



FATİH UNIVERSITY

The Graduate School of Sciences and Engineering

**Master of Science in
Biology**

SAMPLE SPINE

**EVALUATIONS OF CLINICALLY IMPORTANT SNPs
AND POLYMORPHISMS OF SEROTONIN
TRANSPORTER (SERT) GENE**

by

Hatice Kbra ŐENGL

July 2013

**EVALUATIONS OF CLINICALLY IMPORTANT SNPs
AND POLYMORPHISMS OF SEROTONIN
TRANSPORTER (SERT) GENE**



**EVALUATIONS OF CLINICALLY IMPORTANT SNPs
AND POLYMORPHISMS OF SEROTONIN
TRANSPORTER (SERT) GENE**

by

Hatice Kübra ŞENGÜL

A thesis submitted to

the Graduate School of Sciences and Engineering

of

Fatih University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Biology

July 2013
Istanbul, Turkey

APPROVAL PAGE

This is to certify that I have read this thesis written by Hatice Kübra ŞENGÜL and that in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science in Biology.

Assoc. Prof. Dr. Lokman ALPSOY
Thesis Supervisor

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Biology.

Assoc. Prof. Dr. Lokman ALPSOY
Head of Department

Examining Committee Members

Assoc. Prof. Dr. Lokman ALPSOY

Assoc. Prof. Dr. Sevim Işık

Assist. Prof. Dr. Mustafa PETEK

It is approved that this thesis has been written in compliance with the formatting rules laid down by the Graduate School of Sciences and Engineering.

Assoc. Prof. Dr. Nurullah ARSLAN
Director

July 2013

EVALUATIONS OF CLINICALLY IMPORTANT SNPs AND POLYMORPHISMS OF SEROTONIN TRANSPORTER (SERT) GENE

Hatice Kübra ŞENGÜL

M.S. Thesis – Biology
July 2013

Thesis Supervisor: Assoc. Prof. Dr. Lokman ALPSOY

ABSTRACT

Serotonin neurotransmitter regular synthesis is not sufficient for proper function which plays a role in the etiology of many stress related disorders such as depression, bipolar disorder, schizophrenia, obsessive-compulsive disorder, anxiety disorder, ADHD, alcohol and drug addiction, social phobia, Alzheimer's disease, suicide behavior, eating disorder, migraine, sexual dysfunction, irritable bowel sendrom, hypertension, myocardial infarction, nausea and vomiting. Various gene polymorphisms of SERT protein, receptors of SERT neurotransmitters and reuptake mechanisms are searched with related many diseases mentioned above. SERT gene contains approximately 2645 SNPs. Exclusively 5 to 10 of 2645 SNPs and genetic variations are clinically significant that is exhibited with researches.

The aim of this study is to accelerate epidemiological studies of these clinically significant differences with HRM analysis of real-time PCR method which is much faster cheaper and reliable way as they are intended to be identified than the methods currently in used. In this thesis, clinically significant and also in use of clinical and pharmacogenetical rs25533 SNP and 5HTT-VNTR polymorphism in Turkish population from SERT gene were studied with volunteer's DNA. Study was performed with primers used in literature that has been optimized for HRM analysis. DNA sequence analysis and KASP methods were used for validation. As a result of this study, HRM analysis is a method of convenient and advantageous for the detection of polymorphisms in the validation were completed.

Keywords: Serotonin transporter gene (SERT/5HTT/SLC6A4), clinically important SNPs, rs25533, 5HTT-VNTR, High Resolution Melting Analysis (HRM), KASP.

SEROTONİN TAŞIYICI GENİNDE (SERT) KLİNİK ÖNEME SAHİP GENETİK DEĞİŞİMLERİN DEĞERLENDİRİLMESİ

Hatice Kübra ŞENGÜL

Yüksek Lisans Tezi – Biyoloji
Temmuz 2013

Tez Danışmanı: Assoc. Prof. Dr. Lokman ALPSOY

ÖZ

Depresyon, bipolar bozukluk, şizofreni, obsesif kompulsif bozukluk anksiyete, dikkat eksikliği ve hiperaktivite bozukluğu, alkol ve madde bağımlılığı, sosyal fobi, Alzheimer, intihar davranışı, yeme bozuklukları, migren, cinsel işlev bozukluğu, irritabl barsak sendromu, hipertansiyon, miyokard infarktüsü, bulantı ve kusma gibi birçok stres bağlantılı bozukluğun etiolojisinde rolü olan serotonin nörotransmitterinin düzgün fonksiyonunu yapmasında sadece uygun bir şekilde sentezlenmesi yetmemektedir. Bu nörotransmitterin reseptörleri ve geri alımın mekanizmasından sorumlu SERT proteini ve bu proteinin çalışmasını bozan çeşitli gen polimorfizmleri yukarıda geçen hastalıklarla ilgili olarak araştırılmaktadır. SERT geninde yaklaşık 2645 SNP'den klinik öneme sahip olduğu araştırmalarca ortaya konmuş 5-10 arası SNP ve genetik değişim bulunmaktadır.

