents an obstacle for the method. For this reason,

the first approach was to increase the amplicon size to include as many variations as possible to analyze provided that the region allows “SNP-free” spots for primers to bind. However, the amplicon size cannot be extended infinitely as this would finally result in overlapping signals and reduce the number of differentiating signals considerably. One of the simulation runs included larger amplicons with the aim of reducing the practical workload and increasing the throughput, but this trial suffered from conflicting signals for some vital mutations. Therefore, the set-up which uses the previously presented amplicons was kept in practice.

A shorter sub-amplicon, which is named Amp356, is 141 bp-long. It was formed solely to be able to detect the mutation p.Arg356Trp, since this important mutation was impossible to detect within Amp8 in the adopted experimental design and different primer extension assays designed for this position all failed to function. To tackle this difficult problem, a nested PCR primer, which would bind immediately to the 5'-upstream of the mutation, was designed and combined with the reverse nested PCR primer for Amp8. The new amplicon was cleaved *in silico* and observed to generate a differentiating signal for the targeted mutation and the possible interruption of three other secondary SNPs was eliminated automatically (Figure 14). To prevent any mismatch in case of presence of any of the three rare mutations in the primer binding region, universal binding base inosine was incorporated in the forward primer. Since inosine acts as a virtual guanine, the complementary base cytosine was used in the simulation for mass calculation of RNA cleavage fragments.

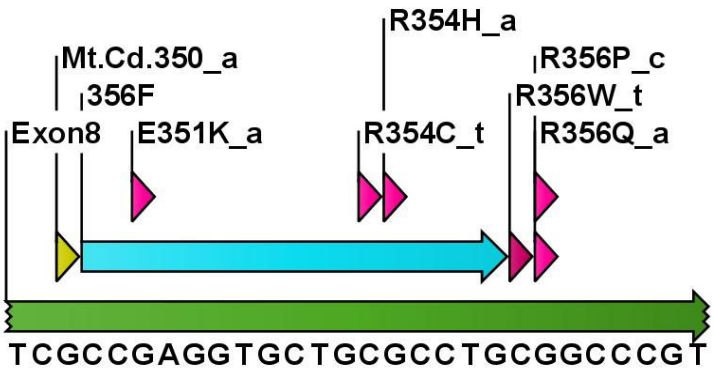


Figure 14. Nested PCR primer positioned adjacent to the mutation site: Blue arrow shows the nested PCR primer named 356F. Red flags indicate the position of rare mutations and the nucleotide change from the wild type. The yellow flag shows a rare mutation found in Bioglobe GmbH laboratories and its base change. Green arrow indicates exon 8 expanding to the left and right. Instead of three wild type bases in the primer binding region, universal base inosine was used in the primer sequence.

This final MassCLEAVE set-up was put in test after the extra amplicon for p.Arg356Trp was added to the design. The test included 6 real samples which possess the mutation p.Arg356Trp in heterozygous form, one wild type sample and one negative control. In all positive samples, the mutation was detected with a decent intensity and the wild type sample resulted in wild type peaks only. The negative control gave primer peaks and a few peaks of negligible intensity. Figure 15 presents the results of these real samples on MS-Analyzer.

Homogenous mass cleave reactions provide genotyping of 103 variations by means of cleaving ten different amplicons in base-specific manner in four reactions separately. A direct detection indicates that the genotype can be assigned analyzing only one of the four cleavage reactions. An indirect detection should be done when a single cleavage reaction does not provide enough information to assign the genotype, but the analysis of two or more reactions need to be combined. Defining all possible genotype combinations in advance, variations are reported automatically via MS-Analyzer tool. The overall setup was tested regularly with different primer mixes, stock solutions, chemical reagents and showed reproducibility, stability and reliability.

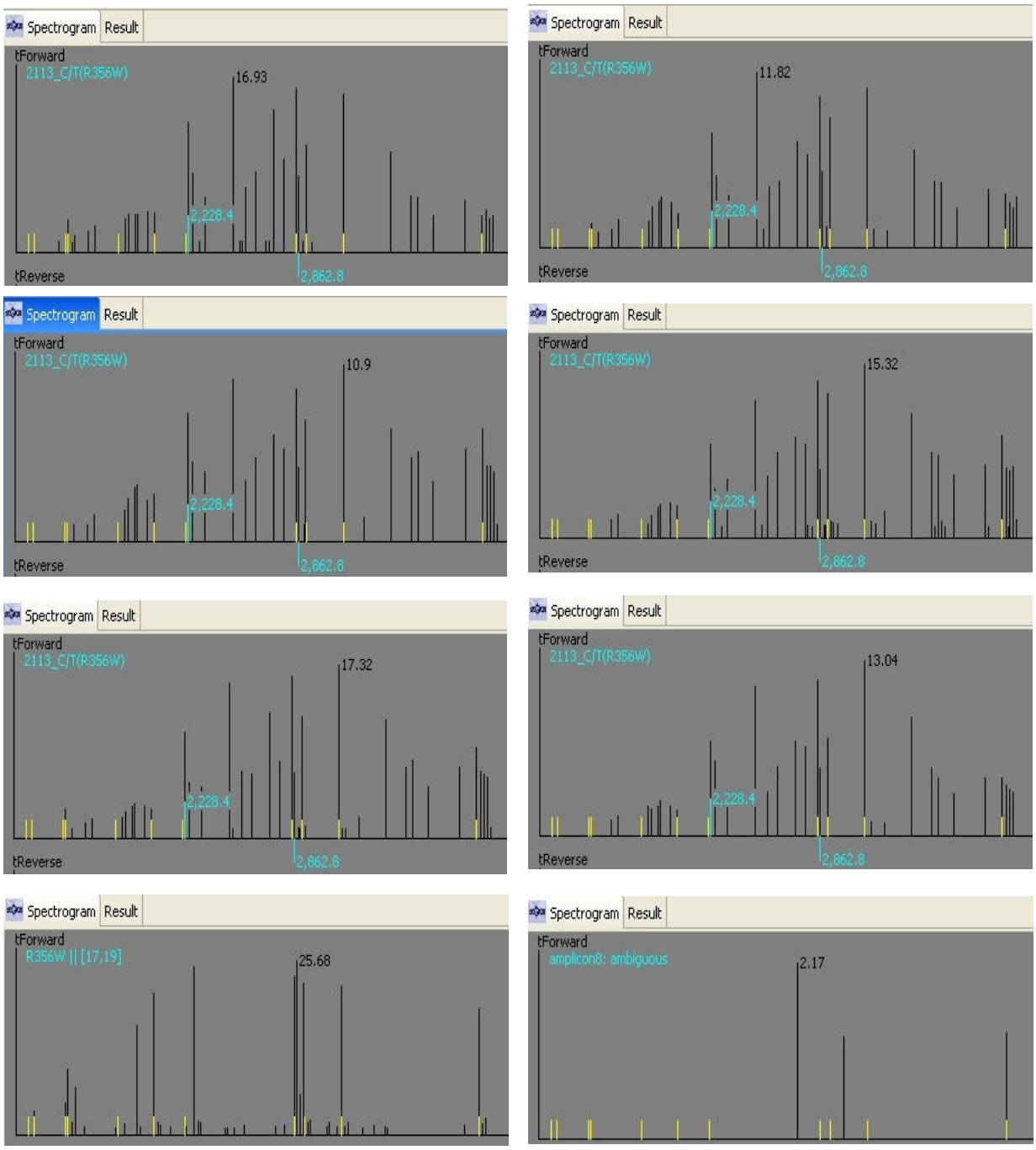


Figure 15. hMC results of the extra amplicon Amp356 with real samples: The first three rows of spectra from the top belong to samples which are heterozygous for the mutation p.Arg356Trp. Left hand side spectrum of the fourth row from the top displays a sample with a wild type allele for this amplicon. Right hand side spectrum of the fourth row from the top is the negative control.

4.1.4 Single Base-extension Analysis (iPLEX)

To analyze CYP21A2 using primer extension approach in a feasible number of multiplex groups is only possible by iPLEX approach; that is, single base extension

reaction. Designing assays using classical mass-extension approach (hME) with different stop mixes is not feasible due to the very high SNP density throughout the analyzed genomic region.

Extension primers are designed using single base extension approach specifically for each position to be analyzed. The size, thus the mass of each extension primer is different and is chosen in such a way that none of the analyte or extension primer masses overlaps with another analyte of any assay or its extension primer in a multiplex group and the complete mass range is efficiently used. AssayDesigner 3.1 is used to design extension assays, where one can set certain preferences such as minimum and/or maximum length, maximum and/or minimum mass cut-off, direction, minimum peak separation. The software displays likely risks such as hair-pin formation, inter- or intra-primer interactions roughly when it has finalized the design. When the designer fails to design an assay, the operator can design it manually by applying expert knowledge and taking into account the special sequences in the region. During the extension reaction the primer is annealed and elongated for one base only after which the elongation is terminated. Due to the chemistry, a possible unextended primer peak (unless the yield is 100%) and at least one analyte peak are expected on the spectrum for each assay.

The primer mix for each group contains extension primers with different masses. To equalize the representation of extension primers on the spectrum, the concentrations of high-mass primers are increased compared to low-mass primers, so that they have similar intensities throughout the spectrum. The concentration of each primer was modified until it reached the best performance during optimization experiments. The concentration of the extension primers in the reaction ranges from 1.3 μM to 2.7 μM.

During the assay design, considering neighboring SNPs and other assays to be multiplexed, some modifications were made in software default parameters, such as reducing the minimum primer length to 15 and the lowering mass limit to 4000 to utilize the low mass range more. Many assays had to be designed and tested completely manually due to very high density of secondary SNPs in the region. As a general rule, extension primers are allowed to have one possible mismatch when the

position is located at least 15 bases away from the extendable 3'-end during the manual design.

All constructed assays are presented below in detail (Table 11). Schematic plot of each assay is displayed in detail in Table 12. Some “special” assays need to be explained in more detail:

- 8bp: To detect the 8-bp deletion, one of the two variations which co-exist with this mutation, 717_C/T (c.339C>T), is used [68]. A “C” at this position would indicate that the eight base pairs are present, and a “T” that they are deleted.
- E351K: The 3rd base from the 3'-end is replaced with inosine to eliminate the possible interference of Mt. Cd. 350 (c.1048G>A).
- R426C: The 3rd base from the 3'-end is replaced with inosine to eliminate the possible interference of G424S (c.1273G>A).
- R426Hlong: In order for the assay not to fail in case of “TG” deletion, two extension primers are used in combination for this position. The wt-probe has the sequence shown in the table, the mt-probe matches to the indicated TG deletion (CA deletion in the primer). The “CA” tag in the 5'-end equals both primer masses.
- V281G: The 5th base from the 3'-end is replaced with inosine to eliminate the possible interference of M283L (c.847A>C).
- L317M/V: The 3rd base from the 3'-end is replaced with inosine to eliminate the possible interference of p.Gln318X (c.952C>T).
- Mt.Cd.450: The 8th base from the 3'-end is replaced with inosine to eliminate the possible interference of p.Pro453Ser (c.1357C>T).
- R483Q/P: In order for the assay to function in case of “G” deletion, two versions of extension primers were combined. The wt-probe has the sequence shown in the table, the mt-probe matches to the indicated “G” deletion (“C” deletion for the primer). The “C” tag in the 5'-end equals both primer masses.
- Mt.Cd.491: The 7th base from the 3'-end is replaced with inosine to eliminate the possible interference of P.Cd.493 (c.1478A>G (rs61732563)).
- 1789_G/C (c.936+11G>C(rs6442)): The 9th and the 10th bases from the 3'-end are replaced with inosine to eliminate the possible interference of 1780_T/G (c.936+2T>G) and/or 1779_G/C (c.936+1G>C).

- P105L: The 9th base from the 3'-end is replaced with inosine to eliminate the possible interference of P.Cd.102 (c.305A>G (rs6474)).
- W302R: The 6th base from the 3'-end is replaced with inosine to eliminate the possible interference of L300F (c.898C>T).
- V304M: The 13th base from the 3'-end is replaced with inosine to eliminate the possible interference of L308F (c.922C>T). In order for the assay to function properly in case of “T” insertion, two versions of extension primers were combined. The wt-probe has the sequence shown in the table and “A” tag in the 5'-end, the mt-probe matches to the indicated “T” insertion (“A” insertion in the primer), which equals both primer masses.
- P.560 (c.290-109G>C): The 14th base from the 3'-end is replaced with inosine to eliminate the possible interference of 546_C/A (c.290-117C>A) .

Table 11. iPLEX assays represented in detail: Unextended primer and analyte (extension product) masses, their sequences, direction of elongation, neighboring SNPs and extension primer binding sites are shown schematically. **UEP**: Unextended primer. **rev**: Primer binding to the downstream of the variation. **fwd**: Primer binding to the upstream of the variation. **i**: Universal binding base inosine to enable annealing in the presence of secondary SNPs. Other annotations for primer binding regions are kept in the figures for better visualization. Red arrows and flags indicate rare mutations. Yellow flags and arrows indicate mutations which are found in Bioglobe laboratories. Since the present nomenclature was not adopted at the time when the design experiments were carried out, screenshots and assay names contain the old nomenclature. Correspondents in the present nomenclature are provided for consistency when necessary.

Multiplex Group 1		
8bp (rev)		
UEP	TTGTGGGCTTTCCAGAGCAG	6164 Da
Analyte-C	TTGTGGGCTTTCCAGAGCAGG	6451.2 Da
Analyte-T	TTGTGGGCTTTCCAGAGCAGA	6435.2 Da
E351K (fwd)		
UEP	TTGCTCAATGCCACCATCiCC	6287.09 Da
Analyte-A	TTGCTCAATGCCACCATCiCCA	6558.3 Da
Analyte-G	TTGCTCAATGCCACCATCiCCG	6574.3 Da
G291C (fwd)		
UEP	GCAGTGGACCTCCTGATC	5475.6 Da
Analyte-A	GCAGTGGACCTCCTGATCA	5746.8 Da
Analyte-C	GCAGTGGACCTCCTGATCC	5722.7 Da
Analyte-G	GCAGTGGACCTCCTGATCG	5762.8 Da
Analyte-T	GCAGTGGACCTCCTGATCT	5802.7 Da
P453S (rev) (p.Pro453Ser)		
UEP	GGGCAGGGCGTCCCCGGAGG	6225 Da
Analyte-C	GGGCAGGGCGTCCCCGGAGGG	6512.2 Da
Analyte-T	GGGCAGGGCGTCCCCGGAGGA	6496.2 Da
R356P/Q (rev)		
UEP	TAAGGGCACAACGGGC	4940.2 Da
Analyte-A	TAAGGGCACAACGGGCT	5267.3 Da
Analyte-C	TAAGGGCACAACGGGCG	5227.4 Da
Analyte-G	TAAGGGCACAACGGGCC	5187.4 Da
R426C (fwd)		
UEP	TCTGGCCTTCGGCTGCGGTiCC	6695.32 Da
Analyte-C	TCTGGCCTTCGGCTGCGGTiCCC	6942.5 Da
Analyte-T	TCTGGCCTTCGGCTGCGGTiCCT	7022.42 Da
R426Hlong (rev)		
UEP	CGCCAGCGGCTCGCCCAGGCACACG	7599 Da
Analyte-A	CGCCAGCGGCTCGCCCAGGCACACGT	7926.2 Da
Analyte-G	CGCCAGCGGCTCGCCCAGGCACACGC	7846.2 Da

Table 11. cont'd

V281G (rev)		
UEP	GGAGGTCCACTGCAGCCAiGTGC	7075.59 Da
Analyte-G	GGAGGTCCACTGCAGCCAiGTGCC	7322.77 Da
Analyte-T	GGAGGTCCACTGCAGCCAiGTGCA	7346.79 Da
V281L (fwd) (p.Val281Leu)		
UEP	CTCCTGGAAGGGCAC	4577.99 Da
Analyte-G	CTCCTGGAAGGGCACG	4865.19 Da
Analyte-T	CTCCTGGAAGGGCACT	4905.08 Da
Multiplex Group 2		
A362V (fwd)		
UEP	GCCCGTTGTGCCCTTAG	5153.4 Da
Analyte-C	GCCCGTTGTGCCCTTAGC	5400.6 Da
Analyte-T	GCCCGTTGTGCCCTTAGT	5480.6 Da
E6 (fwd)		
UEP	AGGCCATAGAGAAGAGGGA	5959.9 Da
Analyte-C	AGGCCATAGAGAAGAGGGAC	6207.1 Da
Analyte-T	AGGCCATAGAGAAGAGGGAT	6287 Da
G178R (rev)		
UEP	TGAGGCACCTTGATCTTGTCTC	6692.4 Da
Analyte-A	TGAGGCACCTTGATCTTGTCTCT	7019.6 Da
Analyte-C	TGAGGCACCTTGATCTTGTCTCG	6979.6 Da
Analyte-G	TGAGGCACCTTGATCTTGTCTCC	6939.6 Da
I172N (rev) (p.Ile172Asn)		
UEP	CGAAGGTGAGGTAACAG	5308.5 Da
Analyte-A	CGAAGGTGAGGTAACAGT	5635.6 Da
Analyte-T	CGAAGGTGAGGTAACAGA	5579.7 Da
L317M/V (rev)		
UEP	CGTGGTCTAGCTCCTCCTiCA	6334.11 Da
Analyte-A	CGTGGTCTAGCTCCTCCTiCAT	6661.2 Da
Analyte-C	CGTGGTCTAGCTCCTCCTiCAG	6621.32 Da
Analyte-G	CGTGGTCTAGCTCCTCCTiCAC	6581.29 Da
Mt.Cd.241 (rev)		
UEP	GTCCCCACCTTGTGCTGCCTCA	6598.4 Da
Analyte-C	GTCCCCACCTTGTGCTGCCTCAG	6885.6 Da
Analyte-G	GTCCCCACCTTGTGCTGCCTCAC	6845.6 Da
Mt.Cd.450 (rev)		
UEP	GTCCCCGGAGGiCAGCAGC	5824.78 Da
Analyte-C	GTCCCCGGAGGiCAGCAGCG	6111.99 Da
Analyte-T	GTCCCCGGAGGiCAGCAGCA	6095.98 Da

Table 11. cont'd

R483Q/P (rev)		
UEP	GTGGGCCCCCATCCCC	4779.2 Da
Analyte-A	GTGGGCCCCCATCCCCT	5106.4 Da
Analyte-C	GTGGGCCCCCATCCCCG	5066.4 Da
Analyte-G	GTGGGCCCCCATCCCCC	5026.4 Da
Mt.Cd.491 (rev)		
UEP	CCCATCACTGGiTCTGGC	5436.53 Da
Analyte-A	CCCATCACTGGiTCTGGCT	5763.63 Da
Analyte-G	CCCATCACTGGiTCTGGCC	5683.72 Da

Multiplex Group 3

1789_G/C (fwd) (rs6442)		
UEP	CCACCCTGAGiiGCGTCCTG	6064.93 Da
Analyte-C	CCACCCTGAGiiGCGTCCTGC	6312.11 Da
Analyte-G	CCACCCTGAGiiGCGTCCTGG	6352.14 Da
P105L (fwd)		
UEP	CTGGTGTCTAiGAACTACC	5788.77 Da
Analyte-C	CTGGTGTCTAiGAACTACCC	6035.96 Da
Analyte-T	CTGGTGTCTAiGAACTACCT	6115.87 Da
W302R (fwd)		
UEP	GACCCCAGCAAACACCiTCTCA	6603.31 Da
Analyte-C	GACCCCAGCAAACACCiTCTCAC	6850.5 Da
Analyte-T	GACCCCAGCAAACACCiTCTCAT	6930.41 Da
R339H (fwd)		
UEP	GGGTCCCCTACAAGGACC	5469.6 Da
Analyte-A	GGGTCCCCTACAAGGACCA	5740.8 Da
Analyte-G	GGGTCCCCTACAAGGACCG	5756.8 Da
V304M (rev)		
UEP	GGTGGTGAAiCAAAAAAAACCA	6819.48 Da
Analyte-A	GGTGGTGAAiCAAAAAAAACCAT	7146.58 Da
Analyte-G	GGTGGTGAAiCAAAAAAAACCAC	7066.66 Da
Q318X (rev) (p.Gln318X)		
UEP	TGGTCTAGCTCCTCCT	4799.1 Da
Analyte-C	TGGTCTAGCTCCTCCTG	5086.3 Da
Analyte-T	TGGTCTAGCTCCTCCTA	5070.3 Da

Table 11. cont'd

Multiplex group Poly		
-4_C/T (fwd) (c.1-4C>T)		
UEP	CCCTGACGGGCGTCT	4544.95 Da
Analyte-C	CCCTGACGGGCGTCTC	4792.14 Da
Analyte-T	CCCTGACGGGCGTCTT	4872.05 Da
P.Cd.493 (rev) (rs61732563)		
UEP	TCCTGCCCCATCACTGG	5082.31 Da
Analyte-A	TCCTGCCCCATCACTGGT	5409.4 Da
Analyte-G	TCCTGCCCCATCACTGGC	5329.49 Da
P.560 (fwd) (c.290-109G>C)		
UEP	GTCiCTGGAAGCTCTTG	5202.38 Da
Analyte-C	GTCiCTGGAAGCTCTTGC	5449.57 Da
Analyte-G	GTCiCTGGAAGCTCTTGG	5489.59 Da
P.601 (rev) (rs6451)		
UEP	GAGGGCCTGACCTTCTTC	5466.55 Da
Analyte-A	GAGGGCCTGACCTTCTTCT	5793.65 Da
Analyte-C	GAGGGCCTGACCTTCTTCG	5753.76 Da
Analyte-G	GAGGGCCTGACCTTCTTCC	5713.74 Da
delCTG (rev) (rs28381641)		
UEP	CAGGGGCAGCAGCAGCAG	5583.64 Da
Analyte-4CTG	CAGGGGCAGCAGCAGCAGG	5870.84 Da
Analyte-5CTG	CAGGGGCAGCAGCAGCAGC	5830.82 Da
P.629-30 (rev) (rs35147842)		
UEP	GCTGGAGGGTGGGAACTGA	5973.88 Da
Analyte-CA	GCTGGAGGGTGGGAACTGAT	6300.98 Da
Analyte-GG	GCTGGAGGGTGGGAACTGAC	6221.07 Da
P.490 (fwd) (rs61732562)		
UEP	GGGATGGGGGCCCACAGCCC	6168.99 Da
Analyte-A	GGGATGGGGGCCCACAGCCCA	6440.2 Da
Analyte-G	GGGATGGGGGCCCACAGCCCG	6456.2 Da

Table 12. Extension primer binding sites are shown together with neighboring SNPs schematically. Annotations follow the same rule as in Table 11. Other annotations for primer binding regions are kept in the figures for better visualization. Red arrows and flags indicate rare mutations. Yellow flags and arrows indicate mutations which are found in Bioglobe laboratories. Since the present nomenclature was not adopted at the time when the design experiments were carried out, screenshots and assay names contain the old nomenclature. Correspondents in the present nomenclature are provided for consistency when necessary.

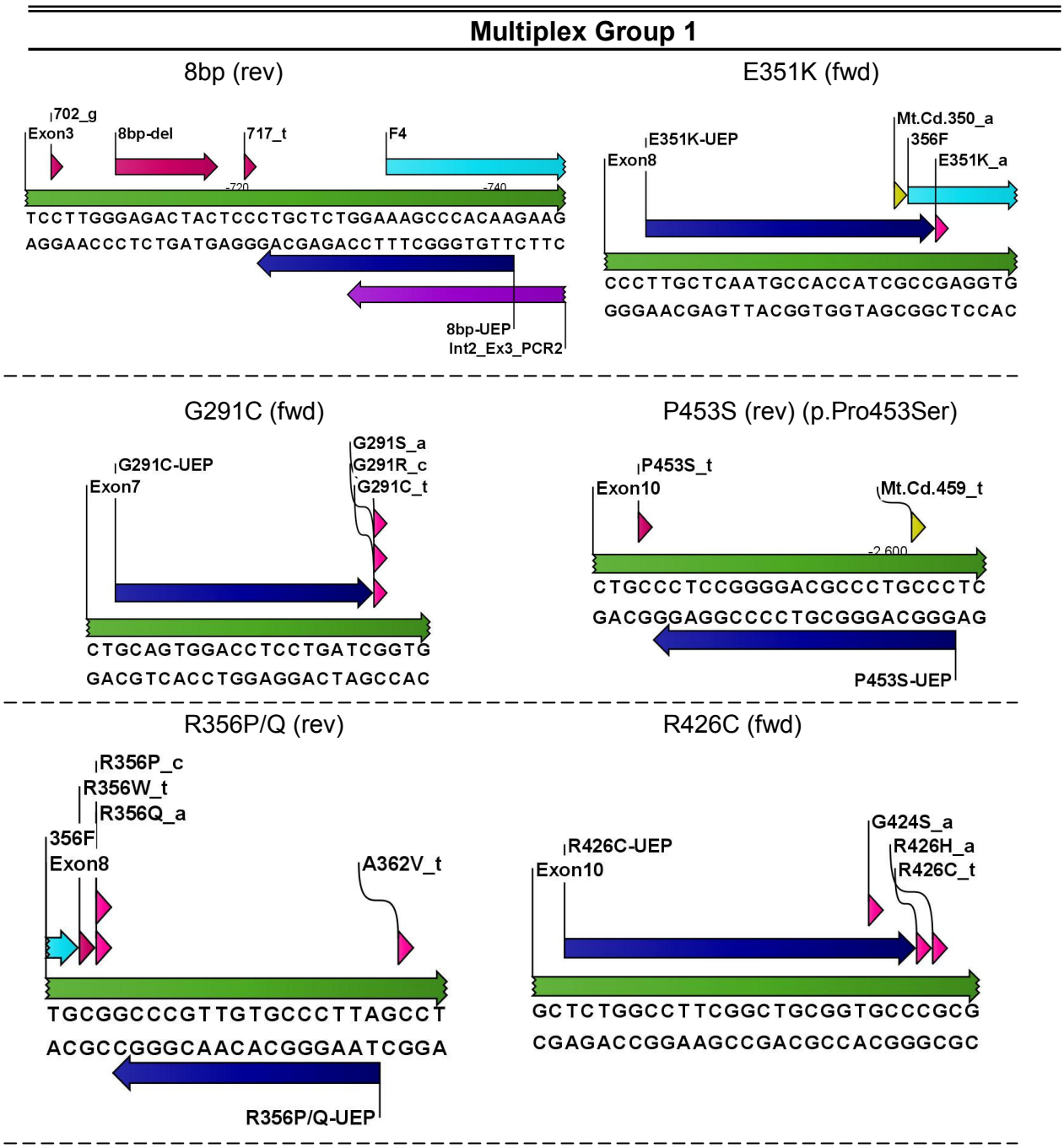
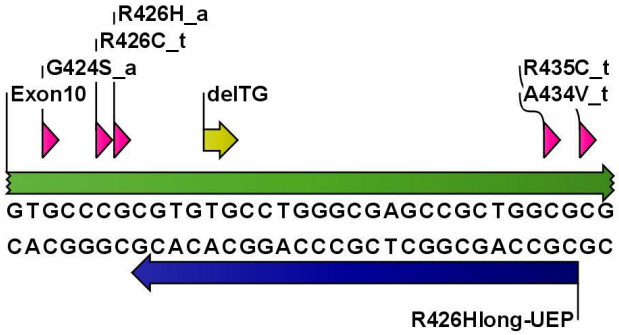
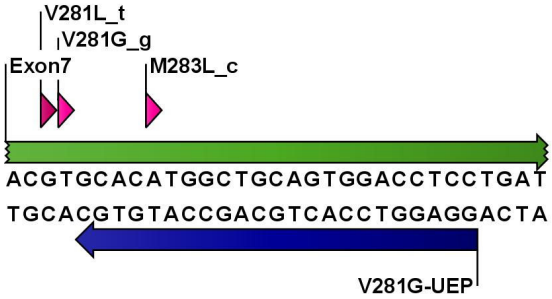


Table 12. cont'd

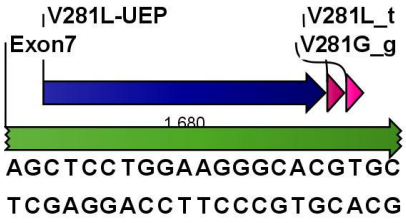
R426Hlong (rev)



V281G (rev)

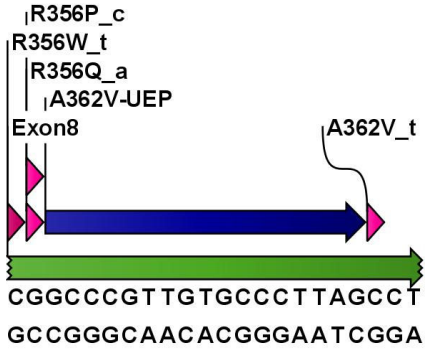


V281L (fwd) (p.Val281Leu)

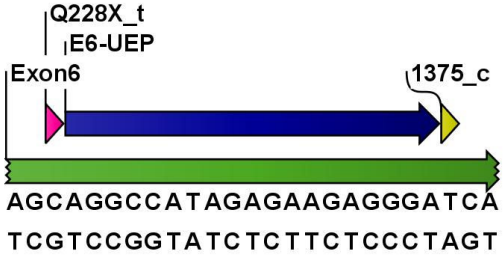


Multiplex Group 2

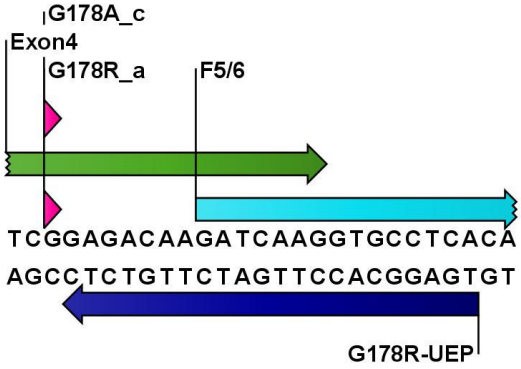
A362V (fwd)



E6 (fwd)



G178R (rev)



I172N (rev) (p.Ile172Asn)

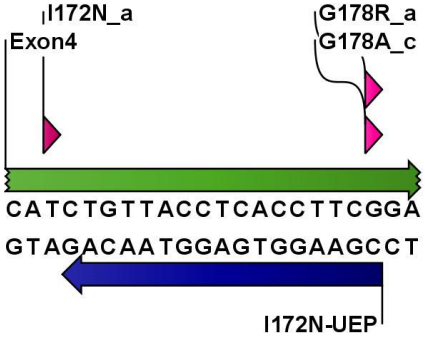
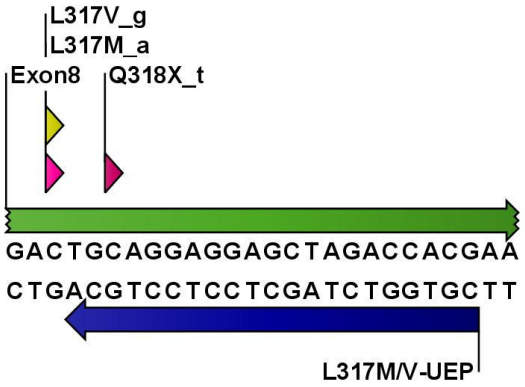
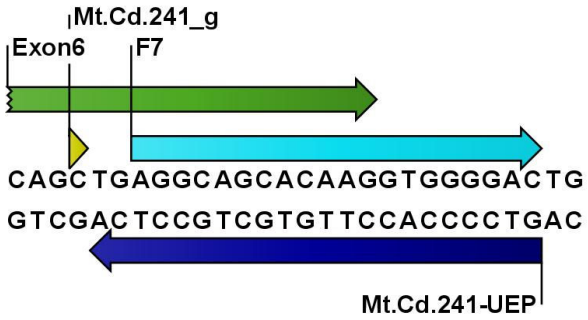


Table 12. cont'd

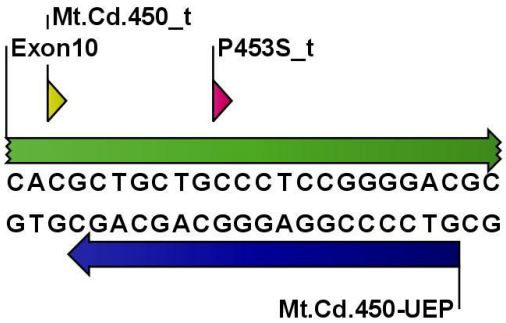
L317M/V (rev)



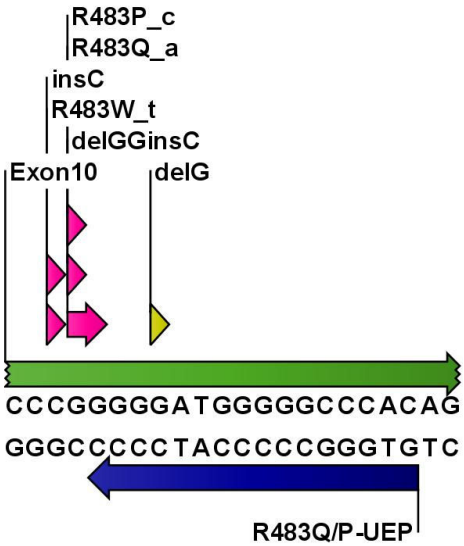
Mt.Cd.241 (rev)



Mt.Cd.450 (rev)



R483Q/P (rev)



Mt.Cd.491 (rev)

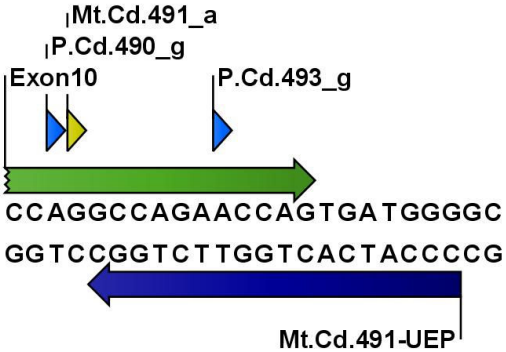
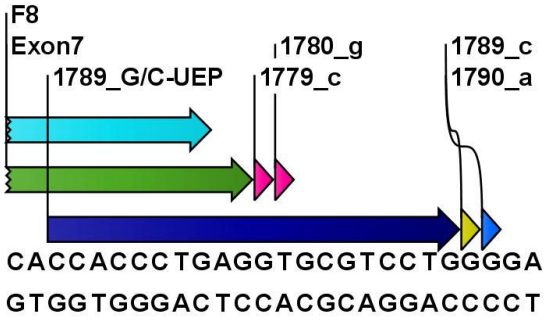


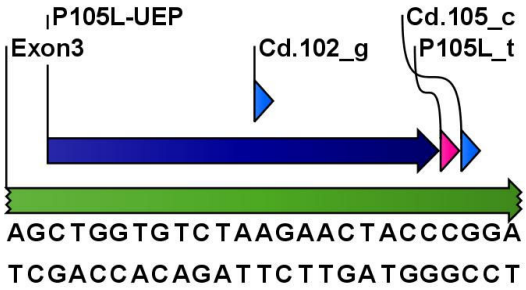
Table 12. cont'd

Multiplex Group 3

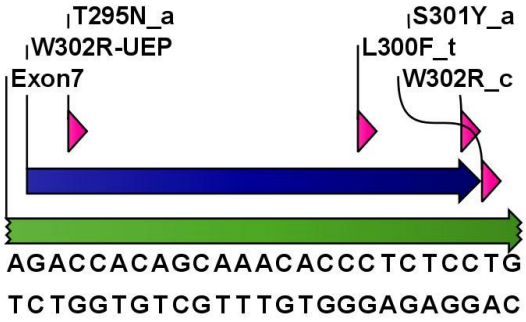
1789_G/C (fwd) (rs6442)



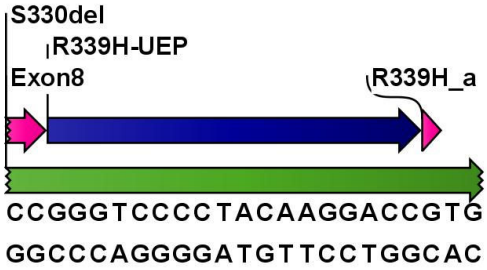
P105L (fwd)



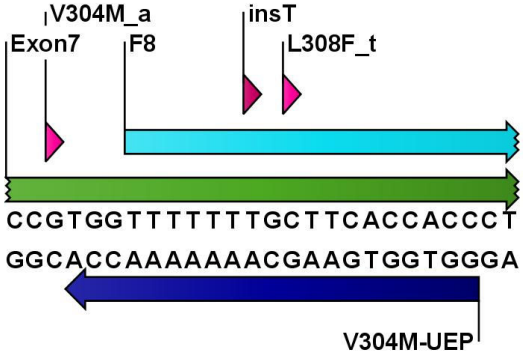
W302R (fwd)



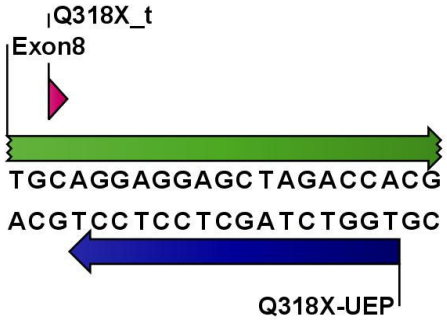
R339H (fwd)



V304M (rev)

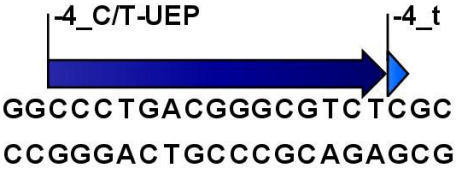


Q318X (rev) (p.Gln318X)



Multiplex group Poly

-4_C/T (fwd) (c.1-4C>T)



P.Cd.493 (rev) (rs61732563)

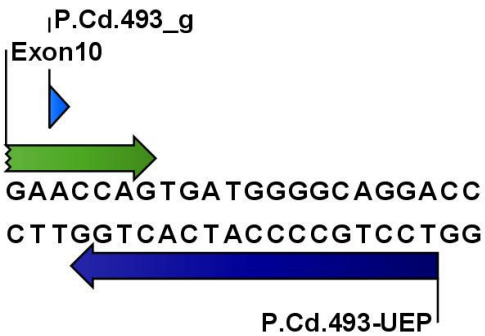
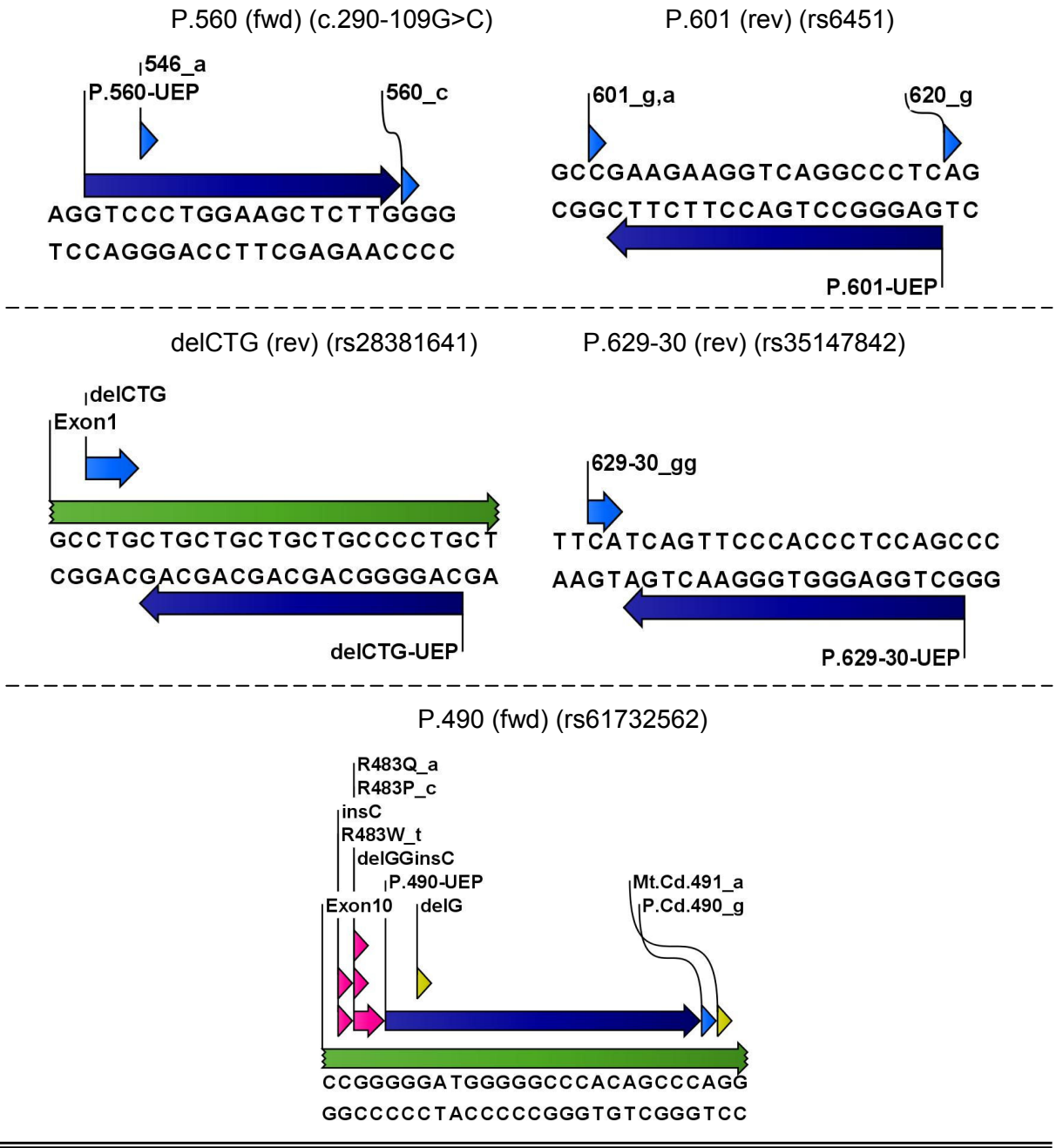


Table 12. cont'd



iPLEX-Gold reaction kit is designed for analyzing high number of SNPs simultaneously; that is, for higher levels of multiplexing. There are, however, two recipes in the manufacturer's manual; one for low multiplexing (1-18) and one for high multiplexing (19-35+). The reaction kit was tested for its performance on CAH analysis with low and high multiplexing recipes. Although in most of the cases the genotype was assigned correctly, a second analyte peak appeared in some cases with too high of an intensity which made automatic calling impossible. An example of this case is shown in Figure 16. This phenomenon is interpreted to be due to the

enzyme in the iPLEX kit being highly processive. This is a useful characteristic for high levels of multiplexing, but when dealing with a lower number of multiplexes, it might sometimes have a negative effect of terminating the elongation with a wrong base, hence necessitating manual user calls. Such an uncertainty could cause important misinterpretations and wrong conclusions in a diagnostic environment. Especially regarding a complex region like the HLA class, where changes in copy numbers is a relatively more frequent event, this undesired result could lead to even more serious problems, which cannot be accepted within the scope of a diagnostic kit.

To avoid this, MassEXTEND enzyme Thermosequenase, which is a less processive enzyme than the original iPLEX enzyme, was implemented into the recipe without changing final concentrations. This test phase was then expanded to include two buffer and enzyme combinations with iPLEX and classical MassEXTEND cycling programs. After direct comparison among parallel extension reactions, Thermosequenase with appropriate buffer and classical MassEXTEND cycling showed the highest efficiency and call rate. The correct genotype obtained from direct sequencing was confirmed with this buffer-enzyme pair and cycling condition. This approach was adopted for further experiments.

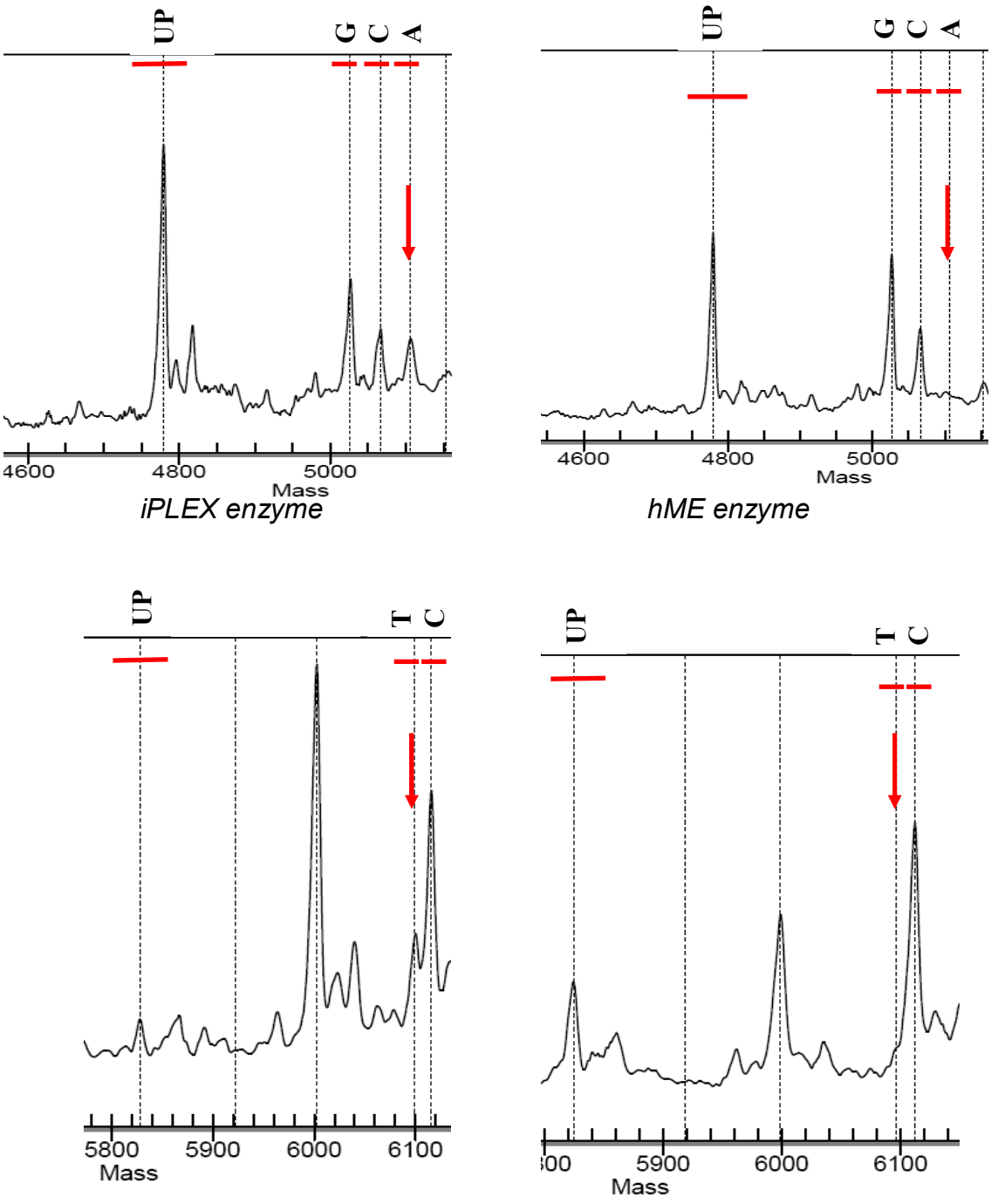


Figure 16. Spectra of two assays obtained from the same test sample with iPLEX enzyme and Thermosequenase: R483Q/P gives the A-analyte peak with iPLEX enzyme while the spectrum shows a clear GC genotype with Thermosequenase (upper row). Mt.Cd.450 gives a small peak for T-analyte with iPLEX enzyme whereas the genotype is clearly CC with Thermosequenase (lower row). Unlabeled peaks belong to other assays.

4.1.5 Routine Analysis – Protocol Evaluation

After the simulations for hMC and assay development for iPLEX had been performed, the protocol was tested in parallel with real samples that were sent to Bioglobe laboratories for CAH molecular genetic testing with conventional sequencing. The results were then compared to those of routine analysis with gene sequencing. The analysis included a total of 36 samples. Detailed results of hMC and iPLEX tests are plotted in section 6.3. All matches between gene sequencing and mass spectrometry tests are indicated with green color. For all positions in the table, the number of mismatches is 6 positions in a total of 10 samples.

One problem was observed to be the inability of MS-Analyzer to pick the mutant signal for 620_A/G (c.290-48A>G). The peak was, however, detected by Discovery-RT, hence was possible to be analyzed manually, which proves that biochemical reactions take place as normal and expected peaks are generated. A similar problem was observed in one case at 395_T/C (rs6462) and at c.290-44G>T (rs6453) (highlighted in red color in the table).

As discussed before, MS-Analyzer is a prototype tool which has many rooms for development. The tool searches for the presence of all identifying signals in all four reactions to call the variation. To enable the automatic calling of variations whose distinguishing signals are in very low or high mass ranges, a best identifying reaction could be defined to be considered in the calling algorithm. This would eliminate the signals with insufficient intensities which hinder the call. In conclusion, these variations are detectable with the protocol, but due to the present algorithms of the scoring tool, they cannot be called automatically most of the time.

At two positions in four samples in the test set, mass spectrometry and gene sequencing delivered discordant results:

1) In samples 3768 and 3792, the mass spectrum tests resulted in a heterozygous genotype for the position rs6472 (c.803G>C), while gene sequencing resulted in wild type in one case and homozygous mutant in the other (indicated with blue color in the table). The chromatogram data and mass spectra for the position in question for

both samples are shown in Figure 17. In 3768, both wild type and mutant signals are clearly present in the mass spectrum. In 3792, due to overall lower signal intensity, the mutant signal is not as strong but it is still present when compared to other wild type samples. On the chromatogram of both samples, it is easily noticed that the homozygous signals at the position in question are much weaker compared to the general intensity behavior in that region. This suggests that the mutant signal is not represented enough to appear as a peak on the chromatogram. This is a relevant example which illustrates the advantage of mass spectrometric techniques resulting in separate signals for heterozygous samples rather than overlapping peaks and how this helps the operator in certain cases.

It can be concluded at this point that in case of a spectrum of insufficient data quality or low intensity, a re-measurement (or a complete repeat) should be performed in routine diagnostics.

2) In two samples at position rs61732563 (c.1478A>G) there was a mismatch between the gene sequencing result and the mass spectrometry result (Figure 18). The genotype was assigned AG in the Typer Software for both samples at this position, whereas it was assigned AA for both by sequencing. Both samples were repeated with both methods and the same results were obtained. As suggested by the lab team too, the decrease in the intensity of the second “A” peak might indicate the failure of representation of the missing allele in the chromatogram. iPLEX result for this position is clear and straight-forward not to leave any dubious thoughts on the final outcome.

It is also worth mentioning that there were 7 discrepancies detected during the comparison. However, post-analytical investigation revealed operator errors (subjective interpretations) of sequencing data, which were corrected in the final result. This fact also shows that data interpretation on mass spectrometry platform is less operator-error-prone than the traditional DNA sequencing.

The settings of the analysis tool MS-Analyzer play an important role in achieving reliable automatic calling of hMC results. During the evaluation phase, it was run with error factor set to 5.0, minimum mass cut-off to 1300 Da and maximum mass cut-off to 7000 Da. Increasing the error factor from the default value of 2.0 to 5.0 minimizes the risk of missing any variation. But it also introduces some false variations to the low score list. This, however, is consistent; that is, false variations that are reported in low score calls are consistent, therefore it can easily be concluded that they are not actually present and can be ignored. Another modification in the settings was to deactivate the check for salt and matrix adducts. This is changed only for the position c.289+33C>A(rs6463) since the wild type signal in T-forward reaction is affected by a nearby strong salt adduct. If MS-Analyzer searches for adduct peaks, the wild type signal for this position might be mistakenly taken as the adduct peak and not reported. No other variation showed decreased call rates due to these modifications.

After the scoring is done and present variations are acquired using the settings described above, the software is run once more with the minimum mass cut-off of 1700. This is necessary for the evaluation of positions c.1122C>T(rs6469) and c.305A>G(rs6474). It eliminates the obligation to detect the mutant signal in T-forward for c.1122C>T(rs6469), which is 1304.8 Da and has a very low intensity, and the mutant signal in C-reverse for c.305A>G(rs6474), which is 1632 Da with a low intensity. With this modification, the reaction which results in low intensity signals for these variations is excluded from the analysis automatically and their automatic calls are made possible. As mentioned before, this would not be necessary if a best identifying reaction could be defined in the software algorithm and other reactions could be discarded. This would save even more time in the data interpretation and make the automation of hMC results faster.

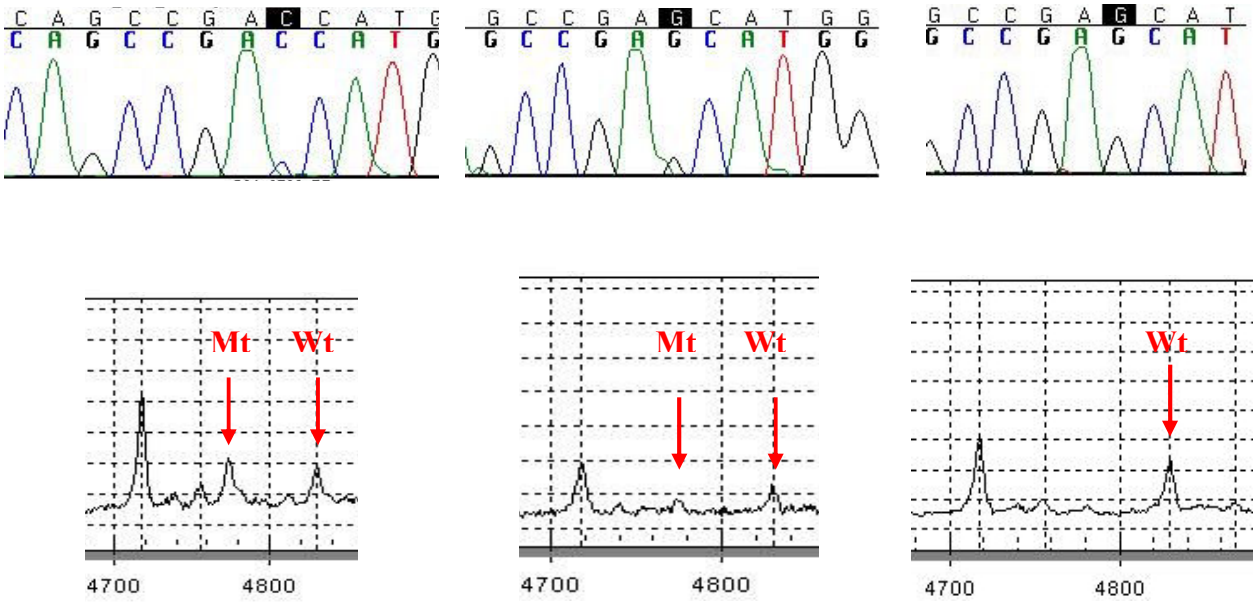


Figure 17. Two examples where the results of two methods do not agree with each other for the position rs6472 (c.803G>C): Chromatogram data are presented in the upper portion and mass spectrum data in the lower portion. Third figure on the right hand side shows a wild type result for comparison purposes. Wild type and mutant signals for MassCLEAVE are indicated with red arrows. Unlabeled signals belong to other variations.

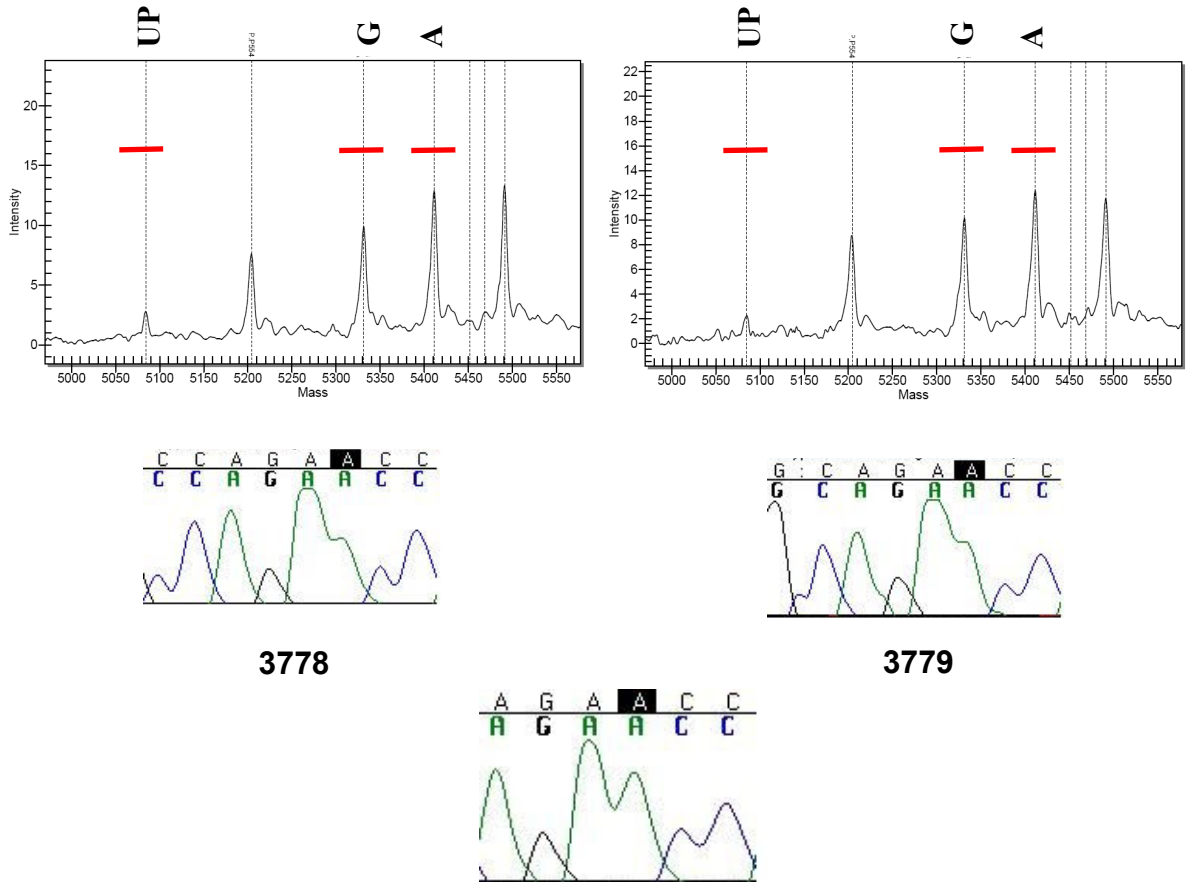


Figure 18. Two cases where iPLEX test resulted in heterozygous genotype and gene sequencing in wild type for the position rs61732563 (c.1478A>G): The upper portion shows iPLEX spectra and the middle portion sequencing chromatogram. A homozygous result is presented in the lower portion for comparison purposes. UEP peak and analyte peaks for the relevant assay are indicated with red dashes in the spectra. Unlabeled peaks belong to other assays.

The expected genotypes after gene sequencing were reproduced up to 98.61% using MassCLEAVE and iPLEX methods of mass spectrometry technology. The percentage given is for testable variations; therefore, it represents the cases where a variation is present and can be tested. Furthermore, all operator errors in the final results related to gene sequencing were detected and corrected based on the high accuracy of mass spectrometry data. The time needed to interpret data was seen to be much less compared to gene sequencing since one does not have to follow base pairs one by one on the computer screen, which ultimately increases the likelihood of missing a mutation or a polymorphism especially when it is present in a homozygous state.

In addition, indications of copy number variations in some samples were possible to be observed already in the mutation detection phase. Assays whose analytes have quite different intensities enable one to hypothesize the occurrence of a duplication/deletion or a large conversion. Figure 19 displays an example case. Sample 39 has a duplicated CYP21A2-gene. The duplicated CYP21A2 gene has a 3'-end conversion: downstream and upstream of exon 3 is CYP21A1P- and CYP21A2-like, respectively. This event is observable in the spectra of iPLEX assays E6 and p.Ile172Asn. Both assays result in CYP21A2-specific genotypes roughly with the double intensity of CYP21A1P-specific genotypes. CYP21A2-specific genotypes come from the intact active gene and CYP21A1P-specific genotypes result due to the converted portion of the duplicated active gene.

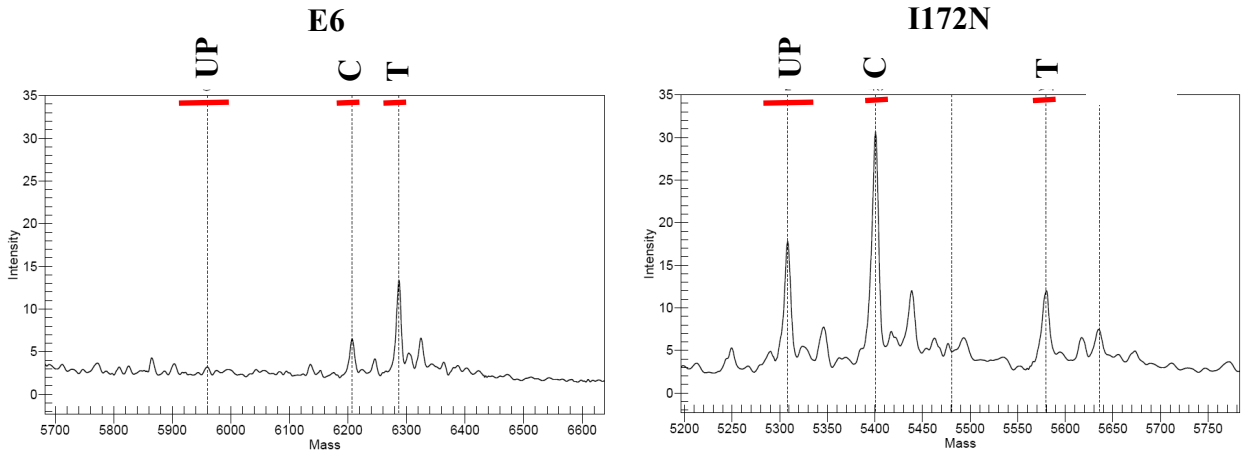


Figure 19. E6 and p.Ile172Asn iPLEX assays of sample 39: UEP peak and analyte peaks of the relevant assays are indicated with red dashes in the spectra. Unlabeled peaks belong to other assays.

Samples may sometimes present excessive background noise on the spectrum. This is an indication of inadequate resin treatment and might easily cause wrong genotypes to be called. In such cases in the experiments, the supernatant was transferred into another well where it was treated with an additional 3 mg of Clean Resin. A much better spectrum with reduced background noise was acquired after this extra conditioning step. Figure 20 shows the spectra of two samples before and after the additional step. This is a very helpful alternative to repeating the samples starting from the beginning and saves considerable amount of time and resources.

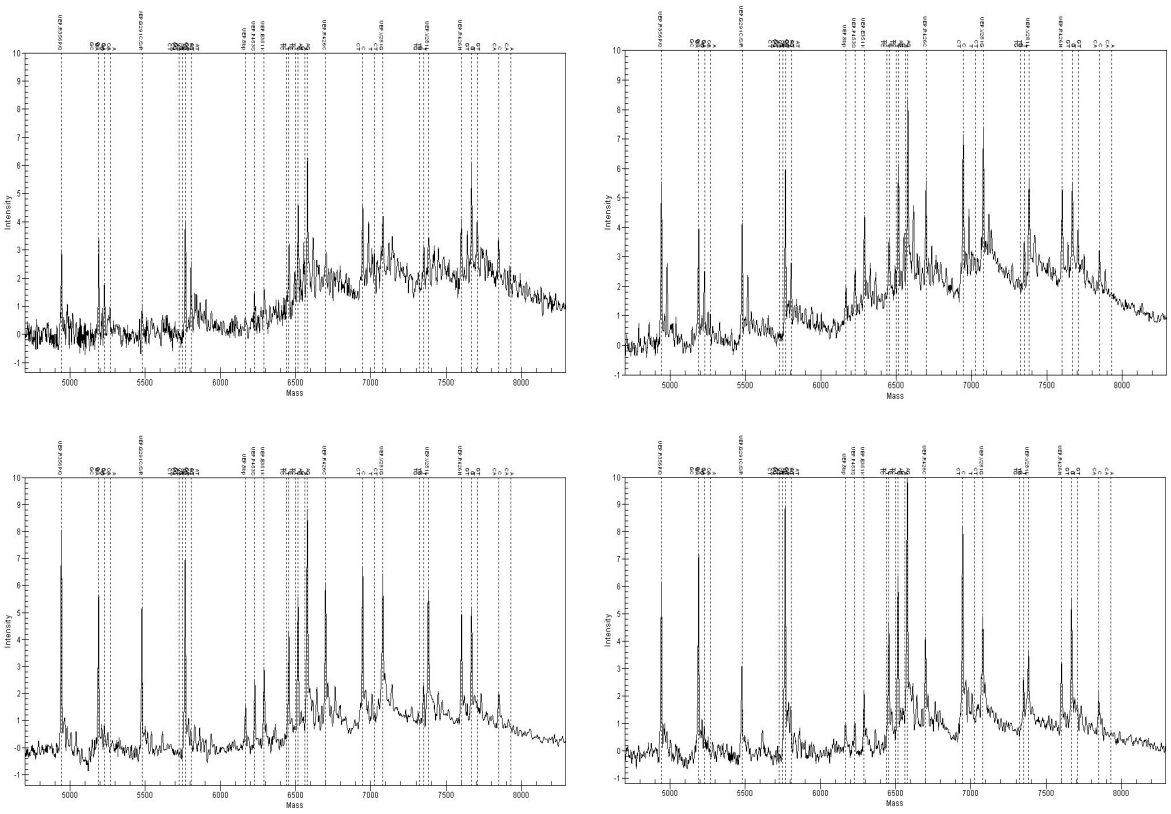


Figure 20. Two samples which resulted in a high background noise and which were treated with an extra 3 mg of Clean Resin: The upper and lower part displays the initial spectra and the spectra after the second conditioning, respectively.

In the designed mass spectrometry-based protocol for CYP21A2 analysis, a nested PCR after the specific long-range PCR is performed. This might seem to be an extra, time-wasting step when compared to standard procedures of mass cleavage and extension reactions. However, for samples with a low genomic DNA concentration or whose long-range PCR result in a low amplification yield, the improvement introduced by the nested PCR step is vital. The concentration is increased drastically in this second amplification assuring there is enough template present in the environment for the cleavage and extension reactions to acquire decent spectra. Figure 21 displays the gel electrophoretic separation after the long-range PCR products of two samples, which were included in the protocol evaluation. The amplification efficiency is very low as it is inspected visually on the gel. Despite of this fact, decent spectra were obtained for both samples after the experiments. No bias or error in the automatic calls was seen in the resulting spectra of these samples, which could have stemmed from a possible amplification from the remaining genomic DNA after the first PCR step. The absence of any CYP21A1P genotypes in the spectra

proves that the nested PCR step does not represent a risk factor for routine diagnosis. Spectra of multiplex group 1 of these samples are shown in Figure 22.