Çalışmamızda epidemiyolojik çalışmaları hızlandıracak şekilde klinik öneme sahip bu değişimlerin real-time PCR HRM analizi yöntemi ile halihazırda kullanımdaki yöntemlere oranla çok daha hızlı, ucuz ve onlar kadar güvenilir bir şekilde tespit edilmeleri amaçlanmıştır. Çalışmamızda klinik öneme sahip, Türk populasyonu açısından da klinik ve farmakogenetik kullanımları olan SERT gen SNP ve polimorfizmlerden rs25533 ve 5HTT-VNTR, literatürde kullanılan primerler ile HRM analizi için optimize edilmek şekliyle gönüllü DNAları ile çalışılmıştır. Validasyon için DNA dizi analizi ve KASP yöntemleri kullanılmıştır. Çalışmamızın sonucunda, HRM analizinin polimorfizm tespiti için uygun ve avantajlı bir yöntem olduğu validasyon da yapılarak gösterilmiştir.

Anahtar Kelimeler: Serotonin taşıyıcı geni (SERT/5-HTT/SLC6A4), klinik öneme sahip SNPler, rs25533, 5HTT-VNTR Yüksek Çözünürlükte Erime Analizi (HRM), KASP.

To Assist. Prof. İffet İrem UZONUR

ACKNOWLEDGEMENT

Foremost, I express sincere appreciation, deep gratitude to Assist. Prof. Dr. İffet İrem UZONUR for her never failing encouraging and all kind support through this thesis also in all respects of my life. This thesis is started by her genius. She will be my gorgeous guide rest of my life. God rest her soul.

I am grateful to my friends Ecehan Özge and Derya Sultan for their moral support.

I thankfully acknowledge that this research is supported by Scientific Research Fund of Fatih University financially under the project number of P50031201_G (2543).

I thank to TÜBİTAK for their economic support.

Lastly, I express my thanks to my family.

TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZ.....	iv
DEDICATION.....	v
ACKNOWLEDGMENT.....	vi
TABLE OF CONTENTS.....	vii
LIST OF TABLES.....	ix
LIST OF FIGURES.....	x
LIST OF SYMBOLS AND ABBREVIATIONS.....	xii
CHAPTER 1 INTRODUCTION.....	1
1.1 Background.....	1
1.2 Purpose of study.....	4
CHAPTER 2 LITERATURE REVIEW.....	5
2.1 SNP (Single Nucleotide Polymorphism) and Polymorphism.....	5
2.1.1 SNP.....	6
2.1.2 Variable Number of Tandem Repeats (VNTR).....	9
2.2 SERT (Serotonin Transporter) Gene.....	10
2.2.1 Clinically Important SNPs and Polymorphisms in SERT.....	11
2.3 Methods for SNP Detection and Polymorphism Detection in SERT Gene.....	12
2.3.1 Real Time PCR.....	12
2.3.2 High Resolution Melting Analysis (HRM Analysis).....	13
2.3.3 KASP Analysis.....	14
CHAPTER 3 METHODOLOGY PROCESS.....	16
3.1 Materials.....	16
3.1.1 Samples.....	16
3.1.2 Chemicals.....	16
3.1.2.1 HRM and KASP Chemicals.....	16
3.1.2.2 Primers.....	17

3.1.2.3 Softwares	17
3.1.3 Instruments	18
3.2 Method	18
3.2.1 HRM Analysis	18
3.2.2 KASP Analysis	20
3.2.3 Sequencing for confirmation	22
CHAPTER 4 RESULTS	23
4.1 HRM Results	23
4.1.1 rs25533 results	23
4.1.2 5HTT-VNTR results	26
4.2 Validation Results	31
CHAPTER 5 DISCUSSION	37
REFERENCES	42

LIST OF TABLES

TABLE

3.1	The designed primers used in multiplex HRM analysis.....	17
3.2	List of instruments.....	18
3.3	The reagents and amounts for HRM kit used in study.....	19
3.4	HRM conditions for rs25533.....	19
3.5	HRM conditions for 5HTT-VNTR.....	20
3.6	The reagents for KASP version 4.0 genotyping.....	21
3.7	Thermal cycling conditions for KASP version 4.0 genotyping.....	21
4.1	SNP genotyping results for 20 samples.....	36
4.2	VNTR genotyping results for 20 samples.....	36
5.1	Comparisons for the different SNP detection methods' prices.....	39