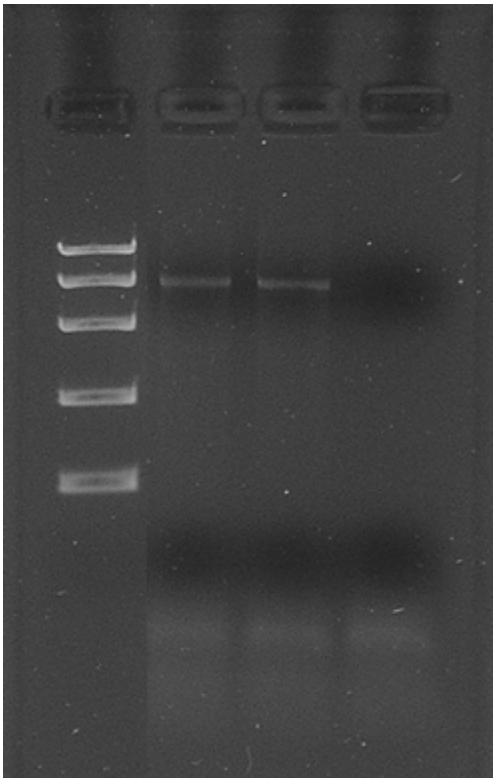


Figure 21. Gel electrophoretic separation picture obtained for the products of two samples after the long-range PCR: The two rows represent weak long range PCR amplification products of two CAH samples.

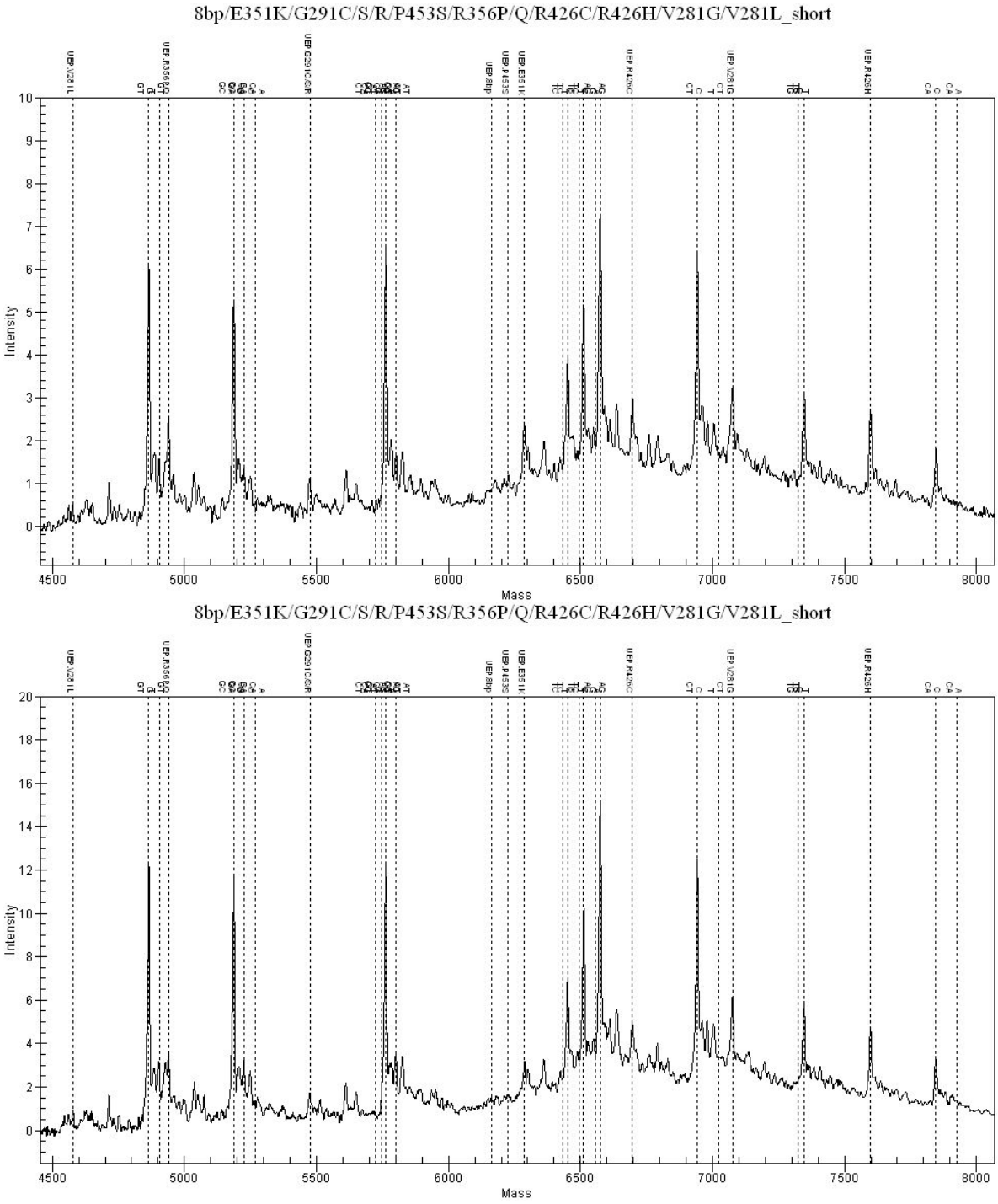


Figure 22. iPLEX spectra of multiplex group 1 of two samples whose gel electrophoresis result is presented in Figure 21.

4.1.6 Cost Analysis

MALDI-TOF MS, being a high-throughput genotyping method, is usually used in genotyping of a high number of positions in different genes and samples simultaneously. This way, costs per genotype are minimized while keeping the time for analysis and data interpretation in an acceptable range. For instance, iPLEX offers 0.05 euros per genotype in a 24-plex set-up and iPLEX Gold assay offers 0.035 euros per genotype in a 36-plex set-up. However, since what was aimed to achieve in this study is not a fast genotyping but to assure a safe molecular genetic diagnosis, the cost is calculated from the first PCR step until the final data acquisition for one sample. Laboratory, operation and initial equipment costs are not taken into account in the calculation (Table 13). The prices for all consumables are taken directly from the manufacturer's catalogue and no discount rates are applied. It was seen at the end of the calculation that a final sum of 73.33 euros per sample is needed for consumables to analyze one sample for CAH on MALDI platform for the target set of mutations and polymorphisms.

Table 13. Cost calculation per one sample for a complete CAH mutation detection analysis on MALDI-TOF MS platform.

Step	**Estimated cost per sample (euro)**
Long-range PCR	0.212
Nested PCR (hMC-PCR)	0.820
SAP Treatment + hMC reactions + Measurement	68.00
iPLEX reactions + Measurement	4.30
Total	73.33

Sanger sequencing, the method currently used for CAH analyses at Bioglobe GmbH, requires approximately 72.00 euros for one sample to provide chromatograms for forward and reverse sequencing (discount rates are taken into account). So, if the discount rates were neglected in traditional sequencing, the difference between the two protocols would be even higher. However, Sanger sequencing results in the full sequence data of the gene, which would make the discovery of any novel mutation possible. But it is less automated and requires more manual inspection. On the other hand, results from MALDI-TOF MS platform are the genotypes of certain positions which are predefined, but obtainable in a shorter time (depending on the number of

samples approximately 70-75% of conventional sequencing) and in a more automated and accurate manner. In addition, discovery of a new variant *is* possible with hMC approach provided that it generates at least one differentiating signal in one of the four cleavage reactions.

4.1.7 Compact Protocol

A predefined set of mutations and polymorphisms in CYP21A2 gene can be analyzed on MALDI-TOF MS platform combining two approaches: homogenous mass-cleave and homogenous mass-extend with single base extension. The presence of a possible novel mutation can be detected in mass-cleave reactions provided that the new variation results in a differentiating signal in at least one of the four reactions.

The designed protocol provides an extensive and automated CYP21A2 analysis. However, the investigator might prefer to analyze only the CYP21A1P-derived mutations in order to correlate the clinical phenotype or manifestations with the genotype more quickly. For this purpose, a “compact protocol” was designed, which has a lower number of targeted variations. This is a common practice and recommendation of experts for routine genetic diagnostics. A similar example is the genetic analysis of cystic fibrosis transmembrane conductance regulator (CFTR) gene, which is reported to possess over 1600 mutations. However, commonly analyzed are only 23, which cover 88% of clinical cases in non-Hispanic Caucasians [150]. Similarly, more than 95% of clinical CAH cases due to 21-OH deficiency are caused by ten specific inactivating mutations. These ten most common CYP21A1P-derived mutations and in addition seven frequent polymorphisms can be detected by the compact protocol in an automated way utilizing the iPLEX approach only.

The long-range PCR using primers shown in Figure 9 assures the specificity like in full analysis experiments. 3 kb-long product is then re-amplified in two separate uniplexes (W1 and W2) and together in one biplex (W3). Table 14 shows the nested PCR primers and their binding regions. All experimental steps are performed as described for the full analysis. Mutations 8bp-del, p.Ile172Asn, IVS2AS,A/C-G,-13, p.Pro30Leu, p.Arg356Trp and the polymorphism rs6464 are analyzed in W1; mutations E6, p.Leu306PhefsX5, p.Pro453Ser, p.Gln318X, p.Val281Leu are

analyzed in W2; polymorphisms rs28381641, rs35147842, rs61732562, rs61732563, rs6463, rs6474 are analyzed in W3.

Table 14. Primers used in nested PCR amplification of CYP21A2 in the compact kit: 10-mer = ACGTTGGATG

Name	Sequence (5' → 3')	Position
1F ⎫ W1	10mer-GGGATGGCTGGGGCTCTTG	c.1-60 - c.1-42
1R ⎭ 2188 bp	10mer-CTAAGGGCACAACGGGC	c.1084 - c.1068
2F ⎫ W2	10mer-AAGCCCACAAGAAGCTCACC	c.350 - c.369
2R ⎭ 2080 bp	10mer-GCGATCTCGCAGCACTGTGT	c.1486+81 - c.1486+100

The mutation p.Arg356Trp is included in the compact kit protocol although this position is analyzed in an additional amplicon in cleavage reactions since it is impossible to analyze using extension primers containing Inosine bases (refer to section 3.1.3 for details). The reason for this is likely to be the close proximity of Inosine bases to the extendable 3'-end, which decreases the stability of the primer binding to its target site. The solution found in this case was to combine three different extension primers with adjusted masses in order to compensate for the secondary SNP in binding region of the extension primer. Figure 23 shows the genomic region in detail. The presence of one of the rare mutations R356P and R356Q would result in a mismatch at the extendable end of a reverse primer which would bind to the downstream of p.Arg356Trp. To compensate for this, three extension primers are used in combination. R356W_wt binds when neither of two rare mutations is present (majority of the cases). R356W_g and R356W_t would bind when R356P and R356Q is present, respectively. The mass difference between the three extension primers was compensated by adding mis-matching bases to the 5'-end. The concentration of R356W_wt in the master mix is double as much as the other two primers since the rare mutations are mostly expected to be absent. Table 15 shows the sequences of the extension primers in detail for this position.

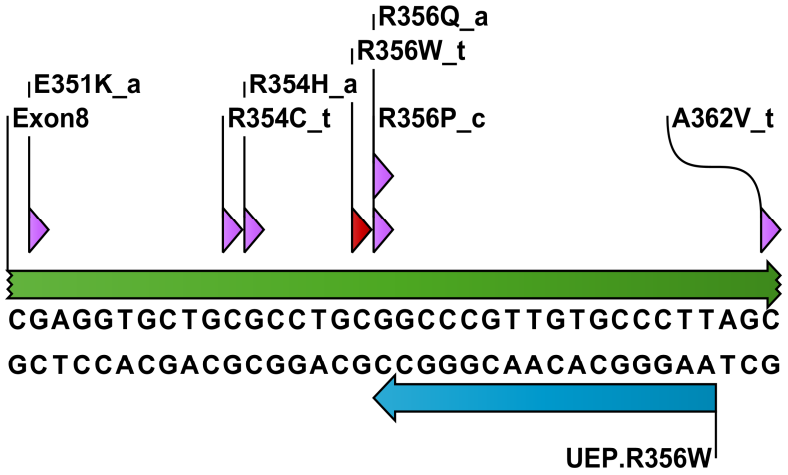


Figure 23. Genomic region surrounding the mutation p.Arg356Trp: Violet arrows represent secondary mutations. The target mutation is shown with a red arrow. The extension primer binding region is shown with a blue arrow.

Table 15. Three extension primers used in combination to detect the presence of the mutation p.Arg356Trp.

Name	**Sequence (5' → 3')**
R356W_wt	TGAAGGGCACAACGGGCC
R356W_t	CGAAGGGCACAACGGGCT
R356W_g	CTAAGGGCACAACGGGCG

Multiplex assay groups W1, W2 and W3 are presented in detail together with their masses in Table 16. Schematic representation of extension primers together with their neighboring SNPs are displayed in Table 17.

Table 16. iPLEX assays for the compact protocol represented in detail: **UEP**: Unextended primer. **rev**: Primer binding to the downstream of the variation. **fwd**: Primer binding to the upstream of the variation. **i**: Universal binding base inosine to enable annealing in the presence of secondary SNPs.

W1		
8bp (rev)		
UEP	TTGTGGGCTTTCCAGAGCAG	6164 Da
Analyte-C	TTGTGGGCTTTCCAGAGCAGG	6451.2 Da
Analyte-T	TTGTGGGCTTTCCAGAGCAGA	6435.2 Da

Table 16. cont'd

I172N (rev) (p.Ile172Asn)		
UEP	CGAAGGTGAGGTAACAG	5308.5 Da
Analyte-A	CGAAGGTGAGGTAACAGT	5635.6 Da
Analyte-T	CGAAGGTGAGGTAACAGA	5579.7 Da
I2G (fwd) (IVS2AS,A/C-G,-13)		
UEP	AGTTCCCACCCTCCAGCCCCCA	6521.24 Da
Analyte-A	AGTTCCCACCCTCCAGCCCCCAA	6792.45 Da
Analyte-C	AGTTCCCACCCTCCAGCCCCCAC	6768.42 Da
Analyte-G	AGTTCCCACCCTCCAGCCCCCAG	6808.45 Da
P30L/Q (rev) (p.Pro30Leu)		
UEP	CCGGGGCAAGAGGC	4338.8 Da
Analyte-A	CCGGGGCAAGAGGCT	4666 Da
Analyte-C	CCGGGGCAAGAGGCG	4626 Da
Analyte-T	CCGGGGCAAGAGGCA	4610 Da
R356W (rev) (p.Arg356Trp)		
UEP	TGAAGGGCACAACGGGCC	5558.62 Da
Analyte-C	TGAAGGGCACAACGGGCCG	5845.83 Da
Analyte-T	TGAAGGGCACAACGGGCCA	5829.83 Da
rs6464 (rev)		
UEP	TGAGTCAGGCCAAGCAGATAGAT	7121.65 Da
Analyte-A	TGAGTCAGGCCAAGCAGATAGATT	7448.75 Da
Analyte-C	TGAGTCAGGCCAAGCAGATAGATG	7408.86 Da

W2		
E6 (fwd)		
UEP	AGGCCATAGAGAAGAGGGA	5959.9 Da
Analyte-C	AGGCCATAGAGAAGAGGGAC	6207.1 Da
Analyte-T	AGGCCATAGAGAAGAGGGAT	6287 Da
F306+t (fwd) (p.Leu306PhefsX5)		
UEP	TCAGGGTGGTGAAiCAAAAAAA	6850.49 Da
Analyte-7T	TCAGGGTGGTGAAiCAAAAAAAC	7097.68 Da
Analyte-8T	TCAGGGTGGTGAAiCAAAAAAAA	7121.7 Da
P453S (rev) (p.Pro453Ser)		
UEP	CAGGGCGTCCCCGGAGG	5237.39 Da
Analyte-C	CAGGGCGTCCCCGGAGGG	5524.6 Da
Analyte-T	CAGGGCGTCCCCGGAGGA	5508.6 Da
Q318X (rev) (p.Gln318X)		
UEP	TGGTCTAGCTCCTCCT	4799.1 Da
Analyte-C	TGGTCTAGCTCCTCCTG	5086.3 Da
Analyte-T	TGGTCTAGCTCCTCCTA	5070.3 Da

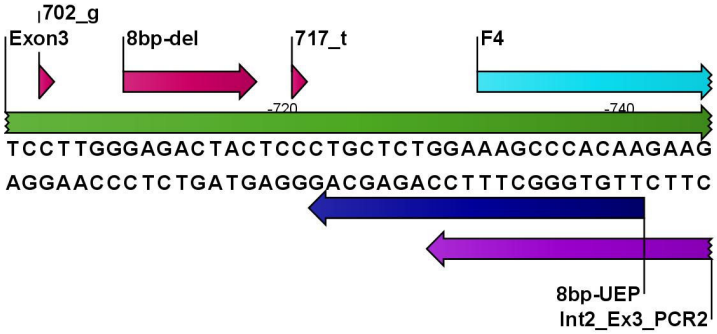
Table 16. cont'd

V281L (fwd) (p.Val281Leu)		
UEP	CTCCTGGAAGGGCAC	4577.99 Da
Analyte-G	CTCCTGGAAGGGCACG	4865.19 Da
Analyte-T	CTCCTGGAAGGGCACT	4905.08 Da
W3		
rs28381641 (rev)		
UEP	CAGGGGCAGCAGCAGCAG	5583.64 Da
Analyte-4CTG	CAGGGGCAGCAGCAGCAGG	5870.84 Da
Analyte-5CTG	CAGGGGCAGCAGCAGCAGC	5830.82 Da
rs35147842 (rev)		
UEP	GCTGGAGGGTGGGAACTGA	5973.88 Da
Analyte-CA	GCTGGAGGGTGGGAACTGAT	6300.98 Da
Analyte-GG	GCTGGAGGGTGGGAACTGAC	6221.07 Da
rs61732562 (fwd)		
UEP	GGGATGGGGGCCCACAGCCC	6168.99 Da
Analyte-A	GGGATGGGGGCCCACAGCCCA	6440.2 Da
Analyte-G	GGGATGGGGGCCCACAGCCCG	6456.2 Da
rs61732563 (rev)		
UEP	TCCTGCCCCATCACTGG	5082.31 Da
Analyte-A	TCCTGCCCCATCACTGGT	5409.4 Da
Analyte-G	TCCTGCCCCATCACTGGC	5329.49 Da
rs6463 (fwd)		
UEP	GGGGCATTTTTTCTTTCTTAAA	6721.4 Da
Analyte-A	GGGGCATTTTTTCTTTCTTAAAA	6992.6 Da
Analyte-C	GGGGCATTTTTTCTTTCTTAAAC	6968.6 Da
rs6474 (fwd)		
UEP	TCCTGCAGAiAAGCTGGTGTCTA	7064.6 Da
Analyte-A	TCCTGCAGAiAAGCTGGTGTCTAA	7335.8 Da
Analyte-G	TCCTGCAGAiAAGCTGGTGTCTAG	7351.81 Da

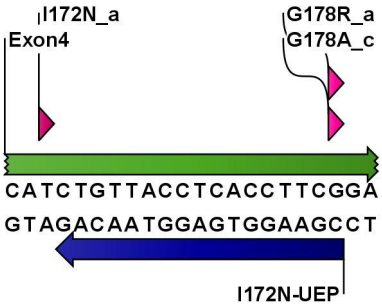
Table 17. iPLEX assays for the compact protocol represented in detail: Unextended primer and analyte masses, their sequences, direction of elongation, neighboring SNPs and extension primer binding sites are shown schematically. **UEP**: Unextended primer. **rev**: Primer binding to the downstream of the variation. **fwd**: Primer binding to the upstream of the variation. **i**: Universal binding base inosine to enable annealing in the presence of secondary SNPs.

W1

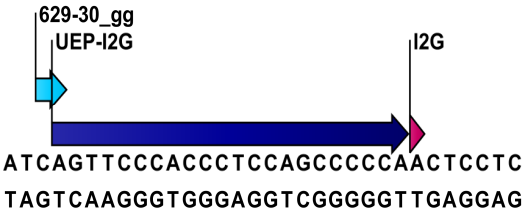
8bp (rev)



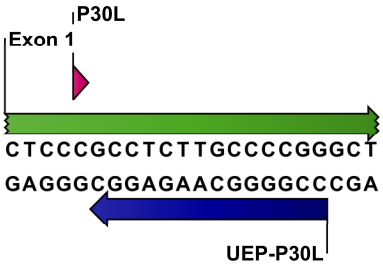
I172N (rev) (p.Ile172Asn)



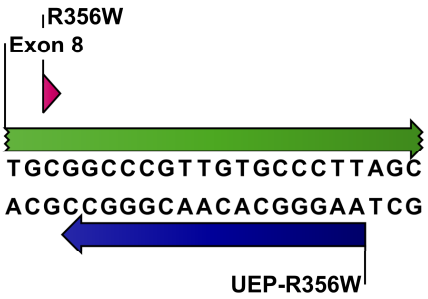
I2G (fwd) (IVS2AS,A/C-G,-13)



P30L/Q (rev) (p.Pro30Leu)



R356W (rev) (p.Arg356Trp)



rs6464 (rev)

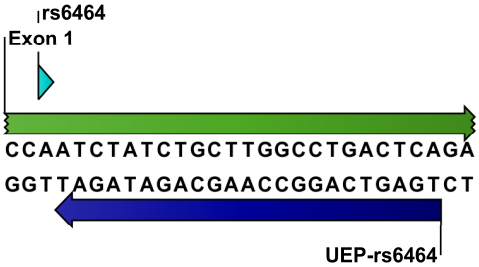
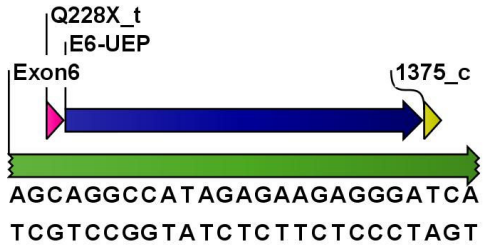


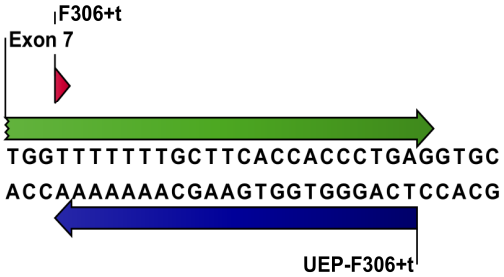
Table 17. cont'd

W2

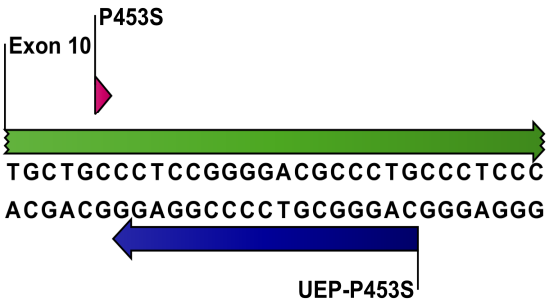
E6 (fwd)



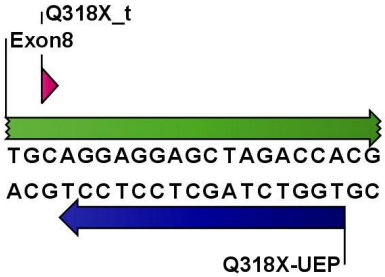
F306+t (fwd) (p.Leu306PhefsX5)



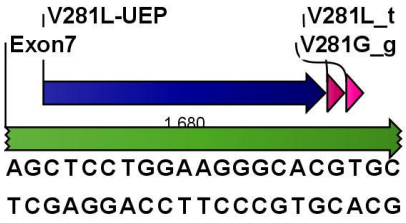
P453S (rev) (p.Pro453Ser)



Q318X (rev) (p.Gln318X)

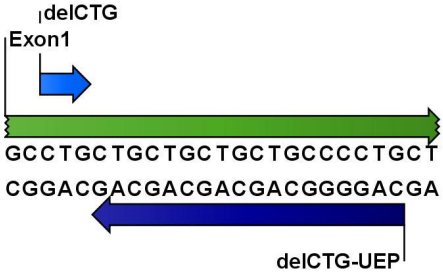


V281L (fwd) (p.Val281Leu)



W3

rs28381641 (rev)



rs35147842 (rev)

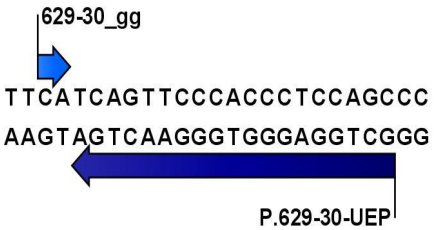
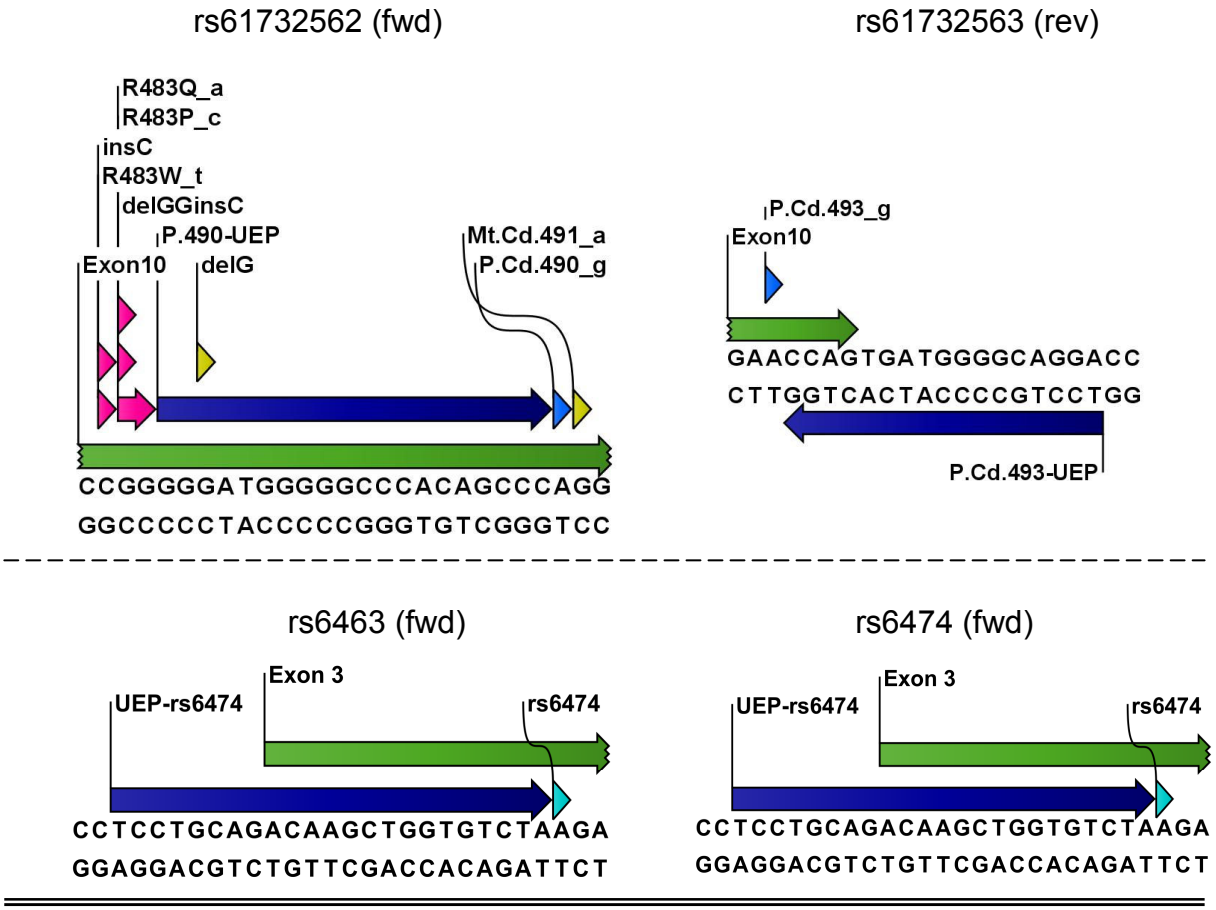


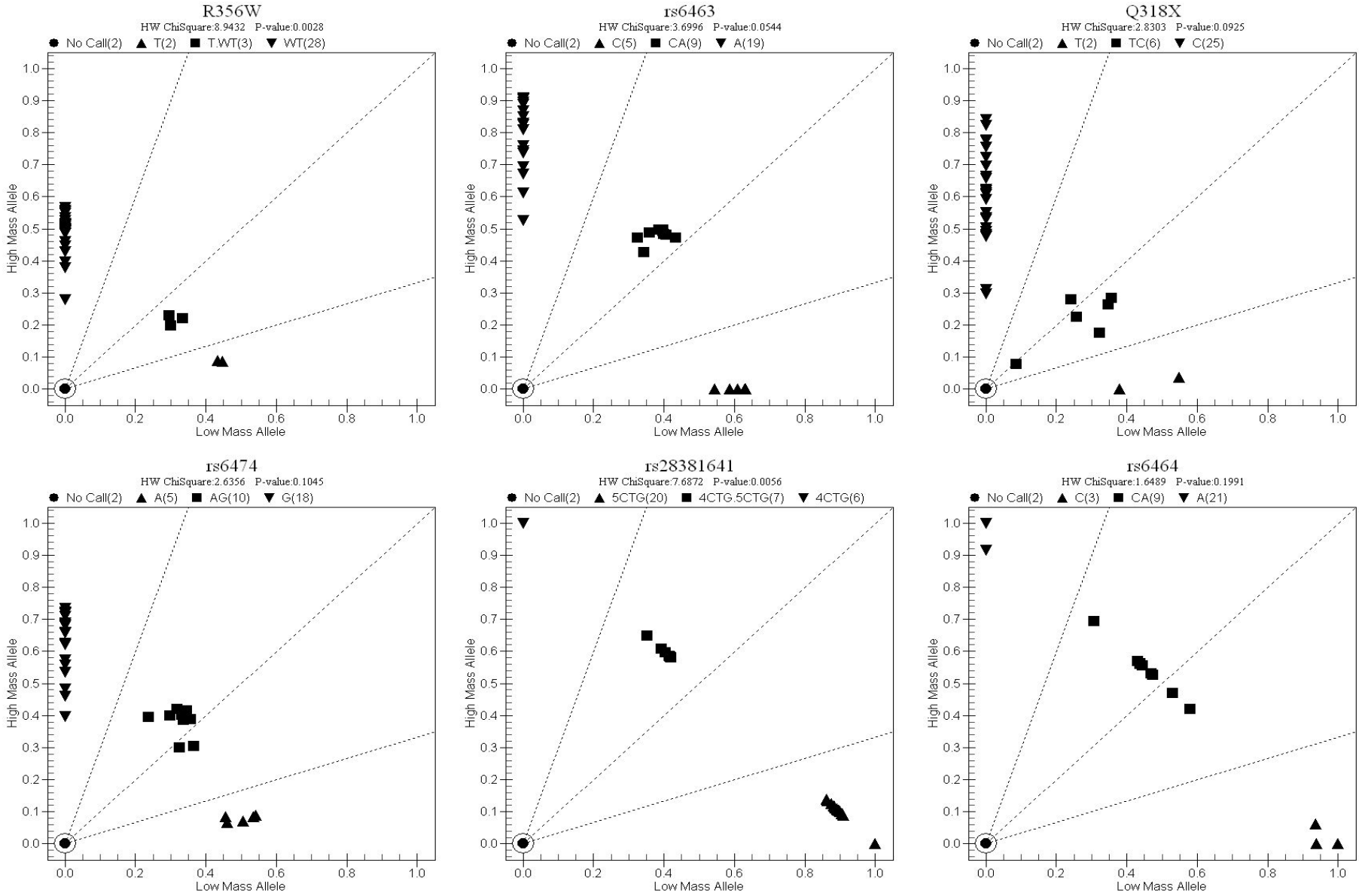
Table 17. cont'd

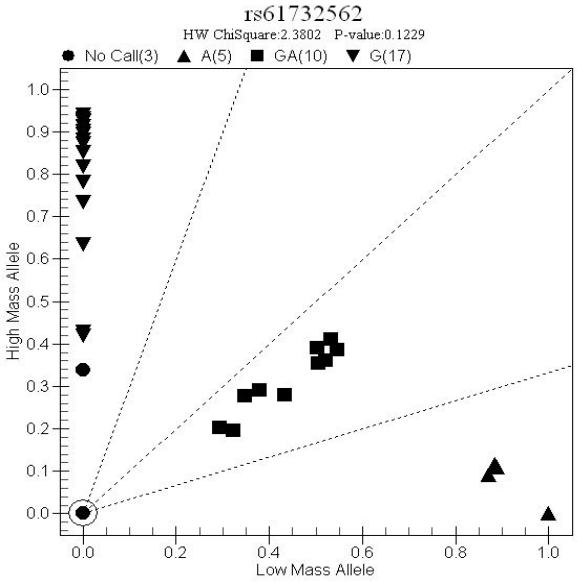


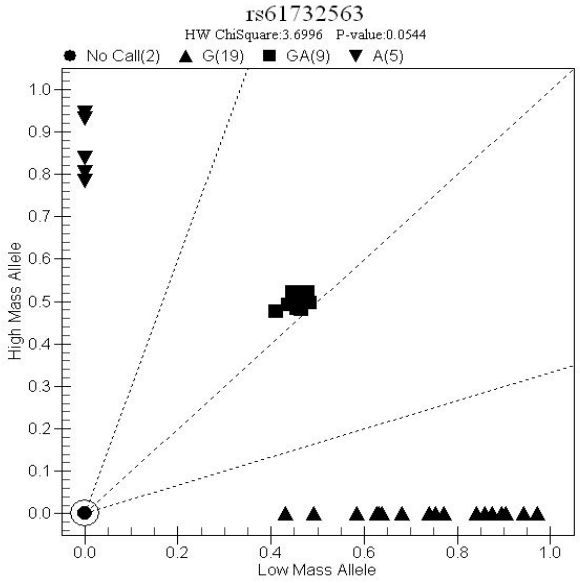
The compact protocol was evaluated using 33 samples, which contain all of the target mutations and polymorphisms in at least one heterozygous form. This guaranteed that each possible allele is represented in the spectrum range. All targeted variations were detected with yields above analytical threshold. Cluster plots of all assays in W1, W2 and W3 are presented in Figure 24. Large deviations from the 45º line in some heterozygous cases are the result of copy number variations in these samples.

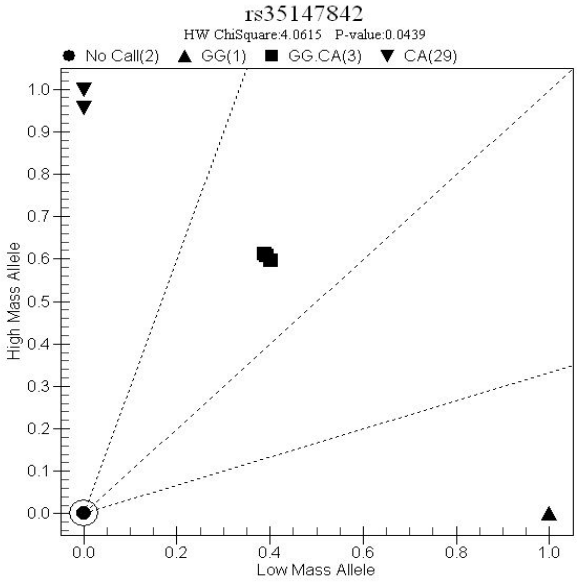
The overall stability of the compact kit was observed to be acceptable to enable a routine diagnosis. All assays in the kit proved reproducibility during independent experiments [151]. Lower yields (especially compared to the normal set-up) might be observed in cases of copy number variations. The reason for this might be the relatively longer size of the amplicons compared to the normal set-up (approximately 5-fold). During the assay optimization experiments, iPLEX assays showed a decrease in yield when the amplicon size was increased. In such cases, it might be

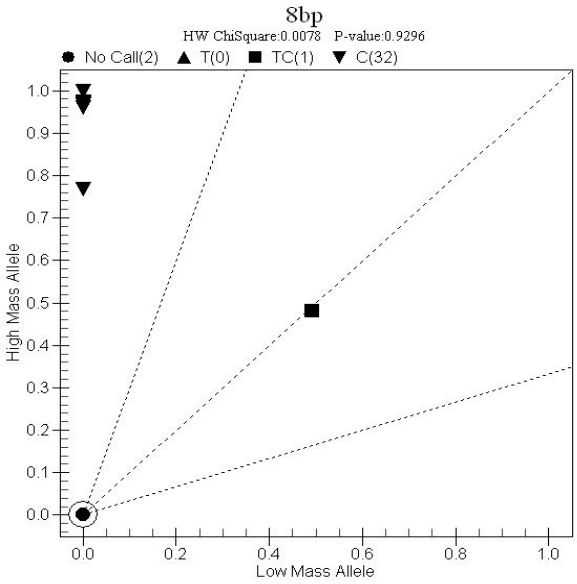
necessary to perform a repeat experiment if the yield does not allow to make a sound conclusion.

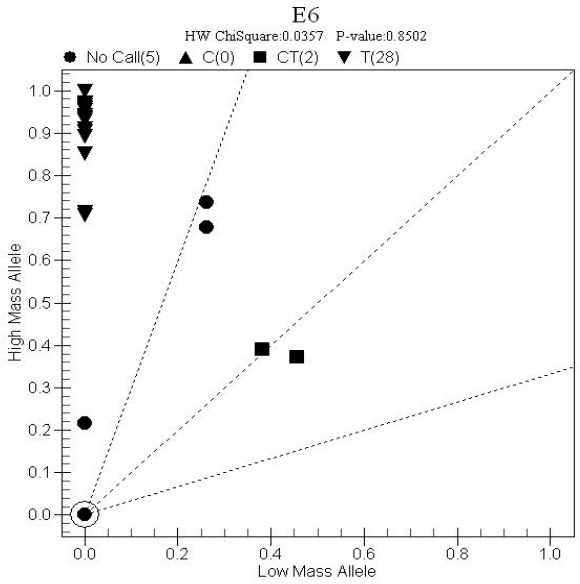

R356W
HW ChiSquare:8.9432 P-value:0.0028
No Call(2) T(2) T.WT(3) WT(28)
rs6463
HW ChiSquare:3.6996 P-value:0.0544
No Call(2) C(5) CA(9) A(19)
Q318X
HW ChiSquare:2.8303 P-value:0.0925
No Call(2) T(2) TC(6) C(25)
rs6474
HW ChiSquare:2.6356 P-value:0.1045
No Call(2) A(5) AG(10) G(18)
rs28381641
HW ChiSquare:7.6872 P-value:0.0056
No Call(2) 5CTG(20) 4CTG.5CTG(7) 4CTG(6)
rs6464
HW ChiSquare:1.6489 P-value:0.1991
No Call(2) C(3) CA(9) A(21)
High Mass Allele
Low Mass Allele

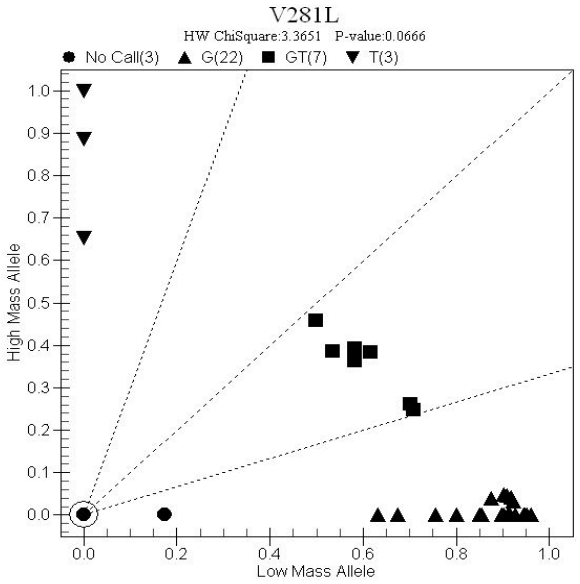

rs61732562
HW ChiSquare:2.3802 P-value:0.1229
No Call(3) A(5) GA(10) G(17)
High Mass Allele
Low Mass Allele


rs61732563
HW ChiSquare:3.6996 P-value:0.0544
No Call(2) G(19) GA(9) A(5)
High Mass Allele
Low Mass Allele


rs35147842
HW ChiSquare:4.0615 P-value:0.0439
No Call(2) GG(1) GG CA(3) CA(29)
High Mass Allele
Low Mass Allele


8bp
HW ChiSquare:0.0078 P-value:0.9296
No Call(2) T(0) TC(1) C(32)
High Mass Allele
Low Mass Allele


E6
HW ChiSquare:0.0357 P-value:0.8502
No Call(5) C(0) CT(2) T(28)
High Mass Allele
Low Mass Allele


V281L
HW ChiSquare:3.3651 P-value:0.0666
No Call(3) G(22) GT(7) T(3)
High Mass Allele
Low Mass Allele

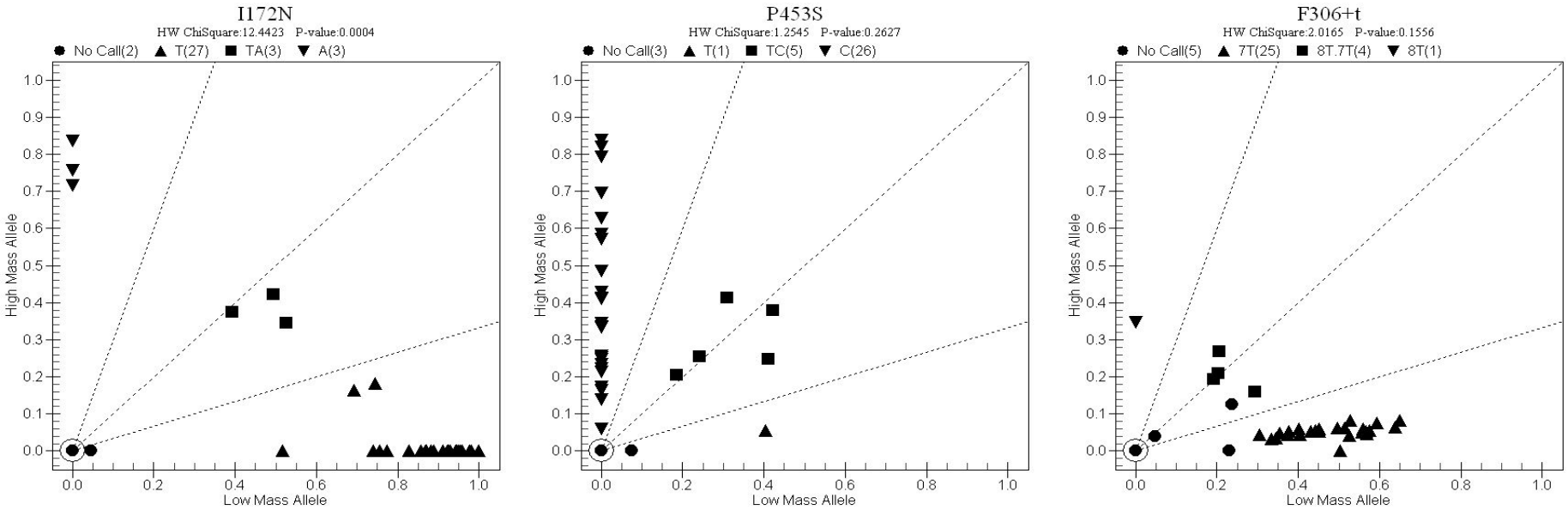


Figure 24. Cluster plots of W1, W2 and W3 of 33 test samples used in the evaluation of the compact protocol.

4.2 Copy Number Variation (CNV)

4.2.1 Copy Numbers of CYP21A2 and CYP21A1P

Both CYP21A2 and CYP21A1P are located in the major histocompatibility complex of chromosome 6p21.3, which is a complex organization of genes with variations in copy number and size. They constitute a genetic unit together with neighboring genes RP1, C4, TNXB, and their pseudogenes RP2 and TNXA. This genetic unit is named the RCCX module. It covers RP-C4-CYP21-TNX, which is a highly variable stretch of DNA with an approximate size of 30 kb [152]. Most chromosomes normally have two RCCX modules, one possessing CYP21A2 and the other CYP21A1P; however, monomodular and trimodular haplotypes have been reported as well [152].

Due to intergenic recombination events in this region of human genome, different haplotype formations are possible. These recombination events, which stem from unequal crossing over between CYP21A2 and CYP21A1P during meiosis, might include 5'-end conversion, 3'-end conversion and 30-kb deletion, which encompasses the 3'-end of CYP21A1P, all of C4B and 5'-end of CYP21A2. Regarding CYP21A2 and CYP21A1P only, the resulting genomic structures can be grouped into three different haplotypes: 1) CYP21A1P-like 5'-end, CYP21A2-like 3'-end (breakpoint between 8-bp deletion and p.Ile172Asn), 2) CYP21A1P-like 5'-end, CYP21A2-like 3'-end (breakpoint between p.Arg356Trp and p.Gln318X), 3) CYP21A2-like 5'-end, CYP21A1P-like 3'-end (breakpoint between 8-bp deletion and p.Ile172Asn) [153]. Genes formed as a result of such deletions or conversions are called fusion or hybrid genes.

4.2.2 CNV Analysis of CAH

To be able to understand the problem fully and develop an approach to tackle the copy number analysis of CYP21A2 and CYP21A1P, one has to have a deep knowledge about the “regime” of both genes. The word “regime” indicates a clear and detailed picture of conserved regions, breakpoints and mutation frequencies in both

genes. Without having this picture at hand, it is impossible to develop a protocol which would not deliver erroneous results.

In the frame of previous studies, CYP21A2 and CYP21A1P genes of 200 individuals were fully sequenced and valuable data giving hints on how to design a CNV analysis of the region using mass spectrometry were obtained [68]. Generally spoken, CYP21A1P was observed to be more conserved than CYP21A2. This means, CYP21A1P deviates less from its reference sequence formed out of 200 individuals. Breakpoints were detected to take place in exon 3 and exon 7 while single mutation transfers took place all over. CYP21A2 varied in a larger extent in the 5'-end (up to exon 3) and less in the 3'-end (from exon 3 onwards), whereas CYP21A1P showed a more conserved behavior in the 5'-end than in the 3'-end. Lastly, the direction of the transfer of variations was rather *from* CYP21A1P *to* CYP21A2 and not the opposite way. Variations with high frequencies were observed in CYP21A2 which were never found to be present in CYP21A1P [68].

CYP21A2 and CYP21A1P are identical up to 98% at exon level and to 96% at intron level. This brings one to the obligation of choosing well-characterized SNPs which distinguish between the two genes at all times and can be used in designing an error-free CNV analysis. Choosing a homozygous, that is, a non-distinguishing position would not bring more than the total copy number of CYP21A2 and CYP21A1P compared to a competitor and this would be the least informative. Therefore, combining the previous detailed research on CYP21A2 and CYP21A1P, it was tried to establish enough number of positions that would serve for this purpose. It was also considered that these positions should not all be located in the same region but distributed in order to provide an insight about the copy numbers of different exons and introns, and conversion events and breakpoints. It is useful to keep in mind during the design that the breakpoint in a gene conversion event lies either in exon 3 or exon 7.

Table 18 shows the positions selected for CNV analysis. Out of a total of three positions, the last one, c.702T>C (rs10947229), is a pre-indicator of exon 6 cluster. In CYP21A1P and CYP21A2, the genotype at this position is a “C” only when the exon 6 cluster is present, when it is not, it is a “T”. In case of a 3'-end conversion,

CYP21A1P genotypes for exon 6 cluster are detected in the formed hybrid gene. Moreover, this position is more conserved between the two genes, whereas T insertion could take place in the active gene without a hybrid gene formation.

Independent of the position of the breakpoint, analyzing CNV data together with that of genotyping helps both analyses immensely. Since the homology is extremely high between the two genes and the number of differentiating positions is limited to three, it is sometimes not possible and not vital to know the exact position if the breakpoint.

Table 18. Well characterized SNPs chosen at three positions to assign copy numbers for CYP21A2 and CYP21A1P: Figures in parenthesis show the frequency of the genotype in the specific gene.

Position	Region	Genotype in CYP21A2	Genotype in CYP21A1P
c.1-126	5'-promoter	C (1.0); T (0.0)	C (0.0); T (1.0)
c.289+138	Intron 2	A (1.0); G (0.0)	A (0.0); G (1.0)
c.702_T>C	Exon 6	T (0.996); C (0.004)	T (0.003); C (0.997)

4.2.3 SNP Allele Ratio (SAR) Analysis

Universal primers which do not differentiate between CYP21A2 and CYP21A1P must be used to analyze copy number ratios of both genes simultaneously. Due to this, one cannot use any fragments amplified during mutation detection analysis unless one would like to assign an absolute copy number for CYP21A2 only together with a competitor.

The first approach was to make use of the differentiating SNP information and try to obtain allele ratios for each position. For this purpose, extension assays and PCR primers were designed. PCR primers had the 10-mer tag in their 5'-end. The incorporation of this tag improves the overall amplification and serves for balancing the PCR efficiencies in the multiplex, which is very significant in a CNV analysis. Increasing the primer mass also assures that the unincorporated primers fall outside the mass range where the analytes are expected.

Three sites in both genes were captured in three amplicons using three non-differentiating primer pairs (Table 19). iPLEX assays were designed for these positions like in the mutation detection section (Table 20, Table 21).

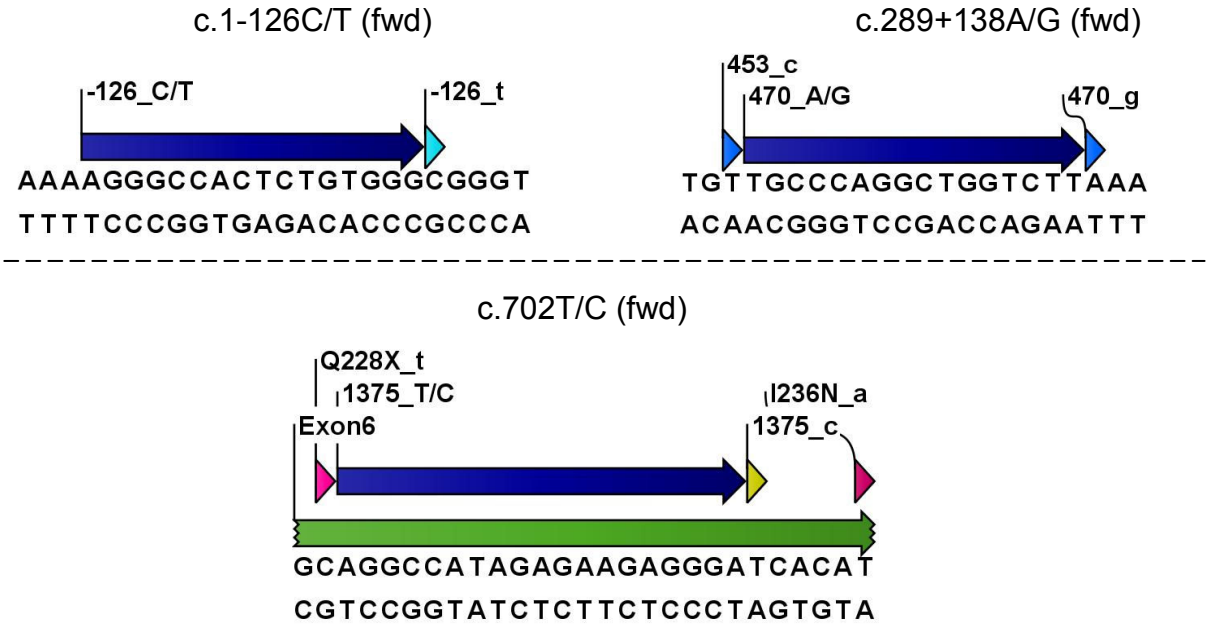
Table 19. Universal primers used to amplify the same regions in CYP21A2 and CYP21A1P in a multiplex PCR: Both forward and reverse primers are tagged with 10mer-tag. **10mer-tag** = acgttggatg

Name		Sequence (5' → 3')	Position
Ex1_PCR1	Amp1	10mer-GGCCAATGAGACTGGTGTCATTCC	c.1-171 – c.1-148
Ex1_PCR2		10mer-AGCCATCCCTCCTGCTGTGTAGACT	c.1-52 – c.1-76
Int2_Ex3_PCR1	Amp2	10mer-TGGCAGACCTGAGCCACTTACCT	c.267 – c.289
Int2_Ex3_PCR2		10mer-GGTGAGCTTCTTGTGGGCTTTCC	c.369 – c.347
Ex6_PCR1	Amp3	10mer-CATTCTCATGCTTCCTGCCGC	c.649-23 – c.649-3
Ex6_PCR2		10mer-CGAGGGGAGGCCGTCCAC	c.649+14 – c.649+31

Table 20. SAR assays for CAH: Unextended primer and analyte masses, their sequences, direction of elongation are shown. **UEP:** Unextended primer. **fwd:** Primer binding on the upstream of the variation.

iPLEX Assays for SAR Analysis		
c.1-126C/T (fwd)		
UEP	AGGGCCACTCTGTGGG	4938.2 Da
Analyte-C	AGGGCCACTCTGTGGGC	5185.4 Da
Analyte-T	AGGGCCACTCTGTGGGT	5265.3 Da
c.289+138A/G (fwd)		
UEP	TGCCCAGGCTGGTCTT	4864.16 Da
Analyte-A	TGCCCAGGCTGGTCTTA	5135.37 Da
Analyte-G	TGCCCAGGCTGGTCTTG	5151.37 Da
c.702T/C (fwd)		
UEP	AGGCCATAGAGAAGAGGGA	5959.9 Da
Analyte-T	AGGCCATAGAGAAGAGGGAT	6287 Da
Analyte-C	AGGCCATAGAGAAGAGGGAC	6207.1 Da

Table 21. SAR assays for CAH: Unextended primers together with neighboring SNPs and extension primer binding sites are shown schematically. **UEP:** Unextended primer. **fwd:** Primer binding on the upstream of the variation.



Multiplex PCR conditions had to be modified from those of used in mutation detection in order to increase data accuracy. The default hME-PCR recipe for low plex multiplexing was tested at first, which contains $MgCl_2$, 10x HotStar PCR Buffer and HotStar Taq DNA Polymerase. However, a constant drop out of analyte-A of the assay 470_A/G in some samples was observed. To overcome this problem, HotStar Taq DNA Polymerase was replaced by Platinum Taq DNA Polymerase High-Fidelity, $MgCl_2$ by $MgSO_4$, and 10x HotStar PCR Buffer by Hi-Fi Taq PCR Buffer. When the final spectrum was inspected, the allele drop-out was observed not to take place anymore with the undertaken modifications (Figure 25). Secondly, since this is a quantitative analysis, the number of PCR cycles was reduced from 45 down to 25. This assures that the reaction is terminated before it reaches plateau thus all allele ratios are correctly represented. This change was coupled with a 60% reduction in dNTPs concentration to lower the load of the SAP treatment and improve final spectrum. This percentage value was found after a titration experiment. Lastly, since Platinum Taq DNA Polymerase High-Fidelity is used, the elongation temperature was changed from 72ºC to 68ºC. SAP treatment and extension reaction recipes and programs are performed as described in the mutation detection section.

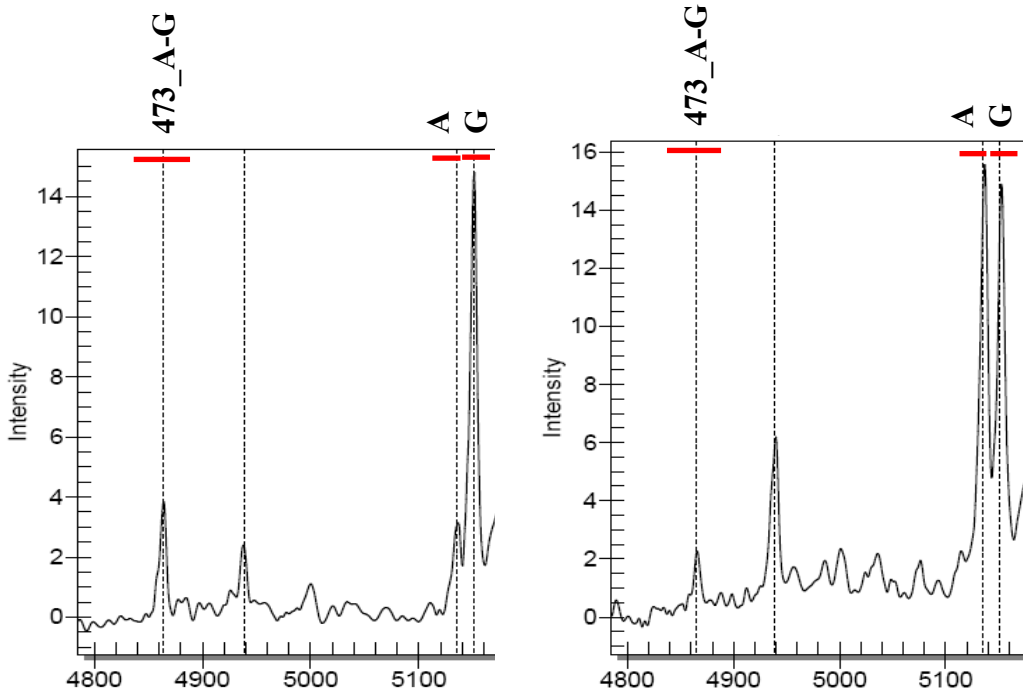


Figure 25. iPLEX spectra of a sample showing the SNP allele ratios for c.289+138A/G after performing PCR with HotStar Taq (left) and Hi-Fi Taq (right): The A-allele in the spectrum on the left could be neglected when compared to the G-allele. In the spectrum on the right, both alleles have the same intensity. Unlabeled peaks belong to other assays.

4.2.3.1 Evaluation of SAR for Routine Applications

The standard approach to evaluate the final outcome of a chromatographic experiment quantitatively is to calculate the area under the curve. This way, depending on the overall performance of experiments, it is possible to conclude on the relative proportions of different analytes in a mixture. For this type of quantitative analysis, mass spectra are treated as chromatogram peaks, wfhere the area under the curve represents the relative ratio of two analytes. Table 22 plots the SNP allele ratios of six samples obtained after iPLEX measurement. As recommended by Sequenom, each sample was dispensed on four different spots on the SpectroCHIP and the area data under the peak for each analyte resulting from all four spots were averaged.

Although the final outcome of MALDI-TOF MS is very much like a chromatography result, everything begins with the PCR step, which is a hybridization-based method.

The biggest obstacle while performing a quantitative analysis using peaks which come from different PCR products is that the PCR efficiency is very likely to be different for each amplicon. This phenomenon is not difficult to notice in Table 22. The CYP21A2 to CYP21A1P ratio shows an increase from the 5'-end to the 3'-end. This observation points out that the non-differentiating primers chosen to simultaneously amplify certain regions of both genes are prone to bind on CYP21A2 more than CYP21A1P towards the 3'-end. This general tendency was observed throughout the following measurements. Homozygous or heterozygous deletion of CYP21A1P is reported to occur with a frequency of 2 to 20% whereas duplicated CYP21A2 haplotypes have been more rarely reported [154]. This observed tendency speculates that the potential heterozygous complexity of CYP21A1P affects *in vitro* amplification as well as cellular replication.

Table 22. SAR results of six samples after iPLEX measurement.

Sample	**CYP21A2-allele/CYP21A1P-allele Ratio**		
	c.1-126_C/T	c.289+138_A/G	c.702_T/C
1	0.87	1.00	1.27
2	0.81	1.05	1.35
3	0.93	1.05	1.23
4	1.80	1.79	2.40
5	1.36	1.63	2.40
6	0.92	1.02	1.27

Data in Table 22 suggest that samples 1, 2, 3, and 6 have the same copy number of both genes whereas in samples 4 and 5 the ratio of CYP21A2 copy number to that of CYP21A1P seems to be higher, and this holds for all three regions. Assuming the interpretation of figures reflects the reality, data also show that the closest figures are given by the assay c.289+138A/G. This assay provides the closest values to a ratio of 1.0 for cases of equal number of copies of both genes.

As mentioned before, it is natural that every PCR takes place with different efficiencies in a multiplex reaction. This can be corrected by two actions: 1) Adjusting single primer pair concentrations in a multiplex reaction 2) Introducing empirical correction factors. The first option is not applicable in our case since the multiplex PCR primers bind universally to both genes in the designed set-up. Adjusting the concentration would not modify amplification efficiencies in a gene-specific manner,

but it would either increase or decrease the amplification efficiency of both genes at the same time. Therefore, to correct for the efficiency difference between three amplifications, correction factors were introduced. 11 more samples, which showed the same behavior as the original test samples, were added to the first analysis to calculate correction factors which are statistically more universal. Calculating the average deviation from the true value for each point using a total of 17 samples, values of 1.15 and 0.8 were calculated as multiplication coefficients for assays c.1-126C/T and c.702T/C, respectively.

The Excel-based software Coffalyser 8.0 is provided by MRC-Holland to analyze the final data. The software calculates the relative peak area (RPA) of a sample by dividing the area of each measured peak (As) by the sum of the area of all peaks of that sample (∑As). Then it calculates the relative peak ratio (RPR) by dividing the RPA (As/∑As) by the mean RPA of the corresponding probe obtained from control wild type DNA samples. The software finally identifies a peak as normal when its RPR value lies in a range from 0.7 to 1.3, as deleted when RPR is <0.7, and duplicated when RPR is >1.3 [155], which is adjustable by the user. Since during the calculation peaks of interest are compared to control peaks, the final result gives the absolute copy number.

At this point, it is useful to compare the results of both methods for the same samples. Correction factors are applied to MALDI-TOF MS figures and MLPA figures are obtained by GeneMarker 1.7, a program developed by Softgenetics which recognizes MLPA probe sets and analyzes the raw data of MLPA from capillary electrophoresis separation directly with a similar algorithm like that of Coffalyser (Table 23). CAH kit from MRC-Holland does not have a probe for exon 6 region of CYP21A1P but rather for exon 10. Summarized in the table are therefore figures representing the probe for exon 10 of CYP21A1P. The same is applied for intron 2 of CYP21A2, that is, Exon 3 probe data was used instead of this region.

Table 23. Comparison of copy number analysis results of MALDI-TOF MS and MLPA: Columns CYP21A2 and CYP21A1P represent MLPA results. Column SAR (MALDI-TOF MS) represents the allele ratio of CYP21A2 to CYP21A1P after the introduction of correction factors.

Sample	CYP21A2 (MLPA)			CYP21A1P (MLPA)			SAR (MALDI-TOF MS)		
	Ex1	Ex3	Ex10	Ex1	Ex3	Ex10	Ex1	Int2	Ex6
1	1.166	0.920	0.802	1.634	0.928	0.971	1.001	1.000	1.013
2	1.104	0.945	0.945	1.467	0.886	0.910	0.932	1.048	1.080
3	1.094	0.897	1.046	1.669	1.020	0.949	1.07	1.053	0.986
4	1.218	0.996	0.851	0.686	0.524	0.537	2.075	1.787	1.917
5	0.956	1.100	0.882	0.963	0.518	1.105	1.560	1.633	1.918
6	0.923	0.730	1.160	1.059	0.912	1.398	1.054	1.020	1.019

For samples 1, 2 and 3, MLPA results suggest a duplication of CYP21A1P in exon 1, and not in exon 3 and exon 10. CYP21A2 seems to be “normal” in all three positions. MALDI-TOF MS results suggest that both genes have the same number of copies at all positions, which disagrees with CYP21A1P duplication in exon 1. For sample 6, MLPA results suggest a duplication of CYP21A1P in exon 10 whereas MALDI-TOF MS gives same copy numbers throughout. According to MALDI-TOF MS results, the CYP21A2 to CYP21A1P ratio for sample 4 is 2, indicating either a homozygous duplication in CYP21A2 or a heterozygous deletion in CYP21A1P. MLPA confirms this by showing a deletion in CYP21A1P and normal behavior in CYP21A2. MALDI-TOF MS analysis shows a ratio of 1.5 for sample 5, which suggests a heterozygous duplication of CYP21A2 and normal CYP21A1P haplotype, whereas MLPA suggests a ratio of 2 only in exon 3, and 1 in exon 1 and exon 10.

In the product manual of the MLPA kit for CAH, the regions analyzed by CYP21A2-specific probes are explained. These are 3 differentiating SNPs before exon 1 for 5'-end, one of which is the selection for this region, 8-bp deletion for exon 3, p.Ile172Asn mutation for exon 4, exon 6 cluster for exon 6, and p.Gln318X mutation for exon 8. When discrepancies in results are examined more closely, it is detected that mass spectrometry method seems to deliver more correct and reasonable results. The main reason for this is the possibility of the presence of variants very close to the ligation site or to regions where probes bind in the MLPA protocol. This could well prevent or reduce the efficiency of hybridization and ligation drastically, hence lead to erroneous results [155].

Both CYP21A2 and CYP21A1P genes of 200 individuals were examined, and the frequencies of variants of all kinds in both genes were studied thoroughly. Therefore, it is easily observable that MLPA probes for exon 4 and exon 8 might not always deliver reliable results. In fact, the likelihood of obtaining false results is not low. For example, in the sample set we have studied, we observed that the possibility of CYP21A1P not to have the p.Ile172Asn mutation is 3.6%. When this is the case, then the MLPA probe for exon 4 binds to the CYP21A1P and is ligated instead of CYP21A2. It finally gives the same result as it would if there was a duplication of CYP21A2 in exon 4, which in reality is not true. The same reasoning holds for exon 8 where p.Gln318X is analyzed. The possibility of CYP21A1P not to have this mutation is even higher, namely 8.5% [68].

It is worthwhile to reconsider Table 23 under the light of this information. In the first three samples, a duplication of exon 1 is reported which is disagreed by mass spectrometry analyses. The same is valid for a deletion of exon 3 of CYP21A1P in sample 5 and a duplication of exon 10 of CYP21A2 in sample 6. Single exon deletion or duplication is a very rare event in CYP21A2 genetics, in fact, there are no cases reported in the literature yet. On the contrary, novel mutations which might hinder probe binding are continuously being reported [155]. Therefore, results obtained by mass spectrometry experiments appear to be more concordant with the genetic knowledge of the region we have so far.

Finally, it is necessary to define ranges of ratios for MALDI-TOF MS analysis to assign copy numbers. Combining the results of further measurements, it was decided that a ratio between 0.1 and 0.7 should represent a heterozygous deletion of CYP21A2 (or a heterozygous duplication of CYP21A1P), a ratio between 0.7 and 1.3 an equal number of copies, a ratio between 1.3 and 1.9 a heterozygous duplication of CYP21A2, and a ratio greater than 1.9 a homozygous duplication of CYP21A2 (or a heterozygous deletion of CYP21A1P). Infinity should indicate the homozygous deletion of CYP21A1P and 0 that of CYP21A2. It can be noticed easily, that the ranges turned out to be quite similar to that of MLPA analysis. This is not unexpected since both methods depend on the amplification and comparison of gene-specific regions. However, it should always be kept in mind that all three area ratios data should be interpreted simultaneously and not separately. Due to the method being so

sensitive in nature, error margins are not too small and ranges summarized here might sometimes shift considerably. Evaluating the outcome as a whole as much as possible decreases the chances of misinterpretation. Combining the CNV results with the results of mutation detection is very helpful for dubious samples.

4.2.4 Absolute Copy Number (ACN) Analysis

Performing an absolute copy number analysis requires the incorporation of a competitor template as a basis for comparison. Within the frame of approach, it was aimed that the sample to be analyzed be kept as isolated as possible. Therefore, similar to MLPA, the idea of using an internal control rather than an artificially synthesized oligonucleotide as a competitor was adopted. This introduces some important advantages: Firstly, one does not have to validate standard curves to determine the competitor concentration for every batch. Second, there is no depletion and re-ordering of the competitor since the sample itself is used for normalization, which saves on extra costs and labor.

The main objective to have an internal control with a constant copy number is to set a reference “intensity” of a normal gene which has a single copy on each allele. SAR analysis provides us with the basic information about copy number ratios of CYP21A2 and CYP21A1P. However, it does not differentiate between a ratio of 4:2 and 2:1. This is where the ACN analysis becomes helpful.

There is a need for a gene which has not been reported to be deleted or duplicated so far, so that it is possible to normalize any measured data against it and obtain absolute figures. The solute-carrier gene (SLC) superfamily encodes membrane-bound transporters. The SLC superfamily comprises 55 gene families having at least 362 putatively functional protein-coding genes. The gene products include passive transporters, symporters and antiporters. SLC30 is involved in Zn^{2+} efflux with ten members [156]. SLC30 member 3, SLC30A3 (also known as ZNT3), has been reported to possess no deletions or duplications so far. Moreover, it does not have a homologous pseudogene in its vicinity. SLC30A3 consists of 8521 base pairs and is located in chromosome 2, 2p23.3. It is conserved in chimpanzee, dog, cow, mouse and rat genomes.

SNPs annotated in this gene were obtained from NCBI dbSNP build 37.2. A fragment of 144 base pairs, out of a region with a relatively lower SNP density, was chosen and amplified with primers tagged with the same 10-mer tag, to equalize PCR efficiencies as much as possible. A non-polymorphic position in this fragment was chosen as a “check point” to normalize CAH data against it (Figure 26). A primer extension assay was designed for the check point with analyte masses that would fit in CNV spectrum of CAH. Since the region is free of duplications and deletions, and the check point of any variations, we expect to obtain the wild type signal with a constant intensity at all times.

Experimentally, SLC30A3 assay was added to the SAR multiplex and the conditions for PCR, SAP and iPLEX reactions were kept the same. A modification in the preparation of PCR primer mix had to be made. The concentration of CYP gene primers in the final mix had to be double the concentration of SLC30A3 primers since each CYP primer is supposed to bind and amplify two fragments, each in either CYP gene, whereas SLC30A3 primer pair is supposed to amplify just one fragment in SLC30A3 gene.

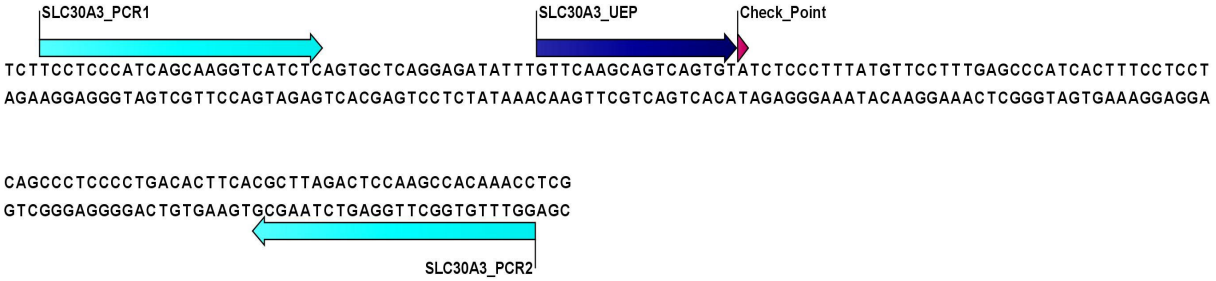


Figure 26. SLC30A3 iPLEX assay designed for ACN analysis of CAH: Primers for multiplex hME-PCR are shown with PCR1 and PCR2, and the extension primer with UEP. The artificial check point is shown with a red flag.

The same way of data interpretation in SAR was followed and results were obtained for the same 17 samples included in the previous SAR analysis. Data were compared to MLPA figures and correction coefficients were introduced. It must be mentioned that to normalize data against the internal control signal correctly is more difficult than carrying out the SAR analysis. The first assay, c.1-126C/T, required different correction factors for different analyte peaks representing CYP21A2 and CYP21A1P. After calculations, C-analyte peak has the coefficient 1.05 and T-analyte

0.9. This is not surprising as we try to achieve absolute numbers in this analysis rather than ratios. Both analytes of c.289+138A/G assay have the coefficient 0.8 since the assay always performed with a higher efficiency than SLC30A3. Assay c.702T/C did not require any correction factor. Figure 27 displays an example spectrum which shows how a typical CNV analysis display on MALDI-TOF MS is.

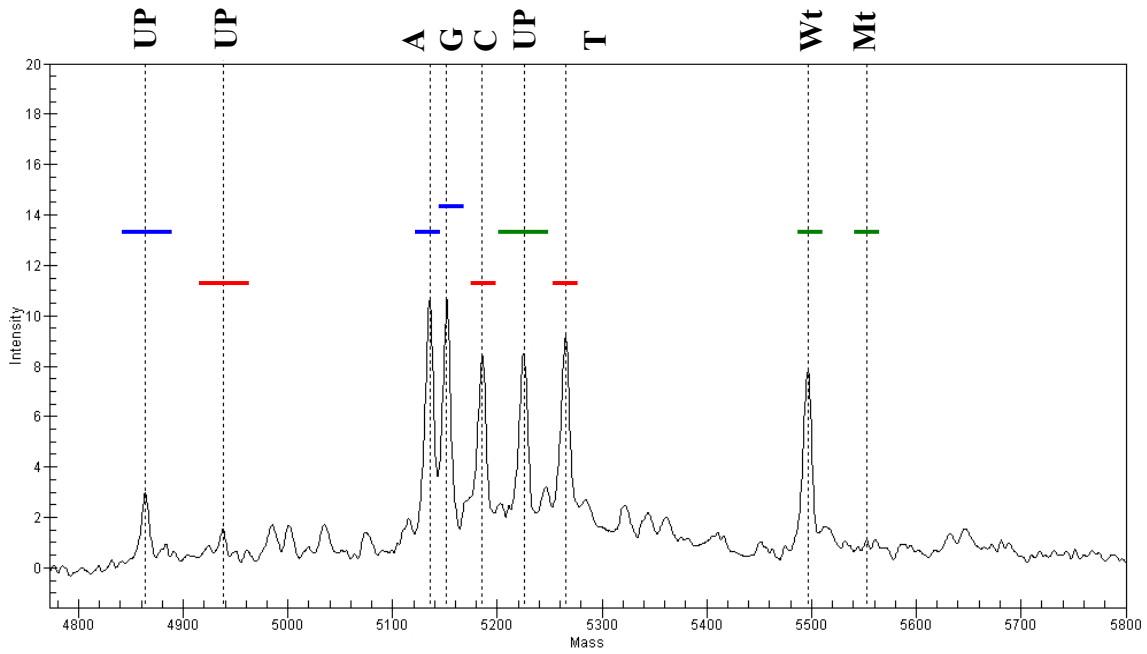


Figure 27. Example spectrum showing assays c.1-126C/T (red), c.289+138A/G (blue) and SLC30A3 (green) of ACN analysis of CAH on MALDI-TOF MS: Unextended primers and the assay analytes are labeled above belonging peaks.

Table 24 combines the results from SAR and ACN, and compares it to MLPA results. Exon 10 data is shown instead of exon 6 data of CYP21A1P from MLPA since there are no probes for exon 6 in the MLPA kit for CAH. As seen in the table, in some cases MLPA probes for CYP21A1P report inconsistent percentages. Percentage values higher than 130 for just one exon is very unlikely to represent the reality as single exon/intron duplication is a very rare phenomenon. MLPA probes for CYP21A1P sometimes result in percentages which are very close to the cut-off limit. There are two such samples in the table where the percentage data turned out to be significantly lower than others but still above the cut-off limit. In such cases, ASCN experiments delivered more consistent results; therefore, they probably represent the

true cases better. Detailed CNV analysis results of all 17 samples are summarized in section 6.2.

Table 24. SAR, ACN, MLPA results of example samples.

Sample 3837	**Ex1**	**Int2**	**Ex6/Ex10**
SAR	1.00	1.00	1.01
ACN-CYP21A2	1.18	1.16	1.19
ACN-CYP21A1P	1.17	1.16	0.95
MLPA-21A2	117%	111%	93%
MLPA-21A1P	170%	89%	99%
Sample 3840	**Ex1**	**Int2**	**Ex6/Ex10**
SAR	1.07	1.05	0.99
ACN-CYP21A2	0.91	0.84	0.98
ACN-CYP21A1P	0.84	0.80	0.80
MLPA-21A2	113%	108%	96%
MLPA-21A1P	156%	86%	105%
Sample 3841	**Ex1**	**Int2**	**Ex6/Ex10**
SAR	2.08	1.79	1.92
ACN-CYP21A2	1.14	0.92	1.24
ACN-CYP21A1P	0.54	0.52	0.52
MLPA-21A2	92%	94%	108%
MLPA-21A1P	75%	82%	68%
Sample 3843	**Ex1**	**Int2**	**Ex6/Ex10**
SAR	1.05	1.02	1.02
ACN-CYP21A2	1.22	1.08	1.10
ACN-CYP21A1P	1.14	1.06	0.86
MLPA-21A2	96%	91%	99%
MLPA-21A1P	95%	96%	154%
Sample 3850	**Ex1**	**Int2**	**Ex6/Ex10**
SAR	1.76	1.74	2.00
ACN-CYP21A2	1.45	1.27	1.53
ACN-CYP21A1P	0.81	0.73	0.61
MLPA-21A2	105%	92%	100%
MLPA-21A1P	87%	76%	81%

Samples 3837 and 3840: MLPA probe for CYP21A1P exon 1 gives an erroneous result. As discussed before, duplication in single exons/introns is very unlikely to occur. Therefore, ACN results of MALDI-TOF MS are more correct in this case.

Samples 3841 and 3850: If the data interpretation rules in MLPA manual are followed, one must assume two copies of CYP21A1P for this sample. However, ACN

results for CYP21A1P are closer to one copy than two copies. When inspected more closely, MLPA percentages are observed to be lower than the percentages of other two copy cases (>90%).

Sample 3843: ACN results present a more consistent behavior whereas MLPA data give two copies for exon 6 only.

Very similar to MLPA and any other hybridization-based quantification method, the experimenter has to “learn” how to interpret data correctly by long experimental practice. Although there are certain intervals for different numbers of gene copies, their error ranges are not narrow as shown above. One very useful indication of a copy number variation in CYP21A2 is the presence of p.Gln318X mutation. This mutation is known to co-exist with the heterozygous duplication of CYP21A2 [157]. In 13 samples with a duplication in CYP21A2, no exceptions were observed. One sample had p.Gln318X mutation and just one copy of CYP21A2, that is, in hemizygous form. In cases of the presence of this mutation in CYP21A2 either in heterozygous or homozygous form together with a copy number variation indication, the experimenter can conclude on a heterozygous deletion or duplication of CYP21A2 by combining results from genotyping with CNV experiments.

4.2.5 Cost Analysis

During the cost analysis, the list prices are used directly in the calculation without applying any discount rates. Labor, operation and initial equipment costs are not taken into account (Table 25). ASCN analysis of one sample using MALDI-TOF MS platform requires approximately 2.87 euros.

Table 25. Cost calculation of one sample for ASCN analysis on MALDI-TOF MS platform.

Step	**Estimated cost per sample (euro)**
hME-PCR	0.11
SAP Treatment + iPLEX + Measurement	2.76
Total	2.87

The list price of one SALSA MLPA KIT P050-B2 CAH from MRC-Holland costs 1100 euros and lasts for 100 reactions. So, by MLPA method one has to spend 11 euros for the analysis of one sample if one chooses *not* to include any control sample (a sample of known copy number). It costs more than 11 euros with the control sample depending on how many samples are processed in one run. When two methods are compared to each other, CNV analysis on MALDI-TOF MS saves on consumable costs considerably.

To summarize, CNV analysis of CAH using mass spectrometry as described above delivers valuable information about the region in a relatively shorter time and with less labor, especially when compared to traditional methods such as Southern blotting, fluorescent in situ hybridization, and MLPA. It has lately become even more important to investigate copy numbers of CYP genes since duplications in CYP21A2 have been reported as a risk factor for *de novo* mutations in the offspring [38].

4.3 Final Discussion

MALDI-TOF MS approaches were used in this study extensively to perform the detection of variations in CYP21A2 and quantitative analysis of CYP21A2 and CYP21A1P on mass spectrometry platform.

As the first step of mutation detection, CYP21A2 gene was amplified in a long range PCR in a single amplicon. Primers were modified with thio-phosphate bonds in order to prevent any miscorrection by the polymerase. This way, any possible allele drop-out was prevented, which might be the case when amplifying CYP21A2 in two or more overlapping fragments.

Detailed *in-silico* simulations were performed to be able to design a feasible layout for mass spectrometry experiments by taking into account CYP21A1P-derived mutations, common SNPs and rare mutations whose effect on enzyme activity is known. Results showed that 103 variations can be detected by homogenous mass cleave and 31 by iPLEX. Different from the default recipe of iPLEX, amplicons from the nested PCR with primers tagged with T7 and SP6 were used. This enabled using

the same amplicons in experiments without any extra PCR step. A total of four iPLEX assay groups enable the automatic detection of 31 variations (9+9+6+7).

The designed setup was evaluated in a large blind study which was carried out in parallel with routine diagnostics with real samples. The performance was proved to be time- and stock solution-independent. All results were obtained with high accuracy and showed reproducibility, which enabled a safe genetic diagnosis environment. The data matched %98.61 to the conventional Sanger sequencing results. It was found out that the discordance between the two platforms was due to wrong operator interpretation during the evaluation of raw sequencing data.

Two experimental kits were designed for detecting variations in CYP21A2 gene: Compact kit, which analyzes 10 major CYP21A1P-derived mutations by iPLEX approach, and extended kit, which analyzes a total of 134 variations including pseudogene-derived mutations, common polymorphisms and rare mutations by combining hMC and iPLEX methods. Both kits give the experimenter the possibility to add new clinically relevant mutations and polymorphisms as they are discovered. This way, contrary to the commercial kits, the protocol can be fully customized and kept up-to-date to meet the operator's needs.

A separate experimental layout was designed for gene copy number analysis. Three universal primer pairs, which do not differentiate between CYP21A2 and CYP21A1P, and one primer pair to create an amplicon out of SLC30A3 gene, which was used for internal normalization purposes, were multiplexed for the first PCR step. By reducing the number of cycling from 45 to 25, it was assured that the reaction is terminated before it reached equilibrium, so that both genes were represented as they are really present on the genomic level. By selecting a check point in a reference gene of the sample itself, the sample was kept in a completely isolated environment. A total of three differentiating gene specific alleles were chosen for quantification. These sites provided quantitative and relative information about the regions exon 1, intron 2 and exon 6 of both CYP genes.

After the set-up was tested with 17 samples, empiric correction factors were determined, which bring results to the true value. Since the deviation from the true

value was systematic and conserved, these factors were successfully used in the following evaluation tests and returned consistent results, which proved them to be universal.

After blind studies were completed, the results were compared to that of MLPA. It was observed that quantification on MALDI-TOF MS platform delivered more consisting results for certain regions in a more gene-specific manner. The approach has major advantages over the ligation-based quantification. First of all, it is more reliable since the selected sites are 100% differentiating. Secondly, the overnight incubation needed for ligation is eliminated, which delivers the final output much faster. The last but not the least, the operating cost is drastically lowered.

To sum up, it was shown within the frame of this study that MALDI-TOF MS can be successfully utilized to fully analyze CYP21A2 and its pseudogene CYP21A1P on genomic level. This approach increases the degree of automation, thus reducing the workload of the operator and the risk of data misinterpretation. Performing the gene copy number analysis on the same platform allows combining mutation detection and quantification on the same platform, which eliminates the necessity of obtaining two different commercial test kits for a full analysis. Finally, unlike a close box, the complete layout gives the operator full customization opportunity to meet his goals.

To our knowledge, this work represents in the up-to-date literature the first deep research project about the detection of certain mutations and polymorphisms in CYP21A2 and assignment of gene copy number of CYP21A2 and its pseudogene CYP21A1P on MALDI-TOF MS platform [151].

5 Final Conclusions and Future Prospects

CAH is among the most common recessive disorders with an average worldwide incidence rate of 1:8000 – 1:16000 for its classical form and 1:1000 for its non-classical form. Since approximately 95% of the cases are due to 21-OHase deficiency, genetic analysis focuses on the detection of mutations in CYP21A2-gene and copy number variations in CYP21A2 and CYP21A1P.

The highly rapid and accurate mass-spectrometry genotyping approach from Sequenom was implemented to achieve a full CAH due to 21-OHase deficiency analysis in this work. It was shown that the MALDI-TOF MS platform can be used for the routine analysis of complex regions with a high density of SNPs also. The novel idea of combining the MassCLEAVE technique with iPLEX without the need for an extra PCR step was successfully performed. Carrying out the PCR for hMC with primers having T7- and SP6-tags allowed to use these amplicons combined together in mass extension reactions with no adverse effect. This approach can be extrapolated to the analysis of other “difficult” regions/genes by increasing the total workload by an acceptable amount rather than doubling it.

Mass spectrometry techniques bring in a high degree of automation while keeping the confidence limit above the threshold for a diagnostic environment. A promising future prospect to make the automation even better is to improve and customize the algorithm behind. If this development was implemented both in hMC and iPLEX analysis, it would give the experimenter new opportunities. Limits regarding signal intensities in assigning a genotype could be defined at the algorithm level and specifically for each variation or variation groups. The criteria of acceptance or rejection of a spectrum could be defined using complex algorithms, which would analyze overall signal intensity and also variation signals. Spectra, which are low in signal intensity, would then be according to this pre-run algorithm either accepted or reported that it must be repeated. Moreover, defining different cut-off levels for signals belonging to different variations specifically would increase automation and decrease the manual work load. Defining best identifying reactions on a variation-specific basis would minimize the number of low score calls of complete manual inspections.

6 Appendix

6.1 Homogenous MassCLEAVE in silico Results

Amplicon 1

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c-1-4C>T(rs6470)		4484,8		1657
c.115T>C(rs6468)		2942,8		1993,2
c.126delC(P42Frameshift)	3465,2	1594		
c.135A>C(rs6464)		2725,6		5644,4
c.141delT(L48Frameshift)		4179,6		5274,2
c.185A>T(H62L)		6122,8		3367
c.185A>T(H62L)+c.191G>A(G64E)		1923,2		3367
		4235,6		
c.191G>A(G64E)		1923,2		
		1963,2		
c.27insCTG(rs28381641)				
c.3G>A(M1I)				2626,6
c.43G>A(A15T)	3425,2	4830		2009,2
c.56G>A(W19X)	1947,2		2610,6	
c.64insT(W22Frameshift)		2814,8	3621,2	
c.64insT(W22Frameshift)+c.66G>A(W22X)	2292,4	2525,6	3605,2	
c.66G>A(W22X)	2292,4	2196,4	3301	
c.82insC(H28Frameshift)	2100,4	2669,6		4738,8
c.82insC(H28Frameshift)+c.89C>A(P30Q)	2100,4	2035,2		4697,8
	2156,4	2669,6		
c.82insC(H28Frameshift)+c.89C>T(p.Pro30Leu)	2100,4	2669,6		4722,8
c.89C>A(P30Q)	2156,4	2035,2		4352,6
c.89C>T (p.Pro30Leu)				4377,6

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c-1-4C>T(rs6470)		4500,8		
c.115T>C(rs6468)	1578	2926,8		
c.126delC(P42Frameshift)	3754,4	1939,2		
c.135A>C(rs6464)		2380,4		5603,4
c.141delT(L48Frameshift)		4508,8		5603,4
c.185A>T(H62L)		2268,4		3342
c.185A>T(H62L)+c.191G>A(G64E)		3850,4		
	1650	2268,4	1968,2	3342
c.191G>A(G64E)		3850,4		
	1650	3850,4	1968,2	
c.27insCTG(rs28381641)				
c.3G>A(M1I)				1648
c.43G>A(A15T)	3441,2			
c.56G>A(W19X)	1963,2		2626,6	

c.64insT(W22Frameshift)		2485,6	3317	
c.64insT(W22Frameshift)+c.66G>A(W22X)		2485,6	3317	
c.66G>A(W22X)		2485,6	3317	
c.82insC(H28Frameshift)		2324,4		4393,6
c.82insC(H28Frameshift)+c.89C>A(P30Q)	2116,4	2324,4		4393,6
	2116,4	2324,4		4393,6
c.82insC(H28Frameshift)+c.89C>T(p.Pro30Leu)	2116,4			4393,6
c.89C>A(P30Q)	2116,4			4393,6
c.89C>T (p.Pro30Leu)	2116,4			4393,6

Amplicon 2

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.200-2A>G	2212,4		4631,8	2363,4
c.220A>T(K74X)	3256	2019,2	2651,6	2930,8
c.230T>C(I77T)	4452,8		2996,8	
c.269G>T(G90V)		3585,2		1302,8
c.289+1G>A			2955,8	4888
c.289+9T>C(rs6462)	4942	3674,4	2050,2	
c.289+9T>C(rs6462)+c.289+15C>A(rs6448)	4982	3329,2	4533,8	
c.289+15C>A(rs6448)	2709,6	3313,2	4838	
c.289+33C>A(rs6463)	2629,6			
c.289+45insTGT		4235,6		
c.289+67T>C(rs6449)		2942,8		

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.200-2A>G	2196,4	6837,2	4615,8	2667,6
c.220A>T(K74X)	4163,6		2676,6	2905,8
c.230T>C(I77T)	4163,6		3630,2	2281,4
c.269G>T(G90V)	2886,8	3545,2		
387G/A(387_G/A), c.289+1G>A			2971,8	3646,2
c.289+9T>C(rs6462)	2669,6	3658,4	2354,4	
c.289+9T>C(rs6462)+401C/A(rs6448)	2669,6	3658,4	2354,4	
			2478,6	
401C/A(rs6448)	2669,6	3658,4	2354,4	
			2478,6	
c.289+33C>A(rs6463)	2589,6			
431insTGT		3288		
c.289+67T>C(rs6449)		2926,8	2560,6	

Amplicon 3

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.289+134C>T(rs11757034)	2252,4			2980,8
c.290-129insTCC		2364,4		5635,4

c.290-109G>C	3304	4718	2050,2	
c.290-109G>C+c.290-105delG	2958,8	4428,8	1705	
c.290-105delG	3014,8	4372,8	3003,8	
c.290-85G>A	5696,4	2718,8	5027	
c.290-85G>A+c.290-84G>T	5006	2758,8	4986	
c.290-84G>T	5022	3337,2	5002	
c.209-74G>A(rs6450)	5985,6		5027	
c.209-74G>A(rs6450)+c.290-67C>A(rs6451)	6025,6		3655,2	
			5027	
c.209-74G>A(rs6450)+c.290-67C>G(rs6451)	6041,6	1827,2	3671,2	
			5027	
c.290-67C>A(rs6451)	6041,6		3655,2	
c.290-67C>G(rs6451)	6057,6	1827,2	3671,2	
c.209-48A>G(rs59064806)	1594	5849,6		2034,2
c.290-44G>T(rs6453)		3312		4304,6
c.290-44G>T(rs6453)+c.290-38_39CA>GG(rs35147842)		4219,6	1927,2	2676,6
c.290-38_39CA>GG(rs35147842)			1927,2	
c.290-13A>C(rs6467)	3931,6	4018,4		5413,2
c.290-13A>C(rs6467)+c.290-12C>T(rs6454)	3642,4	4002,4		5397,2
c.290-13A>G(IVS2AS,A/C-G,-13)	3987,6	3962,4		2354,4
				3037,8
c.290-13A>G(IVS2AS,A/C-G,-13)+c.290-12C>T(rs6454)	3698,4	3946,4		2354,4
				3021,8
c.290-12C>T(rs6454)	3682,4	3657,2		5356,2
c.290-4G>A(P.668_G/A)	3529,2	3038,8		6325,8
c.290-2A>G		4251,6		
c.290-2A>G+c.290-4G>A		3038,8		6021,6
c.290-2A>G+c.291C>A(Y97X)	3601,2	4251,6	2676,6	1541
c.290-2A>G+c.290-2A>G+c.291C>A(Y97X)+c.290-4G>A	3585,2	3038,8	2676,6	1541
				6021,6
c.291C>A(Y97X)	3585,2		2660,6	1541
c.291C>A(Y97X)+c.290-4G>A	3569,2	3038,8	2660,6	1541
				6325,8
c.305A>G(rs6474)	2292,4		2306,4	1632
c.314C>T(P105L)		1923,2		2930,8
c.314C>T(P105L)+c.315G>C(P.Cd.105_G/C)				3276
c.315G>C(P.Cd.105_G/C)	3064	1995,2		3292
c.321G>T(rs6456)		4219,6		3655,2
c.324C>G		4123,6	3317	
c.324C>G+c.329_336delGAGACTAC(8bp-del)				
c.329_336delGAGACTAC(8bp-del)			1927,2	
c.339C>T		4564,8		3630,2
c.370C>T(M.Cd.124_C>T)		2653,6		2667,6
c.370C>T(M.Cd.124_C>T)+c.371G>A(R124H)				4295,6
c.371G>A(R124H)	2429,6			4311,6
c.416T>A(V139E)	3962,4	1522	3687,2	
c.439T>C(C147R)		2870,8	3991,4	

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.289+134C>T(rs11757034)	2830,8	3505,2		2996,8
c.290-129insTCC				4615,8
c.290-109G>C	3360	4662	3349	
c.290-109G>C+c.290-105delG	3360	4662	3349	
c.290-105delG	3360	4662	3349	
c.290-85G>A	5712,4	3297,2	5043	
c.290-85G>A+c.290-84G>T	5712,4	3297,2	5043	
c.290-84G>T	5712,4	3297,2	5043	
c.209-74G>A(rs6450)	6001,6	3297,2	5043	
c.209-74G>A(rs6450)+c.290-67C>A(rs6451)	6001,6	1883,2	3326	
		3297,2	5043	
c.209-74G>A(rs6450)+c.290-67C>G(rs6451)	6001,6	1883,2	3326	
		3297,2	5043	
c.290-67C>A(rs6451)	6001,6	1883,2	3326	
c.290-67C>G(rs6451)	6001,6	1883,2	3326	
c.209-48A>G(rs59064806)		2613,6		2338,4
		3272		
c.290-44G>T(rs6453)		3272		3301
c.290-44G>T(rs6453)+c.290-38_39CA>GG(rs35147842)		3272		3301
c.290-38_39CA>GG(rs35147842)		3272		3301
c.290-13A>C(rs6467)	3971,6	3673,2		5372,2
c.290-13A>C(rs6467)+c.290-12C>T(rs6454)	3971,6	3673,2		5372,2
c.290-13A>G(IVS2AS,A/C-G,-13)	3971,6	3673,2		5372,2
c.290-13A>G(IVS2AS,A/C-G,-13)+c.290-12C>T(rs6454)	3971,6	3673,2		5372,2
c.290-12C>T(rs6454)	3971,6	3673,2		5372,2
c.290-4G>A(P.668_G/A)	3545,2	3673,2		5372,2
c.290-2A>G	3545,2	3673,2		
c.290-2A>G+c.290-4G>A		3673,2		5372,2
c.290-2A>G+c.291C>A(Y97X)	3545,2	3673,2	1327,8	1582
c.290-2A>G+c.290-2A>G+c.291C>A(Y97X)+c.290-4G>A	3545,2	3673,2	1327,8	1582
				5372,2
c.291C>A(Y97X)	3545,2		1327,8	1582
c.291C>A(Y97X)+c.290-4G>A	3545,2	3673,2	1327,8	1582
				5372,2
c.305A>G(rs6474)	2276,4		2290,4	1936,2
c.314C>T(P105L)	3120	1939,2		2946,8
c.314C>T(P105L)+c.315G>C(P.Cd.105_G/C)	3120	1939,2		2946,8
c.315G>C(P.Cd.105_G/C)	3120	1939,2		2946,8
c.321G>T(rs6456)		4179,6		
c.324C>G		4179,6	2971,8	
c.324C>G+c.329_336delGAGACTAC(8bp-del)	2653,6	4179,6	2971,8	3646,2
		4580,8		
c.329_336delGAGACTAC(8bp-del)	2653,6	4179,6	2971,8	3646,2
		4580,8		
c.339C>T		4580,8		3646,2
c.370C>T(M.Cd.124_C>T)	2445,6	2669,6		2683,6
c.370C>T(M.Cd.124_C>T)+c.371G>A(R124H)	2445,6	2669,6		2683,6

c.371G>A(R124H)	2445,6	2669,6		2683,6
c.416T>A(V139E)		2429,6	3662,2	
c.439T>C(C147R)		2854,8	4295,6	1657

Amplicon 4

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.532G>A(G178R)	3914,4	2019,2	2002,2	4954
c.532G>A(G178R)+c.533G>C(rs72552751)	3858,4	2019,2	1327,8	5603,4
c.532G>C(G178A)	3874,4	2653,6	1673	4995
c.532G>C(G178A)+c.533G>C(rs72552751)	3818,4	2709,6	1327,8	5644,4
c.533G>C(rs72552751)	3874,4	2653,6	1327,8	
c.370C>T(Mt.Cd.124_C/T)	3962,4	1522	3687,2	2667,6
		2653,6		
c.371G>A(R124H)	2429,6	1522	3687,2	4311,6
	3962,4			
c.416T>A(V139E)	3962,4	1522	3687,2	
c.439T>C(C147R)		2870,8	3991,4	
c.444+38C>T(rs6466)		4291,6		
c.444+38C>T(rs6466)+c.444+39G>A(rs58693631)		1963,2		4970
		2364,4		
c.444+39G>A(rs58693631)		2364,4		4986
c.497T>C(L166P)	1482	4387,6		6318,8
c.505_506delTGinsA(C169Frameshift)+c.508insA	3762,4		1327,8	
c.505_506delTGinsA(C169Frameshift)	3433,2			6927,2
c.505T>C(C169R)	3738,4	1995,2		6318,8
c.505T>C(C169R)+c.508insA(S170Frameshift)	4067,6	1995,2	1327,8	6318,8
c.508insA(S170Frameshift)	2581,6		1327,8	
c.512T>A(I171N)	3200			2930,8
c.512T>A(I171N)+c.515T>A(p.Ile172Asn)	3818,4	1947,2		2905,8
c.515T>A(p.Ile172Asn)	1562	1947,2		2930,8

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.532G>A(G178R)	3930,4	2597,6	2018,2	4649,8
c.532G>A(G178R)+c.533G>C(rs72552751)	3930,4	2597,6	2018,2	4649,8
c.532G>C(G178A)	3930,4	2597,6	2018,2	4649,8
c.532G>C(G178A)+c.533G>C(rs72552751)	3930,4	2597,6	2018,2	4649,8
c.533G>C(rs72552751)	3930,4	2597,6	2018,2	
c.370C>T(Mt.Cd.124_C/T)	2445,6	2429,6	3662,2	
		2669,6		
c.371G>A(R124H)	2445,6	2429,6	3662,2	1648
		2669,6		
c.416T>A(V139E)		2429,6	3662,2	

c.439T>C(C147R)		2854,8	4295,6	1657
c.444+38C>T(rs6466)	1538	4307,6		
c.444+38C>T(rs6466)+c.444+39G>A(rs58693631)	1538	4307,6		3317
c.444+39G>A(rs58693631)	1538	4307,6		3317
c.497T>C(L166P)		4371,6		6302,8
c.505_506delTGinsA(C169Frameshift)+c.508insA	2252,4	1979,2		6302,8
c.505_506delTGinsA(C169Frameshift)	2252,4	1979,2		6302,8
c.505T>C(C169R)	2252,4	1979,2		6302,8
c.505T>C(C169R)+c.508insA(S170Frameshift)	2252,4	1979,2		6302,8
c.508insA(S170Frameshift)	2252,4			
c.512T>A(I171N)	2252,4			2955,8
c.512T>A(I171N)+c.515T>A(p.Ile172Asn)	2252,4	2276,4		2955,8
c.515T>A(p.Ile172Asn)		2276,4		2955,8

Amplicon 5

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.547-15C>A(rs1040312)		1690		5637,4
		4259,6		
c.547-15C>A(rs1040312)+c.547-8T>C(rs1040311)		1690		5653,4
		4275,6		
c.547-8T>C(rs1040311)		5985,6		5694,4
c.549_C>G(rs1040310)			2692,6	
c.549delC(D183Frameshift)	4564,8		2347,4	
c.587_589delAGG(E196del)			4247,6	
c.587_589delAGG(E196del)+c.594A>T(M.Cd.198)	2220,4	2509,6	4222,6	3548,2
c.594A>T(M.Cd.198)	2220,4	2509,6	5242,2	3548,2
c.636insT(P213Frameshift)		4355,6	2560,6	4963
c.648+30G>A	1344,8	1602	6309,8	
		2613,6		
c.648+30G>A+c.648+35A>G(rs12525076)	1344,8	2613,6		
	4034,4	3048		
c.648+35A>G(rs12525076)	4034,4	5625,6	6341,8	
c.676C>T(rs6457)	2613,6	1851,2	3991,4	
c.676C>T(rs6457)+c.682C>T(Q228X)	1674	1851,2	5315,2	
	2252,4			
	2613,6			
c.682C>T(Q228X)	1674		3301	
	2252,4			
c.702T>C(rs10947229)	5897,6		5004	2946,8
c.702T>C(rs10947229)+c.707T>A(I236N)	6861,2		5004	2921,8
c.702T>C(rs10947229)+c.707T>A(I236N)+c.708C>G	6917,2		4320,6	2576,6
			5004	
c.702T>C(rs10947229)+c.707T>A(I236N)+c.708C>G+c.710T>A(V237E)			4345,6	2576,6
			5004	
c.702T>C(rs10947229)+c.707T>A(I236N)+c.708C>			4370,6	2576,6

G+c.710T>A(V237E)+c.716T>A(M239K)			5004	
c.702T>C(rs10947229)+c.707T>A(I236N)+c.710T>A(V237E)			3342	2921,8
			5004	
c.702_T>C(rs10947229)+c.707T>A(I236N)+c.710T>A(V237E)+c.716T>A(M239K)			3367	2921,8
			5004	
c.707T>A(I236N)	2525,6			2905,8
c.708C>G		2140,4	4295,6	2585,6
c.710T>A(V237E)	2982,8		3342	
c.710T>A(V237E)+c.716T>A(M239K)			3367	
c.716T>A(M239K)	3946,4		3342	
c.721C>G(M.Cd.241)		1522	3012,8	
c.735+12_13AC>GT(rs71552100)		3433,2	3596,2	

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.547-15C>A(rs1040312)		5969,6		5678,4
c.547-15C>A(rs1040312)+c.547-8T>C(rs1040311)		5969,6	1541	5678,4
c.547-8T>C(rs1040311)		5969,6	1541	5678,4
c.549C>G(rs1040310)	4854		1673	
c.549delC(D183Frameshift)	4854		1673	
c.587_589delAGG(E196del)	2942,8		5267,2	
c.587_589delAGG(E196del)+c.594A>T(M.Cd.198)	2549,6		5267,2	
	2942,8			
c.594A>T(M.Cd.198)	2549,6		5267,2	
c.636insT(P213Frameshift)		4026,4	2256,4	4633,8
c.648+30G>A	1360,8	4179,6	6325,8	
c.648+30G>A+c.648+35A>G(rs12525076)	1360,8	1482		
	4018,4	4179,6		
c.648+35A>G(rs12525076)	4018,4	1482	6325,8	
		4179,6		
c.676C>T(rs6457)			1977,2	
			2034,2	
c.676C>T(rs6457)+c.682C>T(Q228X)	3890,4		1977,2	
			2034,2	
c.682C>T(Q228X)	3890,4		1977,2	
c.702T>C(rs10947229)	1562		5308,2	2930,8
	4371,6			
c.702T>C(rs10947229)+c.707T>A(I236N)	1562	2196,4	5308,2	2930,8
	4371,6			
c.702T>C(rs10947229)+c.707T>A(I236N)+c.708C>G	1562	2196,4	3317	2930,8
	4371,6		5308,2	
c.702T>C(rs10947229)+c.707T>A(I236N)+c.708C>G+c.710T>A(V237E)	1562	2196,4	3317	2930,8
	2019,2		5308,2	
	4371,6			
c.702T>C(rs10947229)+c.707T>A(I236N)+c.708C>G+c.710T>A(V237E)+c.716T>A(M239K)	1562	2196,4	3317	2930,8
	2019,2		5308,2	
	4371,6			
c.702T>C(rs10947229)+c.707T>A(I236N)+c.710T>A(V237E)	1562	2196,4	3317	2930,8
	2019,2		5308,2	
	4371,6			

c.702T>C(rs10947229)+c.707T>A(I236N)+c.710T>A(V237E)+c.716T>A(M239K)	1562	2196,4	3317	2930,8
	2019,2		5308,2	
	4371,6			
c.707T>A(I236N)	1562	2196,4		2930,8
c.708C>G		2196,4	3317	2930,8
c.710T>A(V237E)	2019,2	2196,4	3317	
c.710T>A(V237E)+c.716T>A(M239K)	2019,2	2196,4	3317	
c.716T>A(M239K)	2019,2		3317	
c.721C>G(M.Cd.241)				
c.735+12_13AC>GT(rs71552100)				

Amplicon 6

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.735+12_13AC>GT(rs71552100)		3433,2	3596,2	
c.736-74G>A	2469,6		3260	4888
c.736-21C>T(rs6465)		4620,8		
c.737delA(E246Frameshift)			2363,4	
c.744C>G(rs6477)		3465,2	2313,4	
c.782T>C(L261P)	3866,4	2324,4		
c.782T>C(L261P)+c.784C>T(Q262X)	1538		3662,2	
	2364,4			
c.784C>T(Q262X)	2364,4		3662,2	2281,4
c.803G>C(rs6472)	4774			
c.819T>C(rs55695018)	3176	3136		
c.841G>T(p.Val281Leu)	3617,2	1963,2		1977,2
c.841G>T(p.Val281Leu)+c.842T>G(V281G)	2252,4			
	3617,2			
c.842T>G(V281G)	5889,6	1883,2	1359,8	
c.842T>G(V281G)+c.847A>C(M283L)	5849,6	1883,2	1359,8	
		3232		
c.847A>C(M283L)		3232		
c.871G>A(G291S)				
c.871G>C(G291R)				
c.871G>T(G291C)		2870,8		
c.884C>A(T295N)	6348		2306,4	2231,4
c.898C>T(L300F)	6018,8			5274,2
c.898C>T(L300F)+c.903C>A(S301Y)	6018,8	2693,6		5233,2
c.898C>T(L300F)+c.903C>A(S301Y)+c.904T>C(W302R)	3192	2693,6	1359,8	5249,2
	6018,8			
c.898C>T(L300F)+c.904T>C(W302R)	3152		1359,8	
	6018,8			
c.903C>A(S301Y)		2709,6		5249,2
c.903C>A(S301Y)+c.904T>C(W302R)	3192	2709,6	1359,8	5265,2
c.904T>C(W302R)	3152		1359,8	5306,2
c.904T>C(W302R)+c.906G>A(W302X)	3136	3689,2		5306,2
		6444		
c.904T>C(W302R)+c.906G>A(W302X)+c.910G>A(V304M)	3120	3689,2	4122,6	5306,2
		5175,2		
c.904T>C(W302R)+c.910G>A(V304M)	3136	4958	1359,8	5306,2

		5175,2	4122,6	
c.906G>A(W302X)	2268,4	3673,2		
		6444		
c.906G>A(W302X)+c.910G>A(V304M)	2252,4	3673,2	4122,6	
		5175,2		
c.910G>A(V304M)	2268,4	4942	4122,6	
		5175,2		
c.920dupT(p.Leu306PhefsX5)			4442,8	2957,8
c.922C>T(L308F)			5051,2	4320,6
c.936+1G>C		4718		
c.936+1G>C+c.936+2T>G	2958,8	4678		
c.936+2T>G	3014,8	4622	2683,6	
c.936+11G>C(rs6442)	5857,6	4718		
c.936+11G>C(rs6442)+c.936+12G>A	5841,6	4139,6		
c.936+12G>A	5897,6	4083,6		

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.735+12_13AC>GT(rs71552100)		1867,2	1993,2	
c.736-74G>A	2485,6		3276	1616
				3292
c.736-21C>T(rs6465)	2790,8	4636,8		1359,8
c.737delA(E246Frameshift)	3906,4		2692,6	
c.744C>G(rs6477)		3521,2		1689
c.782T>C(L261P)	2942,8	2308,4		
c.782T>C(L261P)+c.784C>T(Q262X)	2942,8		3358	
c.784C>T(Q262X)	2942,8	2308,4	3358	
c.803G>C(rs6472)	4830			
c.819T>C(rs55695018)	2597,6	3120		
c.841G>T(p.Val281Leu)	3962,4	1923,2		
c.841G>T(p.Val281Leu)+c.842T>G(V281G)	3962,4			
c.842T>G(V281G)	3962,4	1923,2		
c.842T>G(V281G)+c.847A>C(M283L)	3962,4	1923,2		1952,2
		2541,6		
c.847A>C(M283L)		2541,6		1952,2
c.871G>A(G291S)		2830,8	2009,2	
c.871G>C(G291R)		2830,8	2009,2	
c.871G>T(G291C)		2830,8	2009,2	
c.884C>A(T295N)	6308		1977,2	2272,4
c.898C>T(L300F)	6308			5290,2
c.898C>T(L300F)+c.903C>A(S301Y)	6308			5290,2
c.898C>T(L300F)+c.903C>A(S301Y)+c.904T>C(W302R)	2284,4			5290,2
	6308			
c.898C>T(L300F)+c.904T>C(W302R)	2284,4			
	6308			
c.903C>A(S301Y)				5290,2
c.903C>A(S301Y)+c.904T>C(W302R)	2284,4			5290,2
c.904T>C(W302R)	2284,4			5290,2
c.904T>C(W302R)+c.906G>A(W302X)	2284,4			5290,2
c.904T>C(W302R)+c.906G>A(W302X)+c.910G>A(V304M)	2284,4		4138,6	5290,2

c.898C>T(L300F)+c.904T>C(W302R)	2284,4		4138,6	5290,2
c.906G>A(W302X)	2284,4			
c.906G>A(W302X)+c.910G>A(V304M)	2284,4		4138,6	
c.910G>A(V304M)	2284,4		4138,6	
c.920dupT(p.Leu306PhefsX5)			4138,6	2628,6
c.922C>T(L308F)			4138,6	4336,6
c.936+1G>C	1690	4662	2642,6	
c.936+1G>C+c.936+2T>G	1690	4662	2642,6	
c.936+2T>G	1690	4662	2642,6	
c.936+11G>C(rs6442)	5913,6	4662	2338,4	
c.936+11G>C(rs6442)+c.936+12G>A	5913,6	4662	2338,4	
c.936+12G>A	5913,6	4662	2338,4	

Amplicon 7

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.936+1G>C	1634	4718		
c.936+2T>G	3014,8	4622	2683,6	
c.936+11G>C(rs6442)	5857,6	4718		
c.936+11G>C(rs6442)+c.936+12G>A	5841,6	4139,6		
c.936+12G>A	5897,6	4083,6		
c.949C>A(L317M)			1977,2	
c.949C>G(L317V)	3561,2	1578	1993,2	
c.943C>T(Q315X)			1977,2	
c.946C>T(R316X)			1977,2	
c.952C>T(p.Gln318X)	3328	1618	3991,4	
c.988_997delTCCAGCTCCC(S330Frameshift)	2284,4			
c.1016G>A(R339H)	3529,2			1952,2
c.1021C>T(R341W)		5199,2		
c.1021C>T(R341W)+c.1022G>C(R341P)		5255,2		
c.1022G>C(R341P)	2501,6	5271,2		
c.1048G>A(M.Cd.350)	2886,8	1304,8		3982,4
c.1051G>A(E351K)	2886,8			1623
c.1060C>T(R354C)				
c.1061G>A(R354H)		2830,8		1993,2
c.1066C>T(p.Arg356Trp)				
c.1067G>A(R356Q)		4774		
		5199,2		
c.1067G>C(R356P)	2806,8			
c.1085C>T(A362V)		3617,2	1911,2	1657
c.1088T>G(L363W)		3593,2		
c.1088T>G(L363W)+c.1093C>T(H365Y)		3577,2		
c.1093C>T(H365Y)		3617,2		
c.1096C>T(R366C)	2445,6			
	6645,2			
c.1116-34G>A(rs6461)		2942,8	2996,8	1977,2

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.936+1G>C	1690	4662	2642,6	

c.936+2T>G	1304,8	4662	2642,6	
	1690			
c.936+11G>C(rs6442)	5913,6	4662		
c.936+11G>C(rs6442)+c.936+12G>A	5913,6	4662		
c.936+12G>A	5913,6	4662		
c.949C>A(L317M)	3505,2			
c.949C>G(L317V)	3505,2			
c.943C>T(Q315X)	3505,2			
c.946C>T(R316X)	3505,2			
c.952C>T(p.Gln318X)	3962,4		3037,8	
c.988_997delTCCAGCTCCC(S330Frameshift)	1867,2	5231,2		2034,2
		6138,8		
c.1016G>A(R339H)	3545,2			
c.1021C>T(R341W)	2557,6	5215,2		
c.1021C>T(R341W)+c.1022G>C(R341P)	2557,6	5215,2		
c.1022G>C(R341P)	2557,6	5215,2		
c.1048G>A(M.Cd.350)				2987,8
c.1051G>A(E351K)				
c.1060C>T(R354C)	1883,2			
c.1061G>A(R354H)	1883,2			1343,8
c.1066C>T(p.Arg356Trp)	2862,8			
c.1067G>A(R356Q)	2862,8			
c.1067G>C(R356P)	2862,8			
c.1085C>T(A362V)		3633,2	1607	
c.1088T>G(L363W)		3633,2		
c.1088T>G(L363W)+c.1093C>T(H365Y)		3633,2		2699,6
c.1093C>T(H365Y)		3633,2		2699,6
c.1096C>T(R366C				2699,6
c.1116-34G>A(rs6461)		3521,2		

Amplicon 8

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.1116-34G>A(rs6461)		2942,8	2996,8	1977,2
c.1122C>T(rs6469)	1304,8	3256		2651,6
c.1122C>T(rs6469)+c.1123G>A(G375S)		2003,2		2955,8
c.1123G>A(G375S)	1867,2	2019,2		
c.1128C>A(Y376X)	2276,4		1961,2	1607
c.1140G>C(E380D)	3561,2			
c.1171G>A(A391T)	4646			3333
c.1214G>A(W405X)	1578	1947,2		1961,2
		2541,6		
c.1219+26G>A(rs2242571)	1578		6277,8	

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.1116-34G>A(rs6461)		3521,2	3012,8	1673
c.1122C>T(rs6469)		3272		2667,6
c.1122C>T(rs6469)+c.1123G>A(G375S)		3272		2667,6
c.1123G>A(G375S)		3272		2667,6

c.1128C>A(Y376X)	2236,4			1648
c.1140G>C(E380D)	3617,2	1538	2338,4	
c.1171G>A(A391T)	4662	2573,6		2683,6
c.1214G>A(W405X)		4452,8		1657
c.1219+26G>A(rs2242571)			6293,8	

Amplicon 9

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.1273G>A(G424S)		6042,8		1993,2
c.1273G>A(G424S)+c.1276C>T(R426C)	1304,8	6026,8		1977,2
c.1273G>A(G424S)+c.1276C>T(R426C)+c.1277G>A(R426H)		4718		
c.1273G>A(G424S)+c.1277G>A(R426H)	2485,6	4718		2642,6
c.1276C>T(R426C)	1304,8			
c.1276C>T(R426C)+c.1277G>A(R426H)		4718		1993,2
c.1277G>A(R426H)		4718		2009,2
c.1303C>T(A434V)		5697,6		
c.1303C>T(A434V)+c.1303C>T(R435C)		5681,6	2617,6	
c.1303C>T(R435C)		5697,6		
c.1330C>T(R444X)	1578			1648
c.1337T>C(L446P)	3481,2	2613,6		
c.1349C>T(M.Cd.450)		6082,8		2651,6
c.1349C>T(M.Cd.450)+c.1357C>T(p.Pro453Ser)		6066,8		2363,4
				2651,6
c.1357C>T(p.Pro453Ser)		6082,8		2363,4
c.1375C>T(M.Cd.459)		5961,6		3037,8
c.1388C>T(M.Cd.463)	1923,2	3978,4		
c.1419G>T(M473I)		1328,8		1936,2
c.1426delT(P475Frameshift)				2971,8
c.1436G>T(R479L)		3505,2		
c.1436G>T(R479L)+c.1442C>T(Q481P)	4806	6966,4		
c.1442C>T(Q481P)	4806	6926,4		
c.1444C>T(P482S)	3248	3425,2		
c.1444C>T(P482S)+c.1448_49delGGinsC(R483Frameshift)	2846,8	3192	3719,2	
c.1444C>T(P482S)+c.1447C>T(R483W)	2380,4	3409,2	4713,8	
c.1444C>T(P482S)+c.1448insC(R483Frameshift)		3770,4		
c.1444C>T(P482S)+c.1448G>A(R483Q)	3232	1482	4393,6	1993,2
		1979,2		
c.1444C>T(P482S)+c.1448G>C(R483P)	3192	3481,2	4064,4	
c.1447C>T(R483W)	2380,4	3425,2	4713,8	
c.1447C>T(R483W)+c.1448_49delGGinsC(R483Frameshift)		3192	3719,2	
c.1447C>T(R483W)+c.1448G>A(R483Q)	2364,4	1482	4697,8	1993,2
		1979,2		
c.1447C>T(R483W)+c.1448G>C(R483P)		3481,2	4064,4	
c.1448insC(R483Frameshift)	5135,2	3786,4		2050,2
c.1448G>A(R483Q)	4830	1482	4393,6	2009,2
		1995,2		
c.1448G>C(R483P)	4790	3497,2	4064,4	2050,2

c.1470A>G(rs61732562)		3208	1359,8	
c.1471G>A(M.Cd.491)				1968,2
c.1471G>A(M.Cd.491)+c.1470A>G(rs61732562)		1939,2		
c.1471G>A(M.Cd.491)+c.1470A>G(rs61732562)+c.1478A>G(rs61732563)		1939,2		
c.1471G>A(M.Cd.491)+c.1478A>G(rs61732563)				1968,2
c.1478A>G(rs61732563)				
c.1486+12C>T	3633,2			
c.1486+12C>T+c.1486+13G>A	3633,2		1936,2	
c.1486+13G>A	4580,8			2256,4

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.1273G>A(G424S)	2517,6			
c.1273G>A(G424S)+c.1276C>T(R426C)	2517,6			
c.1273G>A(G424S)+c.1276C>T(R426C)+c.1277G>A(R426H)	2517,6			
c.1273G>A(G424S)+c.1277G>A(R426H)	2517,6			
c.1276C>T(R426C)	2517,6			
c.1276C>T(R426C)+c.1277G>A(R426H)	2517,6			
c.1277G>A(R426H)	2517,6			
c.1303C>T(A434V)	2862,8	5713,6		
c.1303C>T(A434V)+c.1303C>T(R435C)	2862,8	5713,6		
c.1303C>T(R435C)	2862,8	5713,6		
c.1330C>T(R444X)	2830,8			
c.1337T>C(L446P)	2557,6			
c.1349C>T(M.Cd.450)	1867,2	6098,8		2667,6
c.1349C>T(M.Cd.450)+c.1357C>T(p.Pro453Ser)	1867,2	6098,8		2379,4
				2667,6
c.1357C>T(p.Pro453Ser)		6098,8		2379,4
c.1375C>T(M.Cd.459)		5977,6		3053,8
c.1388C>T(M.Cd.463)	2790,8	3994,4		
c.1419G>T(M473I)	2212,4		2306,4	
c.1426delT(P475Frameshift)		2982,8		3301
c.1436G>T(R479L)	1939,2	3465,2		
c.1436G>T(R479L)+c.1442C>T(Q481P)	1939,2	3441,2		
	4846	3465,2		
c.1442C>T(Q481P)	4846	3441,2		
		3465,2		
c.1444C>T(P482S)	4846	3441,2		1705
c.1444C>T(P482S)+c.1448_49delGGinsC(R483Frameshift)	4846	3441,2	4409,6	1705
c.1444C>T(P482S)+c.1447C>T(R483W)	4846	3441,2	4409,6	1705
c.1444C>T(P482S)+c.1448insC(R483Frameshift)	4846	3441,2		1705
c.1447C>T(R483W)+c.1448G>A(R483Q)	4846	3441,2	4409,6	1705
c.1447C>T(R483W)+c.1448G>C(R483P)	4846	3441,2	4409,6	1705
c.1447C>T(R483W)	4846	3441,2	4409,6	1705

c.1447C>T(R483W)+c.1448_49delGGinsC(R483Frameshift)	4846	3441,2	4409,6	1705
c.1447C>T(R483W)+c.1448G>A(R483Q)	4846	3441,2	4409,6	1705
c.1447C>T(R483W)+c.1448G>C(R483P)	4846	3441,2	4409,6	1705
c.1448insC(R483Frameshift)	4846	3441,2		1705
c.1448G>A(R483Q)	4846	3441,2	4409,6	1705
c.1448G>C(R483P)	4846	3441,2	4409,6	1705
c.1470A>G(rs61732562)		1594		
c.1471G>A(M.Cd.491)		1594		
c.1471G>A(M.Cd.491)+c.1470A>G(rs61732562)		1594		
c.1471G>A(M.Cd.491)+c.1470A>G(rs61732562)+c.1478A>G(rs61732563)		1594	1657	1927,2
c.1471G>A(M.Cd.491)+c.1478A>G(rs61732563)		1594	1657	1927,2
c.1478A>G(rs61732563)			1657	1927,2
c.1486+12C>T	4596,8			
c.1486+12C>T+c.1486+13G>A	4596,8			
c.1486+13G>A	4596,8			

Amplicon R356W

SNP/Mutation Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.1066C>T(p.Arg356Trp)	2228,4		1968,2	
c.1067G>A(R356Q)		4830		
		5199,2		
c.1067G>C(R356P)	2806,8			
c.1085C>T(A362V)		3617,2	1911,2	1657
c.1093C>T(H365Y)		3617,2		2683,6
c.1096C>T(R366C)	2445,6	1634		2683,6
	6645,2			
c.1116-34G>A(rs6461)		2942,8	2996,8	1977,2
c.1066C>T(p.Arg356Trp)+c.1067G>C(R356P)	2172,4			
c.1066C>T(p.Arg356Trp)+c.1067G>A(R356Q)		4814	1952,2	
		5199,2		

Wild Type Signals

Reference	t-forward	t-reverse	c-forward	c-reverse
c.1066C>T(p.Arg356Trp)	2862,8			
c.1067G>A(R356Q)	2862,8			
c.1067G>C(R356P)	2862,8			
c.1085C>T(A362V)	1578	3633,2	1607	
c.1093C>T(H365Y)		3633,2		2699,6
c.1096C>T(R366C)				2699,6
c.1116-34G>A(rs6461)		3521,2	3012,8	
c.1066C>T(p.Arg356Trp)+c.1067G>C(R356P)	2862,8			
c.1066C>T(p.Arg356Trp)+c.1067G>A(R356Q)	2862,8			

6.2 CNV Analysis Results of 17 Samples on MALDI-TOF MS

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
		3837	3838	3839	3840	3841	3842	3843	3844	3845	3846	3847	3848	3849	3850	3851	3852	3853
c.1-126_C/T	1.15x	1,00116001	0,932157001	0,810041416	1,069927428	2,07517692	1,559714187	1,05351141	0,811623367	1,077435249	0,731004165	1,026489502	1,004220281	1,040068927	1,760085072	1,420521219	1,579783552	1,013749837
	Ratio B/A	0,870573922	0,810571305	0,70438384	0,930371677	1,80450167	1,356273206	0,916096878	0,705759449	0,936900217	0,635655796	0,892599567	0,873235027	0,904407763	1,530508758	1,235235842	1,373724828	0,881521597
	MLPA_B	2	2		2	2	2	2	2	2	2	2	2	2	2	3	2	2
	MLPA_A	3	2		3	2	2	2	3	2	3	2	2	2	2	2	1	2
	CYP21A2	54,6119	62,1705	53,5557	58,3217	41,7009	51,3474	80,3397	49,6854	62,5338	48,8373	85,0365	36,5183	80,2648	75,548	88,4513	80,4399	67,2737
		91,8968	64,8993	64,4543	56,4882	77,233	72,4573	53,4524	76,9432	53,7643	47,8649	68,3876	80,0323	97,2856	111,2572	122,2277	61,1549	102,5844
		66,1182	82,2125	50,2253	37,7139	51,1104	58,9485	63,4656	42,5611	66,0494	93,2804	60,8494	62,6504	78,6474	102,5599	106,0335	68,8975	83,0516
		91,5498	51,9072	59,2244	66,9546	61,8163	74,006	132,4856	69,8283	84,4532	38,7842	46,2864	131,7607	79,7848	105,8744	123,3042	91,4439	59,0786
	CYP21A1P	65,4055	75,8227	78,7101	65,1625	24,4068	36,6979	92,3572	76,5437	67,4708	85,9233	93,8021	47,7899	98,4243	51,4453	78,408	57,3167	75,8107
		101,7563	82,6002	92,8407	63,8952	38,0229	51,2728	61,7151	103,7799	59,3843	77,5513	76,2752	87,5754	99,9049	72,2306	98,414	47,0617	117,1905
		79,1743	95,5182	74,2982	41,5422	25,9032	46,5736	72,7548	61,3787	71,0474	133,3997	71,0783	80,3758	84,2997	66,9365	80,9632	50,7374	89,8004
		103,0618	68,2878	77,0711	65,3041	40,1572	54,768	133,1166	96,9655	86,8671	63,0167	50,7557	140,362	88,8657	67,6282	98,4356	64,678	71,1186
c.289+138_A/G		3837	3838	3839	3840	3841	3842	3843	3844	3845	3846	3847	3848	3849	3850	3851	3852	3853
	Ratio B/A	1,000815814	1,048281408	0,364283567	1,053048951	1,786908112	1,633405666	1,020087484	0,801705655	1,109167607	0,819963529	1,100718228	1,059084554	1,018302036	1,741707792	1,418688352	1,791008937	1,057646214
	MLPA_B	2	2		2	2	2	2	2	2	2	2	2	2	2	3	2	2
	MLPA_A	2	2		2	2	2	2	3	2	3	2	2	2	2	2	1	2
	CYP21A2	76,6612	90,2189	36,1118	65,722	55,5253	61,4293	98,7776	74,895	83,7162	75,2072	103,5816	58,8357	103,2297	94,0776	115,1533	63,2322	99,5467
		110,0062	102,9256	41,5632	77,1499	54,7782	87,2699	75,1659	99,1273	72,5417	85,6852	93,5617	95,8343	119,6665	131,0233	167,1199	80,7799	151,6677
		90,2565	101,9397	43,7123	48,7455	52,4179	70,8666	74,9986	60,3873	101,9173	74,4592	53,8162	82,9535	93,2328	114,6102	119,8592	81,1754	109,3519
		112,8206	84,0274	43,1974	74,2584	83,8855	88,8256	136,1955	93,8204	100,0754	51,4814	67,4883	157,0917	100,466	116,1682	148,6071	92,5545	95,2165
	CYP21A1P	79,7555	87,6497	106,6072	66,8486	31,78	42,5703	101,0905	91,1116	77,463	86,6643	91,7781	59,1133	102,0508	51,2038	80,0023	36,585	97,0004
		114,3543	97,1955	119,4324	71,1024	33,622	52,2249	76,3368	129,0006	64,2373	98,9779	84,2435	91,0132	120,7174	78,3083	119,4506	44,9256	141,8095
		88,778	94,5191	113,2521	47,9517	28,3728	43,2886	73,8357	73,3541	89,2705	96,2857	50,1205	79,3749	87,5761	64,0479	89,4838	43,9785	106,787
		106,539	82,2863	112,5121	66,5792	44,2328	50,7189	126,2905	115,9483	92,0197	67,884	63,167	143,1933	98,7632	68,1827	99,2666	51,9204	85,3438
c.702_T/C		3837	3838	3839	3840	3841	3842	3843	3844	3845	3846	3847	3848	3849	3850	3851	3852	3853
	0.8x	1,012772303	1,080173265	0,704094967	0,98599928	1,917009719	1,918463999	1,018869178	0,776949209	0,942134814	0,800712486	0,98700863	1,046963626	1,023800023	2,003308561	1,442701403	1,835328435	1,021699429
	Ratio B/A	1,265965379	1,350216581	0,880118709	1,2324991	2,396262149	2,398079999	1,273586472	0,971186512	1,177668517	1,000890607	1,233760787	1,308704532	1,279750028	2,504135701	1,803376753	2,294160544	1,277124286
	MLPA_B	2	2		2	2	2	2	2	2	2	2	2	2	2	3	2	2
	MLPA_A	2	2		2	2	2	3	3	2	3	2	2	2	2	2	1	2
	CYP21A2	57,4278	66,5722	60,7296	67,3398	85,8584	57,5975	82,8989	59,3028	60,9271	75,1799	85,2957	44,8902	89,8157	91,7142	104,5992	102,1669	70,3349
		95,0611	74,4549	63,0546	61,0659	46,233	81,5156	54,601	84,8247	50,2352	56,4252	71,7944	84,0881	90,8758	135,0552	121,7622	96,4227	93,6348
		75,3823	94,1069	49,8826	55,3762	51,3016	73,5622	71,0272	45,2763	56,8966	117,3338	97,1183	74,8386	70,7832	97,9087	106,0391	86,7968	85,5605
		95,0791	69,8139	57,4742	63,8655	82,3091	76,8178	105,1938	77,8643	79,3205	43,8187	43,7927	119,5174	84,9433	113,0929	123,6623	95,904	72,4419
	CYP21A1P	46,6284	51,5834	66,3315	58,1461	36,7343	23,9297	65,1924	61,1232	52,1288	72,1895	68,9623	32,3644	66,3522	36,1778	56,246	46,9818	54,8218
		77,8245	52,5285	72,6705	53,4162	19,129	33,1997	42,0328	87,1604	40,0023	54,1962	57,5178	60,4418	71,4446	57,7155	64,8435	40,766	70,4874
		60,0592	72,0563	57,943	40,0795	22,9182	28,6961	62,709	49,2879	48,3842	125,607	84,2217	61,8171	58,154	40,2249	60,4126	37,4554	68,0951
		70,5899	49,6829	65,6798	49,2893	32,1004	34,8932	76,3945	77,626	69,5433	40,5044	30,837	92,4411	66,9271	40,701	71,3917	40,9972	58,7028
SLC30A3		3837	3838	3839	3840	3841	3842	3843	3844	3845	3846	3847	3848	3849	3850	3851	3852	3853
		67,38205	66,284475	64,121525	63,35615	53,430925	41,088325	71,207625	74,789725	67,644375	61,80695	51,22005	70,75785	101,443325	71,514025	98,421925	35,971075	69,581225
	WT	51,8998	61,5691	59,7484	66,7314	38,2628	35,0227	69,0819	58,9511	68,9232	51,0321	64,4157	41,5572	96,4108	64,5389	90,3445	35,0179	59,0274
		77,6135	69,5381	69,442	71,9402	65,6215	40,078	55,457	98,0065	53,2018	59,3622	50,8232	74,2388	120,2262	77,3775	111,8577	34,2397	86,3904
		61,3609	75,3835	60,6455	51,4062	48,8273	38,7842	55,1627	54,2831	68,5383	92,1957	52,9986	57,3856	95,4951	76,0489	96,1129	36,8529	70,4636
		78,654	58,6472	66,6502	63,3468	61,0121	50,4684	105,1289	87,9182	79,9142	44,6378	36,6427	109,8498	93,6412	68,0908	95,3726	37,7738	62,4435
B	c.1-126_C/T	1,184980032	1,034363533	0,93117204	0,909352604	1,139104507	1,640351365	1,215566679	0,838915038	1,035343792	0,971594376	1,335355466	1,153616825	0,869405971	1,450769534	1,17356355	2,203388486	1,176997513
	c.289+138_A/G	1,156819954	1,143892593	0,513352419	0,839305419	0,92308677	1,501114489	1,081731345	0,877740893	1,05921771	0,928157756	1,243449782	1,115678896	0,82133546	1,274936769	1,119139866	1,76665279	1,31007409
	c.702_T/C	1,198205976	1,150148281	0,901183339	0,977203476	1,243203725	1,761407285	1,101430149	0,893398458	0,91426449	1,184161328	1,454513906	1,142397275	0,8290787	1,530367645	1,158438021	2,649979185	1,156821039
A	c.1-126_C/T	1,166698364	1,093793117	1,133114387	0,837778179	0,541077522	1,036675686	1,137340734	1,018859944	0,947206032	1,31013543	1,282311175	1,132357717	0,823970281	0,812485873	0,814347819	1,374815876	1,144447299
	c.289+138_A/G	1,155876973	1,091207557	1,409211025	0,797024125	0,516583233	0,919008989	1,060429975	1,094841838	0,954966322	1,131950048	1,129671291	1,053437039	0,806573523	0,732003827	0,788855329	0,986400879	1,238669483
	c.702_T/C	0,946476102	0,851825031	1,023933851	0,792863439	0,518809566	0,734507308	0,864825572	0,919904104	0,776334322	1,183107644	1,178927002	0,872922227	0,647844252	0,611136067	0,642371606	1,155097533	0,905801457
	MALDI-TOF MS	2:2	2:2	2:3	2:2	3:2/2:1	3:2/2:1	2:2	2:3	2:2	2:3	2:2	2:2	2:2	3:2/2:1	3:2/2:1	3:2/2:1	2:2
	MLPA	2:2	2:2	2:3	2:2	2:1	2:2	2:2	2:3	2:2	2:3	2:2	2:2	2:2	2:2	3:2	2:1	2:2

6.3 Complete Results of 36 Samples from the Routine Analysis Assigned genotypes are shown for positions which resulted in at least one non-wild type genotype. Green and red indicate a match and a mismatch with conventional sequencing results, respectively. Two discrepancies at rs6472 are highlighted in blue and at rs61732563 in red. Failure of auto-calling at positions rs59064806 and rs6462 is shown with "not picked" in red. Yellow fields indicate a wild type genotype in the entire experiment set. For some variations, their trivial names are left in parenthesis and variations detected with iPLEX are indicated for easy tracking.

Amplicon 1	**3763**	**3764**	**3765**	**3767**	**3768**	**3769**	**3770**	**3771**	**3772**	**3773**
c.185_A>T(H62L)										
c.191_G>A(G64E)										
c.3_G>A(M1I)										
rs28381641, c.27insCTG (iPLEX)	5CTG	4/5CTG	4CTG	5CTG	5CTG	5CTG	4/5CTG	5CTG	5CTG	4/5CTG
c.43_G>A(A15T)										
c.56_G>A(W19X)										
c.64_insT(W22Frameshift)										
c.66_G>A(W22X)										
c.82_insC(H28Frameshift)										
c.89_C>T (p.Pro30Leu)	C	C	C	C	C	C	C	C	C	C
rs6470, c.1-4C>T (iPLEX)	C	C	C	C	C	C	C	C	C	C
rs6468, c.115T>C	T	T	T	TC	TC	TC	T	T	C	TC
rs6464, c.135A>C	A	A	A	AC	AC	AC	A	AC	C	AC
Amplicon 2	**3763**	**3764**	**3765**	**3767**	**3768**	**3769**	**3770**	**3771**	**3772**	**3773**
c.220_A>T(K74X)										
c.230_T>C(I77T)										
c.269_G>T(G90V)										
rs6462, c.289+9T>C	T	T	T	TC	T	TC	T	T	C	TC
rs6463, c.289+33C>A	A	CA	C	A	A	A	CA	A	A	CA
rs6449, c.289+67T>C	C	TC	T	C	C	C	TC	C	C	TC
c.289+84A>G										
c.289+92A>G										
c.289+100A>G										
c.200-2A>G										
c.289+1G>A										

Amplicon 3	3763	3764	3765	3767	3768	3769	3770	3771	3772	3773
c.290-13_A>G(IVS2AS,A/C-G,-13)	A	A	A	AC	AC	AC	AC	A	A	A
c.291_C>A(Y97X)										
c.314_C>T(P105L)										
c.370_C>T(M.Cd.124_C>T)										
c.371_G>A(R124H)										
c.416_T>A(V139E)										
c.439_T>C(C147R)										
c.290-109_G>C (iPLEX)										
c.290-105_delG	6G	5/6G	6G	6G	5/6G	6G	5/6G	6G	6G	6G
c.209-74_G>A(rs6450)										
c.290-67_C>A(rs6451)	A	C	C	C	C	C	C	CG	C	C
620_A/G - rs59064806, c.209-48A>G	A	A	A	A	A	A	A	A	A	A
c.290-44_G>T(rs6453)	G	G	G	G	G	G	G	G	G	G
c.290-38_39_CA>GG(rs35147842)	CA	CA	CA	CA	CA	CA	CA	CA	CA	CA
c.290-4_G>A(P.668_G/A)										
c.305_A>G(rs6474)	G	AG	A	G	G	G	AG	G	G	AG
c.315_G>C(P.Cd.105_G/C)	G	G	G	G	G	G	G	G	G	G
c.290-2_A>G										
c.329_336_delGAGACTAC(8bp-del) (iPLEX)										
c.532_G>C(G178A) (iPLEX)										
Amplicon 4	**3763**	**3764**	**3765**	**3767**	**3768**	**3769**	**3770**	**3771**	**3772**	**3773**
c.497_T>C(L166P)										
c.505_T>C(C169R)										
c.444+38_C>T(rs6466)	T	C	C	CT	CT	CT	CT	CT	C	C
c.444+39_G>A(rs58693631)										
c.508_insA(S170Frameshift)										
c.515_T>A(p.Ile172Asn)										

Amplicon 5-6	3763	3764	3765	3767	3768	3769	3770	3771	3772	3773
E6 (iPLEX)										
c.547-15_C>A(rs1040312)	C	C	C	C	C	C	C	C	C	C
c.547-8_T>C(rs1040311)	T	T	T	T	T	T	T	T	T	T
c.549_C>G(rs1040310)										
c.587_589delAGG(E196del)										
c.594_A>T(M.Cd.198)										
c.648+30_G>A										
c.648+35_A>G(rs12525076)										
c.721_C>G(M.Cd.241) (iPLEX)										
Amplicon 7	**3763**	**3764**	**3765**	**3767**	**3768**	**3769**	**3770**	**3771**	**3772**	**3773**
c.782_T>C(L261P)										
c.784_C>T(Q262X)										
c.898_C>T(L300F)										
c.920_dupT(p.Leu306PhefsX5)										
c.922_C>T(L308F)										
c.736-74_G>A										
c.736-21_C>T(rs6465)	C	C	C	C	C	C	C	C	C	C
c.744_C>G(rs6477)	C	C	C	C	C	C	C	C	C	C
c.803_G>C(rs6472)	G	G	G	G	GC	G	G	GC	G	G
c.936+11_G>C(rs6442)										
c.936+1_G>C										
c.936+2_T>G										
c.841_G>T(p.Val281Leu) (iPLEX)										
c.842_T>G(V281G) (iPLEX)										
c.871_G>A(G291S) (iPLEX)										
c.871_G>C(G291R) (iPLEX)										
c.871_G>T(G291C) (iPLEX)										
c.904_T>C(W302R) (iPLEX)										
c.910_G>A(V304M) (iPLEX)										

Amplicon 8	3763	3764	3765	3767	3768	3769	3770	3771	3772	3773
c.952_C>T(p.Gln318X)	C	C	C	C	C	C	C	C	C	C
c.988_997_delTCCAGCTCCC(S330Frameshift)										
c.1016_G>A(R339H)										
c.1021_C>T(R341W)										
c.1022_G>C(R341P)										
c.1048_G>A(M.Cd.350)										
c.1061_G>A(R354H)										
c.1116-34_G>A(rs6461)										
c.949_C>A(L317M) (iPLEX)										
c.949_C>G(L317V) (iPLEX)										
c.1051_G>A(E351K) (iPLEX)										
c.1066_C>T(p.Arg356Trp)										
c.1067_G>A(R356Q) (iPLEX)										
c.1067_G>C(R356P) (iPLEX)										
c.1085_C>T(A362V) (iPLEX)										
Amplicon 9	**3763**	**3764**	**3765**	**3767**	**3768**	**3769**	**3770**	**3771**	**3772**	**3773**
c.1123_G>A(G375S)										
c.1128_C>A(Y376X)										
c.1140_G>C(E380D)										
c.1171_G>A(A391T)										
c.1214_G>A(W405X)										
c.1122_C>T(rs6469)	T	CT	CT	C	C	C	C	C	C	C
c.1219+26_G>A(rs2242571)										

Amplicon 10	3763	3764	3765	3767	3768	3769	3770	3771	3772	3773
c.1337_T>C(L446P)										
c.1330_C>T(R444X)										
c.1303_C>T(R435C)										
c.1303_C>T(A434V)										
c.1375_C>T(M.Cd.459)										
c.1388_C>T(M.Cd.463)										
c.1436_G>T(R479L)										
c.1442_C>T(Q481P)										
c.1444_C>T(P482S)										
c.1276_C>T(R426C) (iPLEX)										
c.1277_G>A(R426H) (iPLEX)										
c.1349_C>T(M.Cd.450) (iPLEX)										
c.1357_C>T(p.Pro453Ser) (iPLEX)										
c.1448_G>A(R483Q) (iPLEX)										
c.1470_A>G(rs61732562) (iPLEX)	G	AG	AG	G	G	G	AG	G	G	G
c.1471_G>A(M.Cd.491) (iPLEX)										
c.1478_A>G(rs61732563) (iPLEX)	G	AG	AG	G	G	G	AG	G	G	G
c.1448_G>C(R483P) (iPLEX)										

Amplicon 1	3774	3775	3776	3777	3778	3779	3781	3782
c.185_A>T(H62L)								
c.191_G>A(G64E)								
c.3_G>A(M1I)								
rs28381641, c.27insCTG (iPLEX)	5CTG	5CTG	5CTG	4/5CTG	4/5CTG	4CTG	4CTG	5CTG
c.43_G>A(A15T)								
c.56_G>A(W19X)								
c.64_insT(W22Frameshift)								
c.66_G>A(W22X)								
c.82_insC(H28Frameshift)								
c.89_C>T (p.Pro30Leu)	C	C	C	C	C	C	C	C
rs6470, c.1-4C>T (iPLEX)	C	C	C	C	C	C	C	C
rs6468, c.115T>C	T	T	C	T	T	T	T	TC
rs6464, c.135A>C	A	AC	C	A	A	A	A	AC
Amplicon 2	**3774**	**3775**	**3776**	**3777**	**3778**	**3779**	**3781**	**3782**
c.220_A>T(K74X)								
c.230_T>C(I77T)								
c.269_G>T(G90V)								
rs6462, c.289+9T>C	T	T	C	TC	T	T	T	C
rs6463, c.289+33C>A	A	A	A	CA	CA	C	C	A
rs6449, c.289+67T>C	C	C	C	TC	TC	T	T	C
c.289+84A>G								
c.289+92A>G								
c.289+100A>G								
c.200-2A>G								
c.289+1G>A								

Amplicon 3	3774	3775	3776	3777	3778	3779	3781	3782
c.290-13_A>G(IVS2AS,A/C-G,-13)	AC	C	A	AC	C	AC	A	A
c.291_C>A(Y97X)								
c.314_C>T(P105L)								
c.370_C>T(M.Cd.124_C>T)								
c.371_G>A(R124H)								
c.416_T>A(V139E)								
c.439_T>C(C147R)								
c.290-109_G>C (iPLEX)								
c.290-105_delG	6G	5/6G	6G	6G	5/6G	6G	6G	6G
c.209-74_G>A(rs6450)								
c.290-67_C>A(rs6451)	CA	C	C	CA	C	C	C	CA
620_A/G - rs59064806, c.209-48A>G	A	A	A	not picked	A	A	A	not picked
c.290-44_G>T(rs6453)	G	G	G	GT	G	G	G	not picked
c.290-38_39_CA>GG(rs35147842)	CA	CA	CA	CA/GG	CA	CA	CA	CA/GG
c.290-4_G>A(P.668_G/A)								
c.305_A>G(rs6474)	G	G	G	AG	AG	A	A	G
c.315_G>C(P.Cd.105_G/C)	G	G	G	G	G	G	G	G
c.290-2_A>G								
c.329_336_delGAGACTAC(8bp-del) (iPLEX)								
c.532_G>C(G178A) (iPLEX)								
Amplicon 4	**3774**	**3775**	**3776**	**3777**	**3778**	**3779**	**3781**	**3782**
c.497_T>C(L166P)								
c.505_T>C(C169R)								
c.444+38_C>T(rs6466)	T	CT	C	CT	CT	C	C	CT
c.444+39_G>A(rs58693631)								
c.508_insA(S170Frameshift)								
c.515_T>A(p.Ile172Asn)								

Amplicon 5-6	3774	3775	3776	3777	3778	3779	3781	3782
E6 (iPLEX)								
c.547-15_C>A(rs1040312)	C	C	C	C	C	C	C	C
c.547-8_T>C(rs1040311)	T	T	T	T	T	T	T	T
c.549_C>G(rs1040310)								
c.587_589delAGG(E196del)								
c.594_A>T(M.Cd.198)								
c.648+30_G>A								
c.648+35_A>G(rs12525076)								
c.721_C>G(M.Cd.241) (iPLEX)								
Amplicon 7	**3774**	**3775**	**3776**	**3777**	**3778**	**3779**	**3781**	**3782**
c.782_T>C(L261P)								
c.784_C>T(Q262X)								
c.898_C>T(L300F)								
c.920_dupT(p.Leu306PhefsX5)								
c.922_C>T(L308F)								
c.736-74_G>A								
c.736-21_C>T(rs6465)	C	C	C	C	C	C	C	C
c.744_C>G(rs6477)	C	C	C	C	C	C	C	C
c.803_G>C(rs6472)	G	GC	G	G	G	G	G	G
c.936+11_G>C(rs6442)								
c.936+1_G>C								
c.936+2_T>G								
c.841_G>T(p.Val281Leu) (iPLEX)								
c.842_T>G(V281G) (iPLEX)								
c.871_G>A(G291S) (iPLEX)								
c.871_G>C(G291R) (iPLEX)								
c.871_G>T(G291C) (iPLEX)								
c.904_T>C(W302R) (iPLEX)								
c.910_G>A(V304M) (iPLEX)								

Amplicon 8	3774	3775	3776	3777	3778	3779	3781	3782
c.952_C>T(p.Gln318X)	C	C	C	C	C	C	T	C
c.988_997_delTCCAGCTCCC(S330Frameshift)								
c.1016_G>A(R339H)								
c.1021_C>T(R341W)								
c.1022_G>C(R341P)								
c.1048_G>A(M.Cd.350)								
c.1061_G>A(R354H)								
c.1116-34_G>A(rs6461)								
c.949_C>A(L317M) (iPLEX)								
c.949_C>G(L317V) (iPLEX)								
c.1051_G>A(E351K) (iPLEX)								
c.1066_C>T(p.Arg356Trp)								
c.1067_G>A(R356Q) (iPLEX)								
c.1067_G>C(R356P) (iPLEX)								
c.1085_C>T(A362V) (iPLEX)								
Amplicon 9	**3774**	**3775**	**3776**	**3777**	**3778**	**3779**	**3781**	**3782**
c.1123_G>A(G375S)								
c.1128_C>A(Y376X)								
c.1140_G>C(E380D)								
c.1171_G>A(A391T)								
c.1214_G>A(W405X)								
c.1122_C>T(rs6469)	C	C	C	C	C	CT	T	C
c.1219+26_G>A(rs2242571)								

Amplicon 10	3774	3775	3776	3777	3778	3779	3781	3782
c.1337_T>C(L446P)								
c.1330_C>T(R444X)								
c.1303_C>T(R435C)								
c.1303_C>T(A434V)								
c.1375_C>T(M.Cd.459)								
c.1388_C>T(M.Cd.463)								
c.1436_G>T(R479L)								
c.1442_C>T(Q481P)								
c.1444_C>T(P482S)								
c.1276_C>T(R426C) (iPLEX)								
c.1277_G>A(R426H) (iPLEX)								
c.1349_C>T(M.Cd.450) (iPLEX)								
c.1357_C>T(p.Pro453Ser) (iPLEX)								
c.1448_G>A(R483Q) (iPLEX)								
c.1470_A>G(rs61732562) (iPLEX)	G	G	G	AG	AG	AG	G	G
c.1471_G>A(M.Cd.491) (iPLEX)								
c.1478_A>G(rs61732563) (iPLEX)	G	G	G	AG	AG	AG	A	G
c.1448_G>C(R483P) (iPLEX)								

Amplicon 1	3783	3785	3786	3787	3788	3789	3790	3791	3792
c.185_A>T(H62L)									
c.191_G>A(G64E)									
c.3_G>A(M1I)									
rs28381641, c.27insCTG (iPLEX)	4CTG	4/5CTG	5CTG	4/5CTG	4/5CTG	4CTG	5CTG	5CTG	4/5CTG
c.43_G>A(A15T)									
c.56_G>A(W19X)									
c.64_insT(W22Frameshift)									
c.66_G>A(W22X)									
c.82_insC(H28Frameshift)									
c.89_C>T (p.Pro30Leu)	C	C	C	C	C	C	C	C	C
rs6470, c.1-4C>T (iPLEX)	C	CT	C	C	C	C	C	C	C
rs6468, c.115T>C	T	T	T	TC	T	T	T	T	T
rs6464, c.135A>C	A	AC	A	AC	A	A	A	A	A
Amplicon 2	**3783**	**3785**	**3786**	**3787**	**3788**	**3789**	**3790**	**3791**	**3792**
c.220_A>T(K74X)									
c.230_T>C(I77T)									
c.269_G>T(G90V)									
rs6462, c.289+9T>C	T	T	TC	T	T	T	T	T	TC
rs6463, c.289+33C>A	C	CA	A	CA	CA	C	A	A	CA
rs6449, c.289+67T>C	T	TC	C	TC	TC	T	C	C	TC
c.289+84A>G									
c.289+92A>G									
c.289+100A>G									
c.200-2A>G									
c.289+1G>A									

Amplicon 3	3783	3785	3786	3787	3788	3789	3790	3791	3792
c.290-13_A>G(IVS2AS,A/C-G,-13)	A	AC	A	A	AC	AC	C	AC	AC
c.291_C>A(Y97X)									
c.314_C>T(P105L)									
c.370_C>T(M.Cd.124_C>T)									
c.371_G>A(R124H)									
c.416_T>A(V139E)									
c.439_T>C(C147R)									
c.290-109_G>C (iPLEX)									
c.290-105_delG	6G	6G	5/6G	5/6G	6G	5/6G	6G	5/6G	5/6G
c.209-74_G>A(rs6450)									
c.290-67_C>A(rs6451)	C	C	CA	C	C	C	C	C	CG
620_A/G - rs59064806, c.209-48A>G	A	A	not picked	A	A	not picked	A	A	A
c.290-44_G>T(rs6453)	G	G	not picked	G	G	GT	G	G	G
c.290-38_39_CA>GG(rs35147842)	CA	CA	CA	CA	CA	CA/GG	CA	CA	CA
c.290-4_G>A(P.668_G/A)									
c.305_A>G(rs6474)	A	AG	G	AG	AG	A	G	G	AG
c.315_G>C(P.Cd.105_G/C)	G	G	G	G	G	G	C	G	G
c.290-2_A>G									
c.329_336_delGAGACTAC(8bp-del) (iPLEX)									
c.532_G>C(G178A) (iPLEX)									
Amplicon 4	**3783**	**3785**	**3786**	**3787**	**3788**	**3789**	**3790**	**3791**	**3792**
c.497_T>C(L166P)									
c.505_T>C(C169R)									
c.444+38_C>T(rs6466)	C	C	CT	C	CT	C	C	C	CT
c.444+39_G>A(rs58693631)									
c.508_insA(S170Frameshift)									
c.515_T>A(p.Ile172Asn)									

Amplicon 5-6	3783	3785	3786	3787	3788	3789	3790	3791	3792
E6 (iPLEX)									
c.547-15_C>A(rs1040312)	C	C	C	C	C	C	A	C	C
c.547-8_T>C(rs1040311)	T	T	T	T	T	T	C	T	T
c.549_C>G(rs1040310)									
c.587_589delAGG(E196del)									
c.594_A>T(M.Cd.198)									
c.648+30_G>A									
c.648+35_A>G(rs12525076)									
c.721_C>G(M.Cd.241) (iPLEX)									
Amplicon 7	**3783**	**3785**	**3786**	**3787**	**3788**	**3789**	**3790**	**3791**	**3792**
c.782_T>C(L261P)									
c.784_C>T(Q262X)									
c.898_C>T(L300F)									
c.920_dupT(p.Leu306PhefsX5)									
c.922_C>T(L308F)									
c.736-74_G>A									
c.736-21_C>T(rs6465)	C	C	C	C	C	C	T	C	C
c.744_C>G(rs6477)	CG	C	C	C	C	CG	G	C	CG
c.803_G>C(rs6472)	G	G	G	GC	G	G	G	G	GC
c.936+11_G>C(rs6442)									
c.936+1_G>C									
c.936+2_T>G									
c.841_G>T(p.Val281Leu) (iPLEX)									
c.842_T>G(V281G) (iPLEX)									
c.871_G>A(G291S) (iPLEX)									
c.871_G>C(G291R) (iPLEX)									
c.871_G>T(G291C) (iPLEX)									
c.904_T>C(W302R) (iPLEX)									
c.910_G>A(V304M) (iPLEX)									

Amplicon 8	3783	3785	3786	3787	3788	3789	3790	3791	3792
c.952_C>T(p.Gln318X)	C	C	C	C	C	C	C	C	CT
c.988_997_delTCCAGCTCCC(S330Frameshift)									
c.1016_G>A(R339H)									
c.1021_C>T(R341W)									
c.1022_G>C(R341P)									
c.1048_G>A(M.Cd.350)									
c.1061_G>A(R354H)									
c.1116-34_G>A(rs6461)									
c.949_C>A(L317M) (iPLEX)									
c.949_C>G(L317V) (iPLEX)									
c.1051_G>A(E351K) (iPLEX)									
c.1066_C>T(p.Arg356Trp)									
c.1067_G>A(R356Q) (iPLEX)									
c.1067_G>C(R356P) (iPLEX)									
c.1085_C>T(A362V) (iPLEX)									
Amplicon 9	**3783**	**3785**	**3786**	**3787**	**3788**	**3789**	**3790**	**3791**	**3792**
c.1123_G>A(G375S)									
c.1128_C>A(Y376X)									
c.1140_G>C(E380D)									
c.1171_G>A(A391T)									
c.1214_G>A(W405X)									
c.1122_C>T(rs6469)	CT	C	C	C	C	CT	C	C	C
c.1219+26_G>A(rs2242571)									

Amplicon 10	3783	3785	3786	3787	3788	3789	3790	3791	3792
c.1337_T>C(L446P)									
c.1330_C>T(R444X)									
c.1303_C>T(R435C)									
c.1303_C>T(A434V)									
c.1375_C>T(M.Cd.459)									
c.1388_C>T(M.Cd.463)									
c.1436_G>T(R479L)									
c.1442_C>T(Q481P)									
c.1444_C>T(P482S)									
c.1276_C>T(R426C) (iPLEX)									
c.1277_G>A(R426H) (iPLEX)									
c.1349_C>T(M.Cd.450) (iPLEX)									
c.1357_C>T(p.Pro453Ser) (iPLEX)									
c.1448_G>A(R483Q) (iPLEX)									
c.1470_A>G(rs61732562) (iPLEX)	A	G	G	AG	AG	A	G	G	AG
c.1471_G>A(M.Cd.491) (iPLEX)									
c.1478_A>G(rs61732563) (iPLEX)	A	G	G	AG	AG	A	G	G	AG
c.1448_G>C(R483P) (iPLEX)									

Amplicon 1	3794	3795	3796	3797	3798	3799	3800	3801	3802
c.185_A>T(H62L)									
c.191_G>A(G64E)									
c.3_G>A(M1I)									
rs28381641, c.27insCTG (iPLEX)	4/5CTG	5CTG	5CTG	4CTG	4/5CTG	4/5CTG	4/5CTG	5CTG	5CTG
c.43_G>A(A15T)									
c.56_G>A(W19X)									
c.64_insT(W22Frameshift)									
c.66_G>A(W22X)									
c.82_insC(H28Frameshift)									
c.89_C>T (p.Pro30Leu)	C	C	C	C	C	C	C	C	C
rs6470, c.1-4C>T (iPLEX)	C	C	C	C	C	C	C	C	C
rs6468, c.115T>C	T	TC	C	T	T	TC	T	T	TC
rs6464, c.135A>C	A	AC	C	A	A	AC	AC	A	AC
Amplicon 2	**3794**	**3795**	**3796**	**3797**	**3798**	**3799**	**3800**	**3801**	**3802**
c.220_A>T(K74X)									
c.230_T>C(I77T)									
c.269_G>T(G90V)									
rs6462, c.289+9T>C	not picked	T	T	T	T	TC	TC	T	C
rs6463, c.289+33C>A	CA	A	A	C	CA	CA	CA	A	A
rs6449, c.289+67T>C	TC	C	C	T	C	TC	TC	C	C
c.289+84A>G									
c.289+92A>G									
c.289+100A>G									
c.200-2A>G									
c.289+1G>A									

Amplicon 3	3794	3795	3796	3797	3798	3799	3800	3801	3802
c.290-13_A>G(IVS2AS,A/C-G,-13)	A	AC	A	A	C	A	AC	AC	A
c.291_C>A(Y97X)									
c.314_C>T(P105L)									
c.370_C>T(M.Cd.124_C>T)									
c.371_G>A(R124H)									
c.416_T>A(V139E)									
c.439_T>C(C147R)									
c.290-109_G>C (iPLEX)									
c.290-105_delG	6G	5/6G	5G	6G	6G	6G	5/6G	5/6G	6G
c.209-74_G>A(rs6450)									
c.290-67_C>A(rs6451)	CA	C	C	C	C	C	C	C	CA
620_A/G - rs59064806, c.209-48A>G	A	A	A	A	A	A	A	A	not picked
c.290-44_G>T(rs6453)	G	G	G	G	G	G	G	G	GT
c.290-38_39_CA>GG(rs35147842)	CA	CA	CA	CA	CA	CA	CA	CA	CA/GG
c.290-4_G>A(P.668_G/A)									
c.305_A>G(rs6474)	AG	AG	G	A	AG	AG	AG	G	G
c.315_G>C(P.Cd.105_G/C)	G	G	G	G	G	G	G	G	G
c.290-2_A>G									
c.329_336_delGAGACTAC(8bp-del) (iPLEX)									
c.532_G>C(G178A) (iPLEX)									
Amplicon 4	**3794**	**3795**	**3796**	**3797**	**3798**	**3799**	**3800**	**3801**	**3802**
c.497_T>C(L166P)									
c.505_T>C(C169R)									
c.444+38_C>T(rs6466)	CT	C	C	C	C	C	C	CT	CT
c.444+39_G>A(rs58693631)									
c.508_insA(S170Frameshift)									
c.515_T>A(p.Ile172Asn)									

Amplicon 5-6	3794	3795	3796	3797	3798	3799	3800	3801	3802
E6 (iPLEX)									
c.547-15_C>A(rs1040312)	C	C	C	C	C	C	C	C	C
c.547-8_T>C(rs1040311)	T	T	T	T	T	T	T	T	T
c.549_C>G(rs1040310)									
c.587_589delAGG(E196del)									
c.594_A>T(M.Cd.198)									
c.648+30_G>A									
c.648+35_A>G(rs12525076)									
c.721_C>G(M.Cd.241) (iPLEX)									
Amplicon 7	**3794**	**3795**	**3796**	**3797**	**3798**	**3799**	**3800**	**3801**	**3802**
c.782_T>C(L261P)									
c.784_C>T(Q262X)									
c.898_C>T(L300F)									
c.920_dupT(p.Leu306PhefsX5)									
c.922_C>T(L308F)									
c.736-74_G>A									
c.736-21_C>T(rs6465)	C	CT	C	C	CT	C	C	C	C
c.744_C>G(rs6477)	C	CG	C	C	CG	C	C	C	C
c.803_G>C(rs6472)	G	GC	C	G	G	G	G	G	G
c.936+11_G>C(rs6442)									
c.936+1_G>C									
c.936+2_T>G									
c.841_G>T(p.Val281Leu) (iPLEX)									
c.842_T>G(V281G) (iPLEX)									
c.871_G>A(G291S) (iPLEX)									
c.871_G>C(G291R) (iPLEX)									
c.871_G>T(G291C) (iPLEX)									
c.904_T>C(W302R) (iPLEX)									
c.910_G>A(V304M) (iPLEX)									

Amplicon 8	3794	3795	3796	3797	3798	3799	3800	3801	3802
c.952_C>T(p.Gln318X)	C	C	C	C	C	C	C	C	C
c.988_997_delTCCAGCTCCC(S330Frameshift)									
c.1016_G>A(R339H)									
c.1021_C>T(R341W)									
c.1022_G>C(R341P)									
c.1048_G>A(M.Cd.350)									
c.1061_G>A(R354H)									
c.1116-34_G>A(rs6461)									
c.949_C>A(L317M) (iPLEX)									
c.949_C>G(L317V) (iPLEX)									
c.1051_G>A(E351K) (iPLEX)									
c.1066_C>T(p.Arg356Trp)									
c.1067_G>A(R356Q) (iPLEX)									
c.1067_G>C(R356P) (iPLEX)									
c.1085_C>T(A362V) (iPLEX)									
Amplicon 9	**3794**	**3795**	**3796**	**3797**	**3798**	**3799**	**3800**	**3801**	**3802**
c.1123_G>A(G375S)									
c.1128_C>A(Y376X)									
c.1140_G>C(E380D)									
c.1171_G>A(A391T)									
c.1214_G>A(W405X)									
c.1122_C>T(rs6469)	CT	C	C	CT	CT	C	C	C	C
c.1219+26_G>A(rs2242571)									

Amplicon 10	3794	3795	3796	3797	3798	3799	3800	3801	3802
c.1337_T>C(L446P)									
c.1330_C>T(R444X)									
c.1303_C>T(R435C)									
c.1303_C>T(A434V)									
c.1375_C>T(M.Cd.459)									
c.1388_C>T(M.Cd.463)									
c.1436_G>T(R479L)									
c.1442_C>T(Q481P)									
c.1444_C>T(P482S)									
c.1276_C>T(R426C) (iPLEX)									
c.1277_G>A(R426H) (iPLEX)									
c.1349_C>T(M.Cd.450) (iPLEX)									
c.1357_C>T(p.Pro453Ser) (iPLEX)									
c.1448_G>A(R483Q) (iPLEX)									
c.1470_A>G(rs61732562) (iPLEX)	AG	G	G	A	AG	G	G	G	G
c.1471_G>A(M.Cd.491) (iPLEX)									
c.1478_A>G(rs61732563) (iPLEX)	AG	G	G	A	AG	G	G	G	G
c.1448_G>C(R483P) (iPLEX)									

References

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Mar ;409(6822):860-921.
2. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science (New York, N.Y.). 2001 Feb ;291(5507):1304-51.
3. http://en.wikipedia.org/wiki/File:Steroidogenesis.svg
4. Merke DP, Bornstein SR. Congenital Adrenal Hyperplasia. Lancet. 2005;365:2125-36
5. White PC, Speiser PW. Congenital adrenal hyperplasia due to 21-hydroxylase deficiency. Endocrine reviews. 2000;21(3):245-91.
6. Speiser PW, Dupont B, Rubinstein P, et al. High frequency of nonclassical steroid 21-hydroxylase deficiency. American journal of human genetics. 1985;37(4):650-67.
7. Olgemöller B, Roscher AA, Liebl B, Fingerhut R. Screening for congenital adrenal hyperplasia: adjustment of 17-hydroxyprogesterone cut-off values to both age and birth weight markedly improves the predictive value. The Journal of clinical endocrinology and metabolism. 2003;88(12):5790-4.
8. Kösel S, Burggraf S, Fingerhut R, et al. Rapid second-tier molecular genetic analysis for congenital adrenal hyperplasia attributable to steroid 21-hydroxylase deficiency. Clinical chemistry. 2005;51(2):298-304.
9. Forest MG. Recent advances in the diagnosis and management of congenital adrenal hyperplasia due to 21-hydroxylase deficiency. Human reproduction update. 10(6):469-85.
10. Cutfield WS, Webster D. Newborn screening for congenital adrenal hyperplasia in New Zealand. The Journal of pediatrics. 1995;126(1):118-21.
11. Ibáñez L, Bonnin MR, Zampolli M, et al. Usefulness of an ACTH test in the diagnosis of nonclassical 21-hydroxylase deficiency among children presenting with premature pubarche. Hormone research. 1995;44(2):51-6.
12. Török D, Halász Z, Garami M, et al. Limited value of serum steroid measurements in identification of mild form of 21-hydroxylase deficiency. Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association. 2003;111(1):27-32.
13. Nordenström A, Thilén A, Hagenfeldt L, Larsson A, Wedell A. Genotyping is a valuable diagnostic complement to neonatal screening for congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency. The Journal of clinical endocrinology and metabolism. 1999;84(5):1505-9.
14. Huynh T, McGown I, Cowley D, et al. The clinical and biochemical spectrum of congenital adrenal hyperplasia secondary to 21-hydroxylase deficiency. The Clinical biochemist. Reviews / Australian Association of Clinical Biochemists. 2009;30(2):75-86.
15. Clayton PE, Miller WL, Oberfield SE, et al. Consensus statement on 21-hydroxylase deficiency from the European Society for Paediatric Endocrinology and the Lawson Wilkins Pediatric Endocrine Society. Hormone research. 2002;58(4):188-95.
16. Young MC, Hughes IA. Response to treatment of congenital adrenal hyperplasia in infancy. Archives of disease in childhood. 1990;65(4):441-4.

17. Punthakee Z, Legault L, Polychronakos C. Prednisolone in the treatment of adrenal insufficiency: a re-evaluation of relative potency. The Journal of Pediatrics. 2003;143(3):402-405.
18. Rivkees SA, Crawford JD. Dexamethasone treatment of virilizing congenital adrenal hyperplasia: the ability to achieve normal growth. Pediatrics. 2000;106(4):767-73.
19. Mullis PE, Hindmarsh PC, Brook CG. Sodium chloride supplement at diagnosis and during infancy in children with salt-losing 21-hydroxylase deficiency. European journal of pediatrics. 1990;150(1):22-5.
20. Tajima T, Okada T, Ma XM, et al. Restoration of adrenal steroidogenesis by adenovirus-mediated transfer of human cytochromeP450 21-hydroxylase into the adrenal gland of 21-hydroxylase-deficient mice. Gene therapy. 1999;6(11):1898-903.
21. White PC, Grossberger D, Onufer BJ, et al. Two genes encoding steroid 21-hydroxylase are located near the genes encoding the fourth component of complement in man. Proceedings of the National Academy of Sciences of the United States of America. 1985;82(4):1089-93.
22. Yang Z, Mendoza AR, Welch TR, Zipf WB, Yu CY. Modular variations of the human major histocompatibility complex class III genes for serine/threonine kinase RP, complement component C4, steroid 21-hydroxylase CYP21, and tenascin TNX (the RCCX module). A mechanism for gene deletions and disease associat. The Journal of biological chemistry. 1999;274(17):12147-56.
23. White PC, New MI, Dupont B. Structure of human steroid 21-hydroxylase genes. Proceedings of the National Academy of Sciences of the United States of America. 1986;83(14):5111-5.
24. Rodrigues NR, Dunham I, Yu CY, et al. Molecular characterization of the HLA-linked steroid 21-hydroxylase B gene from an individual with congenital adrenal hyperplasia. The EMBO journal. 1987;6(6):1653-61.
25. White PC, Tusie-Luna MT, New MI, Speiser PW. Mutations in steroid 21-hydroxylase (CYP21). Human mutation. 1994;3(4):373-8.
26. Koppens PFJ, Smeets HJM, Wijs IJ de, Degenhart HJ. Mapping of a de novo unequal crossover causing a deletion of the steroid 21-hydroxylase (CYP21A2) gene and a non-functional hybrid tenascin-X (TNXB) gene. Journal of medical genetics. 2003;40(5):e53.
27. Tusié-Luna MT, White PC. Gene conversions and unequal crossovers between CYP21 (steroid 21-hydroxylase gene) and CYP21P involve different mechanisms. Proceedings of the National Academy of Sciences of the United States of America. 1995;92(23):10796-800.
28. Kharrat M, Tardy V, M'Rad R, et al. Molecular genetic analysis of Tunisian patients with a classic form of 21-hydroxylase deficiency: identification of four novel mutations and high prevalence of Q318X mutation. The Journal of clinical endocrinology and metabolism. 2004;89(1):368-74.
29. Wilson RC, Nimkarn S, Dumic M, et al. Ethnic-specific distribution of mutations in 716 patients with congenital adrenal hyperplasia owing to 21-hydroxylase deficiency. Molecular genetics and metabolism. 2007;90(4):414-21.
30. Speiser PW, Dupont J, Zhu D, et al. Disease expression and molecular genotype in congenital adrenal hyperplasia due to 21-hydroxylase deficiency. The Journal of clinical investigation. 1992;90(2):584-95.
31. Jääskeläinen J, Levo A, Voutilainen R, Partanen J. Population-wide evaluation of disease manifestation in relation to molecular genotype in steroid 21-

hydroxylase (CYP21) deficiency: good correlation in a well defined population. The Journal of clinical endocrinology and metabolism. 1997;82(10):3293-7.
32. Krone N, Braun A, Roscher AA, Knorr D, Schwarz HP. Predicting phenotype in steroid 21-hydroxylase deficiency? Comprehensive genotyping in 155 unrelated, well defined patients from southern Germany. The Journal of clinical endocrinology and metabolism. 2000;85(3):1059-65.
33. Iafrate AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. Nature genetics. 2004;36(9):949-51.
34. Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. Science (New York, N.Y.). 2004;305(5683):525-8.
35. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll S a, Altshuler DM, et al. Copy number variation: new insights in genome diversity. Genome research. 2006 Aug ;16(8):949-61.
36. Freeman JL, Perry GH, Feuk L, et al. Copy number variation: new insights in genome diversity. Genome research. 2006;16(8):949-61.
37. Qian W, Zhang J. Gene dosage and gene duplicability. Genetics. 2008;179(4):2319-24.
38. Baumgartner-Parzer SM, Fischer G, Vierhapper H. Predisposition for de novo gene aberrations in the offspring of mothers with a duplicated CYP21A2 gene. The Journal of clinical endocrinology and metabolism. 2007;92(3):1164-7.
39. Collier S, Sinnott PJ, Dyer PA, et al. Pulsed field gel electrophoresis identifies a high degree of variability in the number of tandem 21-hydroxylase and complement C4 gene repeats in 21-hydroxylase deficiency haplotypes. The EMBO journal. 1989;8(5):1393-402.
40. Chung EK, Yang Y, Rupert KL, et al. Determining the one, two, three, or four long and short loci of human complement C4 in a major histocompatibility complex haplotype encoding C4A or C4B proteins. American journal of human genetics. 2002;71(4):810-22.
41. Wedell A, Stengler B, Luthman H. Characterization of mutations on the rare duplicated C4/CYP21 haplotype in steroid 21-hydroxylase deficiency. Human genetics. 1994;94(1):50-4.
42. Baumgartner-Parzer SM, Fischer G, Vierhapper H. Predisposition for de novo gene aberrations in the offspring of mothers with a duplicated CYP21A2 gene. The Journal of clinical endocrinology and metabolism. 2007;92(3):1164-7.
43. Levo A, Partanen J. Mutation-haplotype analysis of steroid 21-hydroxylase (CYP21) deficiency in Finland. Implications for the population history of defective alleles. Human genetics. 1997;99(4):488-97.
44. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic acids research. 2002 Jun ;30(12):e57.
45. Park C, Correll D, Oeth P. Measuring Allele-Specific Expression Using MassARRAY™. SEQUENOM, Inc. 2004
46. Park C, Correll D, Jurinke C. Gene Expression Analysis Using Competitive PCR and MassARRAY™. SEQUENOM, Inc. 2004
47. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Analytical chemistry. 1988;60(20):2299-301.
48. Tanaka K, Waki H, Ido Y, et al. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry. 1988;2(8):151-153.

49. Fitzgerald MC, Zhu L, Smith LM. The analysis of mock DNA sequencing reactions using matrix-assisted laser desorption/ionization mass spectrometry. Rapid Communications in Mass Spectrometry. 1993;7(10):895-897.
50. Nordhoff E, Ingendoh A, Cramer R, et al. Matrix-assisted laser desorption/ionization mass spectrometry of nucleic acids with wavelengths in the ultraviolet and infrared. Rapid communications in mass spectrometry : RCM. 1992;6(12):771-6.
51. Mouradian S, Rank DR, Smith LM. Analyzing Sequencing Reactions from Bacteriophage M13 by Matrix-assisted Laser Desorption/Ionization Mass Spectrometry. Rapid Communications in Mass Spectrometry. 1996;10(12):1475-1478.
52. Wu KJ, Steding A, Becker CH. Matrix-assisted laser desorption time-of-flight mass spectrometry of oligonucleotides using 3-hydroxypicolinic acid as an ultraviolet-sensitive matrix. Rapid communications in mass spectrometry : RCM. 1993;7(2):142-6.
53. Nordhoff E, Kirpekar F, Karas M, et al. Comparison of IR- and UV-matrix-assisted laser desorption/ionization mass spectrometry of oligodeoxynucleotides. Nucleic acids research. 1994;22(13):2460-5.
54. Pieles U, Zürcher W, Schär M, Moser HE. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: a powerful tool for the mass and sequence analysis of natural and modified oligonucleotides. Nucleic acids research. 1993;21(14):3191-6.
55. Kirpekar F, Nordhoff E, Kristiansen K, et al. Matrix assisted laser desorption/ionization mass spectrometry of enzymatically synthesized RNA up to 150 kDa. Nucleic acids research. 1994;22(19):3866-70.
56. Vestal ML, Juhasz P, Martin SA. Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry. 1995;9(11):1044-1050.
57. Haff LA, Smirnov IP. Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. Genome research. 1997;7(4):378-88.
58. Haff LA, Smirnov IP. Multiplex genotyping of PCR products with MassTag-labeled primers. Nucleic acids research. 1997;25(18):3749-50.
59. Little DP, Braun A, Darnhofer-Demar B, Köster H. Identification of apolipoprotein E polymorphisms using temperature cycled primer oligo base extension and mass spectrometry. European journal of clinical chemistry and clinical biochemistry : journal of the Forum of European Clinical Chemistry Societies. 1997;35(7):545-8.
60. Hartmer R, Storm N, Boecker S, et al. RNase T1 mediated base-specific cleavage and MALDI-TOF MS for high-throughput comparative sequence analysis. Nucleic acids research. 2003;31(9):e47.
61. Storm N, Darnhofer-Patel B, Boom D van den, Rodi CP. MALDI-TOF mass spectrometry-based SNP genotyping. Methods in molecular biology (Clifton, N.J.). 2003;212:241-62.
62. Ragoussis J, Elvidge GP, Kaur K, Colella S. Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research. PLoS genetics. 2006;2(7):e100.
63. Misra A, Hong J-Y, Kim S. Multiplex genotyping of cytochrome p450 single-nucleotide polymorphisms by use of MALDI-TOF mass spectrometry. Clinical chemistry. 2007;53(5):933-9.

64. Ghebranious N, Ivacic L, Mallum J, Dokken C. Detection of ApoE E2, E3 and E4 alleles using MALDI-TOF mass spectrometry and the homogeneous mass-extend technology. Nucleic acids research. 2005;33(17):e149.
65. Honisch C, Raghunathan A, Cantor CR, Palsson BØ, Boom D van den. High-throughput mutation detection underlying adaptive evolution of Escherichia coli-K12. Genome research. 2004;14(12):2495-502.
66. Sasayama T, Kato M, Aburatani H, Kuzuya A, Komiyama M. Simultaneous genotyping of indels and SNPs by mass spectroscopy. Journal of the American Society for Mass Spectrometry. 2006;17(1):3-8.
67. Blievernicht JK, Schaeffeler E, Klein K, et al. MALDI-TOF mass spectrometry for multiplex genotyping of CYP2B6 single-nucleotide polymorphisms. Clinical chemistry. 2007;53(1):24-33.
68. Cantürk C, Baade U, Salazar R, Storm N, Pörtner R, Höppner W. Sequence Analysis of CYP21A1P in a German Population to Aid in the Molecular Biological Diagnosis of Congenital Adrenal Hyperplasia. Clinical chemistry. 2011 Mar;57(3):511-7. Epub 2010 Dec 9.
69. Ehrich M, Correll D, van den Boom D. SNP Discovery Using the MassARRAY™ System, August 2004. Sequenom, Inc.
70. iPLEX® Gold Application Guide. 2009. Sequenom, Inc.
71. http://www.cypalleles.ki.se/cyp21.htm
72. http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=9606
73. Tusie-Luna MT, Speiser PW, Dumic M, New MI, White PC. A mutation (Pro-30 to Leu) in CYP21 represents a potential nonclassic steroid 21-hydroxylase deficiency allele. Molecular endocrinology (Baltimore, Md.). 1991 May ;5(5):685-92.
74. Rodrigues NR, Dunham I, Yu CY, Carroll MC, Porter RR, Campbell RD. Molecular characterization of the HLA-linked steroid 21-hydroxylase B gene from an individual with congenital adrenal hyperplasia. The EMBO journal. 1987 Jun ;6(6):1653-61.
75. Higashi Y, Hiromasa T, Tanae A, Miki T, Nakura J, Kondo T, et al. Effects of individual mutations in the P-450(C21) pseudogene on the P-450(C21) activity and their distribution in the patient genomes of congenital steroid 21-hydroxylase deficiency. Journal of biochemistry. 1991 Apr ;109(4):638-44.
76. Higashi Y, Tanae A, Inoue H, Fujii-Kuriyama Y. Evidence for frequent gene conversion in the steroid 21-hydroxylase P-450(C21) gene: implications for steroid 21-hydroxylase deficiency. American journal of human genetics. 1988 Jan ;42(1):17-25.
77. Amor M, Parker KL, Globerman H, New MI, White PC. Mutation in the CYP21B gene (Ile-172----Asn) causes steroid 21-hydroxylase deficiency. Proceedings of the National Academy of Sciences of the United States of America. 1988 Mar ;85(5):1600-4.
78. Tusie-Luna MT, Traktman P, White PC. Determination of functional effects of mutations in the steroid 21-hydroxylase gene (CYP21) using recombinant vaccinia virus. The Journal of biological chemistry. 1990 Dec ;265(34):20916-22.
79. Robins T, Barbaro M, Lajic S, Wedell A. Not all amino acid substitutions of the common cluster E6 mutation in CYP21 cause congenital adrenal hyperplasia. The Journal of clinical endocrinology and metabolism. 2005 Apr ;90(4):2148-53.

80. Speiser PW, New MI, White PC. Molecular genetic analysis of nonclassic steroid 21-hydroxylase deficiency associated with HLA-B14,DR1. The New England journal of medicine. 1988 Jul ;319(1):19-23.
81. Globerman H, Amor M, Parker KL, New MI, White PC. Nonsense mutation causing steroid 21-hydroxylase deficiency. The Journal of clinical investigation. 1988 Jul ;82(1):139-44.
82. Chiou SH, Hu MC, Chung BC. A missense mutation at Ile172----Asn or Arg356----Trp causes steroid 21-hydroxylase deficiency. The Journal of biological chemistry. 1990 Feb ;265(6):3549-52.
83. Owerbach D, Sherman L, Ballard AL, Azziz R. Pro-453 to Ser mutation in CYP21 is associated with nonclassic steroid 21-hydroxylase deficiency. Molecular endocrinology (Baltimore, Md.). 1992 Aug ;6(8):1211-5.
84. Helmberg A, Tusie-Luna MT, Tabarelli M, Kofler R, White PC. R339H and P453S: CYP21 mutations associated with nonclassic steroid 21-hydroxylase deficiency that are not apparent gene conversions. Molecular endocrinology (Baltimore, Md.). 1992 Aug ;6(8):1318-22.
85. Nikoshkov A, Lajic S, Holst M, Wedell A, Luthman H. Synergistic effect of partially inactivating mutations in steroid 21-hydroxylase deficiency. The Journal of clinical endocrinology and metabolism. 1997 Jan ;82(1):194-9.
86. Lobato MN, Ordóñez-Sánchez ML, Tusié-Luna MT, Meseguer A. Mutation analysis in patients with congenital adrenal hyperplasia in the Spanish population: identification of putative novel steroid 21-hydroxylase deficiency alleles associated with the classic form of the disease. Human heredity. 1999 Jun ;49(3):169-75.
87. Nunez BS, Lobato MN, White PC, Meseguer A. Functional analysis of four CYP21 mutations from spanish patients with congenital adrenal hyperplasia. Biochemical and biophysical research communications. 1999 Sep ;262(3):635-7.
88. Ohlsson G, Müller J, Skakkebaek NE, Schwartz M. Steroid 21-hydroxylase deficiency: mutational spectrum in Denmark, three novel mutations, and in vitro expression analysis. Human mutation. 1999 Jan ;13(6):482-6.
89. Grischuk Y, Rubtsov P, Riepe FG, Grötzinger J, Beljelarskaia S, Prassolov V, et al. Four novel missense mutations in the CYP21A2 gene detected in Russian patients suffering from the classical form of congenital adrenal hyperplasia: identification, functional characterization, and structural analysis. The Journal of clinical endocrinology and metabolism. 2006 Dec ;91(12):4976-80.
90. Lajic S, Levo A, Nikoshkov A, Lundberg Y, Partanen J, Wedell A. A cluster of missense mutations at Arg356 of human steroid 21-hydroxylase may impair redox partner interaction. Human genetics. 1997 Jun ;99(6):704-9.
91. Lajic S, Nikoshkov A, Holst M, Wedell A. Effects of missense mutations and deletions on membrane anchoring and enzyme function of human steroid 21-hydroxylase (P450c21). Biochemical and biophysical research communications. 1999 Apr ;257(2):384-90.
92. Robins T, Bellanne-Chantelot C, Barbaro M, Cabrol S, Wedell A, Lajic S. Characterization of novel missense mutations in CYP21 causing congenital adrenal hyperplasia. Journal of molecular medicine (Berlin, Germany). 2007 Mar ;85(3):247-55.

93. Barbaro M, Baldazzi L, Balsamo A, Lajic S, Robins T, Barp L, et al. Functional studies of two novel and two rare mutations in the 21-hydroxylase gene. Journal of molecular medicine (Berlin, Germany). 2006 Jun ;84(6):521-8.
94. Baumgartner-Parzer SM, Schulze E, Waldhäusl W, Pauschenwein S, Rondot S, Nowotny P, et al. Mutational spectrum of the steroid 21-hydroxylase gene in Austria: identification of a novel missense mutation. The Journal of clinical endocrinology and metabolism. 2001 Oct ;86(10):4771-5.
95. Balsamo A, Cacciari E, Baldazzi L, Tartaglia L, Cassio A, Mantovani V, et al. CYP21 analysis and phenotype/genotype relationship in the screened population of the Italian Emilia-Romagna region. Clinical endocrinology. 2000 Jul ;53(1):117-25.
96. Pinto G, Tardy V, Trivin C, Thalassinos C, Lortat-Jacob S, Nihoul-Fékété C, et al. Follow-up of 68 children with congenital adrenal hyperplasia due to 21-hydroxylase deficiency: relevance of genotype for management. The Journal of clinical endocrinology and metabolism. 2003 Jun ;88(6):2624-33.
97. Wedell A, Ritzén EM, Haglund-Stengler B, Luthman H. Steroid 21-hydroxylase deficiency: three additional mutated alleles and establishment of phenotype-genotype relationships of common mutations. Proceedings of the National Academy of Sciences of the United States of America. 1992 Aug ;89(15):7232-6.
98. Nikoshkov A, Lajic S, Vlamis-Gardikas A, Tranebjaerg L, Holst M, Wedell A, et al. Naturally occurring mutants of human steroid 21-hydroxylase (P450c21) pinpoint residues important for enzyme activity and stability. The Journal of biological chemistry. 1998 Mar ;273(11):6163-5.
99. Lajić S, Clauin S, Robins T, Vexiau P, Blanché H, Bellanne-Chantelot C, et al. Novel mutations in CYP21 detected in individuals with hyperandrogenism. The Journal of clinical endocrinology and metabolism. 2002 Jun ;87(6):2824-9.
100. Wedell A, Luthman H. Steroid 21-hydroxylase (P450c21): a new allele and spread of mutations through the pseudogene. Human genetics. 1993 Apr ;91(3):236-40.
101. Krone N, Riepe FG, Grötzinger J, Partsch C-J, Brämswig J, Sippell WG. The residue E351 is essential for the activity of human 21-hydroxylase: evidence from a naturally occurring novel point mutation compared with artificial mutants generated by single amino acid substitutions. Journal of molecular medicine (Berlin, Germany). 2005 Jul ;83(7):561-8.
102. Stikkelbroeck NMML, Hoefsloot LH, Wijs IJ de, Otten BJ, Hermus ARMM, Sistermans EA. CYP21 gene mutation analysis in 198 patients with 21-hydroxylase deficiency in The Netherlands: six novel mutations and a specific cluster of four mutations. The Journal of clinical endocrinology and metabolism. 2003 Aug ;88(8):3852-9.
103. Robins T, Bellanne-Chantelot C, Barbaro M, Cabrol S, Wedell A, Lajic S. Characterization of novel missense mutations in CYP21 causing congenital adrenal hyperplasia. Journal of molecular medicine (Berlin, Germany). 2007 Mar ;85(3):247-55.
104. Krone N, Braun A, Roscher AA, Knorr D, Schwarz HP. Predicting phenotype in steroid 21-hydroxylase deficiency? Comprehensive genotyping in 155 unrelated, well defined patients from southern Germany. The Journal of clinical endocrinology and metabolism. 2000 Mar ;85(3):1059-65.

105. Lajić S, Robins T, Krone N, Schwarz HP, Wedell A. CYP21 mutations in simple virilizing congenital adrenal hyperplasia. Journal of molecular medicine (Berlin, Germany). 2001 Oct ;79(10):581-6.
106. Krone N, Riepe FG, Grötzinger J, Partsch C-J, Sippell WG. Functional characterization of two novel point mutations in the CYP21 gene causing simple virilizing forms of congenital adrenal hyperplasia due to 21-hydroxylase deficiency. The Journal of clinical endocrinology and metabolism. 2005 Jan ;90(1):445-54.
107. Krone N, Braun A, Roscher AA, Knorr D, Schwarz HP. Predicting phenotype in steroid 21-hydroxylase deficiency? Comprehensive genotyping in 155 unrelated, well defined patients from southern Germany. The Journal of clinical endocrinology and metabolism. 2000 Mar ;85(3):1059-65.
108. Wedell A, Ritzén EM, Haglund-Stengler B, Luthman H. Steroid 21-hydroxylase deficiency: three additional mutated alleles and establishment of phenotype-genotype relationships of common mutations. Proceedings of the National Academy of Sciences of the United States of America. 1992 Aug ;89(15):7232-6.
109. Kirby-Keyser L, Porter CC, Donohoue PA. E380D: a novel point mutation of CYP21 in an HLA-homozygous patient with salt-losing congenital adrenal hyperplasia due to 21-hydroxylase deficiency. Human mutation. 1997 Jan ;9(2):181-2.
110. Hsu NC, Guzov VM, Hsu LC, Chung BC. Characterization of the consequence of a novel Glu-380 to Asp mutation by expression of functional P450c21 in Escherichia coli. Biochimica et biophysica acta. 1999 Feb ;1430(1):95-102.
111. Helmberg A, Tusie-Luna MT, Tabarelli M, Kofler R, White PC. R339H and P453S: CYP21 mutations associated with nonclassic steroid 21-hydroxylase deficiency that are not apparent gene conversions. Molecular endocrinology (Baltimore, Md.). 1992 Aug ;6(8):1318-22.
112. Barbaro M, Lajic S, Baldazzi L, Balsamo A, Pirazzoli P, Cicognani A, et al. Functional analysis of two recurrent amino acid substitutions in the CYP21 gene from Italian patients with congenital adrenal hyperplasia. The Journal of clinical endocrinology and metabolism. 2004 May ;89(5):2402-7.
113. Zeng X, Witchel SF, Dobrowolski SF, Moulder PV, Jarvik JW, Telmer CA. Detection and assignment of CYP21 mutations using peptide mass signature genotyping. Molecular genetics and metabolism. 2004 May ;82(1):38-47.
114. Deneux C, Tardy V, Dib A, Mornet E, Billaud L, Charron D, et al. Phenotype-genotype correlation in 56 women with nonclassical congenital adrenal hyperplasia due to 21-hydroxylase deficiency. The Journal of clinical endocrinology and metabolism. 2001 Jan ;86(1):207-13.
115. Bojunga J, Welsch C, Antes I, Albrecht M, Lengauer T, Zeuzem S. Structural and functional analysis of a novel mutation of CYP21B in a heterozygote carrier of 21-hydroxylase deficiency. Human genetics. 2005 Oct ;117(6):558-64.
116. Kharrat M, Tardy V, M'Rad R, Maazoul F, Jemaa LB, Refaï M, et al. Molecular genetic analysis of Tunisian patients with a classic form of 21-hydroxylase deficiency: identification of four novel mutations and high prevalence of Q318X mutation. The Journal of clinical endocrinology and metabolism. 2004 Jan ;89(1):368-74.

117. Wedell A, Luthman H. Steroid 21-hydroxylase deficiency: two additional mutations in salt-wasting disease and rapid screening of disease-causing mutations. Human molecular genetics. 1993 May ;2(5):499-504.
118. Lajic S, Wedell A. An intron 1 splice mutation and a nonsense mutation (W23X) in CYP21 causing severe congenital adrenal hyperplasia. Human genetics. 1996 Aug ;98(2):182-4.
119. Levo A, Partanen J. Novel nonsense mutation (W302X) in the steroid 21-hydroxylase gene of a Finnish patient with the salt-wasting form of congenital adrenal hyperplasia. Human mutation. 1997 Jan ;9(4):363-5.
120. Krone N, Roscher AA, Schwarz HP, Braun A. Comprehensive analytical strategy for mutation screening in 21-hydroxylase deficiency. Clinical chemistry. 1998 Oct ;44(10):2075-82.
121. Witchel SF, Smith R, Suda-Hartman M. Identification of CYP21 mutations, one novel, by single strand conformational polymorphism (SSCP) analysis. Mutations in brief no. 218. Online. Human mutation. 1999 Jan ;13(2):172.
122. Lee HH, Chao HT, Lee YJ, Shu SG, Chao MC, Kuo JM, et al. Identification of four novel mutations in the CYP21 gene in congenital adrenal hyperplasia in the Chinese. Human genetics. 1998 Sep ;103(3):304-10.
123. Ordoñez-Sánchez ML, Ramírez-Jiménez S, López-Gutierrez AU, Riba L, Gamboa-Cardiel S, Cerrillo-Hinojosa M, et al. Molecular genetic analysis of patients carrying steroid 21-hydroxylase deficiency in the Mexican population: identification of possible new mutations and high prevalence of apparent germ-line mutations. Human genetics. 1998 Feb ;102(2):170-7.
124. Krone N, Braun A, Roscher AA, Schwarz HP. A novel frameshift mutation (141delT) in exon 1 of the 21-hydroxylase gene (CYP21) in a patient with the salt wasting form of congenital adrenal hyperplasia. Mutation in brief no. 255. Online. Human mutation. 1999 Jan ;14(1):90-1.
125. Billerbeck AE, Bachega TA, Frazatto ET, Nishi MY, Goldberg AC, Marin ML, et al. A novel missense mutation, GLY424SER, in Brazilian patients with 21-hydroxylase deficiency. The Journal of clinical endocrinology and metabolism. 1999 Aug ;84(8):2870-2.
126. Lau IF, Soardi FC, Lemos-Marini SH, Guerra Jr G, Baptista MT, De Mello MP. H28+C insertion in the CYP21 gene: a novel frameshift mutation in a Brazilian patient with the classical form of 21-hydroxylase deficiency. The Journal of clinical endocrinology and metabolism. 2001 Dec ;86(12):5877-80.
127. Loke KY, Lee YS, Lee WW, Poh LK. Molecular analysis of CYP-21 mutations for congenital adrenal hyperplasia in Singapore. Hormone research. 2001 Jan ;55(4):179-84.
128. Levo A, Partanen J. Novel mutations in the human CYP21 gene. Prenatal diagnosis. 2001 Oct ;21(10):885-9.
129. Billerbeck AEC, Mendonca BB, Pinto EM, Madureira G, Arnhold IJP, Bachega TASS. Three novel mutations in CYP21 gene in Brazilian patients with the classical form of 21-hydroxylase deficiency due to a founder effect. The Journal of clinical endocrinology and metabolism. 2002 Sep ;87(9):4314-7.
130. Koyama S, Toyoura T, Saisho S, Shimozawa K, Yata J. Genetic analysis of Japanese patients with 21-hydroxylase deficiency: identification of a patient with a new mutation of a homozygous deletion of adenine at codon 246 and patients without demonstrable mutations within the structural gene for CYP21. The Journal of clinical endocrinology and metabolism. 2002 Jun ;87(6):2668-73.

131. Dolzan V, Stopar-Obreza M, Zerjav-Tansek M, Breskvar K, Krzisnik C, Battelino T. Mutational spectrum of congenital adrenal hyperplasia in Slovenian patients: a novel Ala15Thr mutation and Pro30Leu within a larger gene conversion associated with a severe form of the disease. European journal of endocrinology / European Federation of Endocrine Societies. 2003 Aug ;149(2):137-44.
132. Usui T, Nishisho K, Kaji M, Ikuno N, Yorifuji T, Yasuda T, et al. Three novel mutations in Japanese patients with 21-hydroxylase deficiency. Hormone research. 2004 Jan ;61(3):126-32.
133. Ezquieta B, Oyarzábal M, Jariego CM, Varela JM, Chueca M. A novel frameshift mutation in the first exon of the 21-OH gene found in homozygosity in an apparently nonconsanguineous family. Hormone research. 1999 Jan ;51(3):135-41.
134. Ezquieta B, Cueva E, Oyarzábal M, Oliver A, Varela JM, Jariego C. Gene conversion (655G splicing mutation) and the founder effect (Gln318Stop) contribute to the most frequent severe point mutations in congenital adrenal hyperplasia (21-hydroxylase deficiency) in the Spanish population. Clinical genetics. 2002 Aug ;62(2):181-8.
135. Ezquieta B, Cueva E, Varela J, Oliver A, Fernández J, Jariego C. Non-classical 21-hydroxylase deficiency in children: association of adrenocorticotropic hormone-stimulated 17-hydroxyprogesterone with the risk of compound heterozygosity with severe mutations. Acta paediatrica (Oslo, Norway : 1992). 2002 Jan ;91(8):892-8.
136. Kharrat M, Tardy V, M'rad R, Maazoul F, Morel Y, Chaabouni H. A novel 13-bp deletion in exon 1 of CYP21 gene causing severe congenital adrenal hyperplasia. Diagnostic molecular pathology : the American journal of surgical pathology, part B. 2005 Dec ;14(4):250-2.
137. Di Pasquale G, Wasniewska M, Caruso M, Salzano G, Coco M, Lombardo F, et al. Salt wasting phenotype in a compound heterozygous girl with P482S mutation associated with anovel mutation of CYP21 gene (Q481P). Journal of endocrinological investigation. 2005 Dec ;28(11):1038-9.
138. Friães A, Rêgo AT, Aragüés JM, Moura LF, Mirante A, Mascarenhas MR, et al. CYP21A2 mutations in Portuguese patients with congenital adrenal hyperplasia: identification of two novel mutations and characterization of four different partial gene conversions. Molecular genetics and metabolism. 2006 May ;88(1):58-65.
139. Robins T, Carlsson J, Sunnerhagen M, Wedell A, Persson B. Molecular model of human CYP21 based on mammalian CYP2C5: structural features correlate with clinical severity of mutations causing congenital adrenal hyperplasia. Molecular endocrinology (Baltimore, Md.). 2006 Nov ;20(11):2946-64.
140. Krone N, Braun A, Weinert S, et al. Multiplex minisequencing of the 21-hydroxylase gene as a rapid strategy to confirm congenital adrenal hyperplasia. Clinical chemistry. 2002;48(6 Pt 1):818-25.
141. Fitness J, Dixit N, Webster D, et al. Genotyping of CYP21, linked chromosome 6p markers, and a sex-specific gene in neonatal screening for congenital adrenal hyperplasia. The Journal of clinical endocrinology and metabolism. 1999;84(3):960-6.
142. Stikkelbroeck NMML, Hoefsloot LH, Wijs IJ de, et al. CYP21 gene mutation analysis in 198 patients with 21-hydroxylase deficiency in The Netherlands: six novel mutations and a specific cluster of four mutations. The Journal of clinical

endocrinology and metabolism. 2003;88(8):3852-9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12915679
143. Koppens PF, Hoogenboezem T, Degenhart HJ. CYP21 and CYP21P variability in steroid 21-hydroxylase deficiency patients and in the general population in the Netherlands. European journal of human genetics : EJHG. 2000;8(11):827-36.
144. Oriola J, Plensa I, Machuca I, Pavía C, Rivera-Fillat F. Rapid screening method for detecting mutations in the 21-hydroxylase gene. Clinical chemistry. 1997;43(4):557-61.
145. Keen-Kim D, Redman JB, Alanes RU, et al. Validation and clinical application of a locus-specific polymerase chain reaction- and minisequencing-based assay for congenital adrenal hyperplasia (21-hydroxylase deficiency). The Journal of molecular diagnostics : JMD. 2005;7(2):236-46.
146. Barbaro M, Lajic S, Baldazzi L, et al. Functional analysis of two recurrent amino acid substitutions in the CYP21 gene from Italian patients with congenital adrenal hyperplasia. The Journal of clinical endocrinology and metabolism. 2004;89(5):2402-7.
147. Parajes S, Quinterio C, Domínguez F, Loidi L. A simple and robust quantitative PCR assay to determine CYP21A2 gene dose in the diagnosis of 21-hydroxylase deficiency. Clinical chemistry. 2007;53(9):1577-84.
148. Concolino P, Mello E, Toscano V, et al. Multiplex ligation-dependent probe amplification (MLPA) assay for the detection of CYP21A2 gene deletions/duplications in congenital adrenal hyperplasia: first technical report. Clinica chimica acta; international journal of clinical chemistry. 2009;402(1-2):164-70.
149. Griffin TJ, Smith LM. Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. Trends in biotechnology. 2000;18(2):77-84.
150. Amos J, Feldman GL, Grody WW, Monaghan K, Palomaki GE, Prior TW, Richards CS, Watson MS. Technical Standards and Guidelines for CFTR Mutation Testing, 2006 Edition. The Board of Directors of the American College of Medical Genetics
October 26, 2002. Genetics in Medicine 2002;3 (5).
151. Cumhur Cantürk, Niels Storm, Ulrike Baade, Ramona Salazar, Wolfgang Höppner. Complete genetic analysis of CYP21A2 and its pseudogene for diagnostics of CAH using MALDI-TOF MS**.** 54th Deutsche Gesellschaft für Endokrinologie (German Society of Endocrinology) Symposium 2011, PS2-06-5.
152. Parajes S, Quinteiro C, Domínguez F, Loidi L. High frequency of copy number variations and sequence variants at CYP21A2 locus: implication for the genetic diagnosis of 21-hydroxylase deficiency. PloS one. 2008;3(5):e2138.
153. Levo A, Partanen J. Mutation-haplotype analysis of steroid 21-hydroxylase (CYP21) deficiency in Finland. Implications for the population history of defective alleles. Human genetics. 1997;99(4):488-97.
154. Baumgartner-Parzer SM, Nowotny P, Heinze G, Waldhäusl W, Vierhapper H. Carrier frequency of congenital adrenal hyperplasia (21-hydroxylase deficiency) in a middle European population. The Journal of clinical endocrinology and metabolism. 2005;90(2):775-8.
155. Concolino P, Mello E, Toscano V, et al. Multiplex ligation-dependent probe amplification (MLPA) assay for the detection of CYP21A2 gene deletions/duplications in congenital adrenal hyperplasia: first technical report.

Clinica chimica acta; international journal of clinical chemistry. 2009;402(1-2):164-70.
156. He L, Vasiliou K, Nebert DW. Analysis and update of the human solute carrier (SLC) gene superfamily. Human genomics. 2009;3(2):195-206.
157. Kleinle S, Lang R, Fischer GF, et al. Duplications of the functional CYP21A2 gene are primarily restricted to Q318X alleles: evidence for a founder effect. The Journal of clinical endocrinology and metabolism. 2009;94(10):3954-8.

Lebenslauf

Name	Cantürk
Vorname	Cumhur
Geburtsdatum	17.02.1981
Geburtsort, -land	Uzunköprü, Türkei
09.1991 - 06.1996	Grundschule in Uzunköprü
09.1996 - 06.1999	Gymnasium in Uzunköprü
10.1999 - 06.2004	Studium an der Middle East Technical University in Ankara, Türkei Abschluss: Bachelor of Food Engineering
10.2004 - 05.2007	Studium an der Technischen Universität Hamburg-Harburg in Hamburg, Deutschland Abschluss: Master of Biotechnology
06.2007 - 12.2007	Werkstudent an der Bioglobe GmbH, Grandweg 64, 22529, Hamburg
seit 01.2008	Studium an der Technischen Universität Hamburg-Harburg in Zusammenarbeit mit der Bioglobe GmbH, Grandweg 64, 22529, Hamburg