LIST OF FIGURES

FIGURE

2.1	Polymorphism types	5
2.2	Single Nucleotide Polymorphism (SNP).....	6
2.3	Variable Number of Tandem Repeat (VNTR)	9
2.4	Serotonin Transporter (SERT) gene.....	10
2.5	Serotonin Transporter (SERT) gene SNPs	11
4.1	Amplification curves of all 20 samples for rs25533 analysis on HRM show that HRM has been done well.....	24
4.2	Normalisation curves of all 20 samples for analysis rs25533 on HRM show similarities of samples	24
4.3	Melting curves of some of 20 samples for analysis rs25533 on HRM.	25
4.4	T _m of all 20 samples for analysis rs25533 samples on HRM show proximity of melting temperatures	25
4.5	HRM genotyping results of all 20 samples for analysis rs25533 on HRM show affinity ratios.....	26
4.6	Amplification curves of same 12 samples for 5HTT-VNTR analysis on HRM show that HRM has been done well.....	26
4.7	Melting curves of same 12 samples for analysis 5HTT-VNTR on HRM.....	27
4.8	T _m of same 12 samples for analysis rs25533 samples on HRM show proximity of melting temperatures.....	27
4.9	Normalisation curves of same 12 samples for analysis rs25533 on HRM show similarities of samples	28
4.10	HRM genotyping results of same 12 samples for analysis rs25533 on HRMs show affinity ratios.....	28
4.11	Amplification curves of 8 samples for rs25533 analysis on HRM show that HRM has been done well.....	29
4.12	Normalisation curves of same 12 samples for analysis rs25533 on HRM show	

similarities of samples	29
4.13 Melting curves of same 8 samples for analysis 5HTT-VNTR on HRM.....	30
4.14 T _m of same 8 samples for analysis rs25533 samples on HRM show proximity of melting temperatures	30
4.15 HRM genotyping results of same 8 samples for analysis rs25533 on HRM show affinity ratios.....	31
4.16 Genotyping results of KASP analysis of rs25533 show all 20 samples have same genotypes	32
4.17 KASPar version 4.0 genotyping graph for rs25533 SNP shows all 20 samples have same genotype. Sequencing results of all 20 samples show all samples are CC genotype	32
4.18 Sequencing results of all 20 samples which show all samples are CC genotype ..	33
4.19 Sequencing results of same 12 samples of 5HTT-VNTR of chromatograph analyzing program (Finch TV 1.4.0). The results show that these 12 samples include 7 repeats and they are 7/7 genotype according to 5HTT-VNTR gene	34
4.20 Sequencing results of same 12 samples of 5HTT-VNTR of chromatograph analyzing program (Finch TV 1.4.0). The results show that these 8 samples include 7 repeats and they are 9/9 genotype according to 5HTT-VNTR gene	35

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOL/ABBREVIATION

ADHD	Attention deficit hyperactivity disorder
CNS	Central nervous system
dbSNP	The Single Nucleotide Polymorphism Database
DNA	Deoxyribonucleic acid
dsDNA	double-stranded DNA
FRET	Fluorescence resonance energy transfer
GMO	Genetically modified organism
HRM	High resolution melting
KASP	KBioscience Competitive Allele-Specific PCR
Kb	Kilobase
LD	Linkage disequilibrium
mRNA	Messenger RNA
PCR	Polymerase chain reaction
rs	SNP ID number
RFLP	Restriction fragment length polymorphisms
SERT	Serotonin transporter gene
SLC6A4	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4: Serotonin transporter gene
SES	Socioeconomic status
SNP	Single nucleotide polymorphisms
SSRP	Simple sequence repeats polymorphism
STin2	A variable number of tandem repeats (VNTR) in the functional intron 2
TM	Transmembrane
T _m	Melting temperature
VNTR	A variable number tandem repeat

3' UTR	3' Untranslated Region
5HT	5-hydroxytryptamine: Serotonin
5HTT	Serotonin transporter gene
5HTTLPR	5HTT gene-linked polymorphic region

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

At the present time, we are at a stage where we can read nearly the entire genetic code of the human genome, as recently a rough draft of the human genome sequence has been determined (Venter et. al 2001). It represents the sequence of A, G, C, and T letters that are symbols for nucleotides. The specific sequence of these nucleotides constitutes all our genes that have specific characteristics and expression in the human being. These genes are responsible for various physical, physiological and pharmacological activities in the body. Since mutations in these genes are often the cause of many heritable diseases, it is sometimes necessary to find specific genetic mutations responsible for a particular disease. This is one of the current aims of the Human Genome Project (Pungliya, 2001).

99.9 % of human genome is the same and the difference between humans is made by remained 0.1 % of genome (Brown 2000). The occurrence of Single nucleotide polymorphism (SNP) is approximately one in every 1000-2000 base pairs, and the total number of SNPs in the human genome estimated by November 2000 is 1.433.393 (The International SNP Map Working Group 2001) and 2.104.820 (Venter et. al 2001). These polymorphisms are present in coding as well non-coding regions less than 1% of all the SNPs are present in the protein-coding regions of the genome (Venter et. al 2001). This suggests that a very small proportion of SNPs may be responsible for phenotypic variation.

In this situation, SNPs, have important role because most sequence variations are constituted by SNPs. SNPs are single base substitutions and they affect molecular world from forensic science to cancer. And if the amount of SNP (approximately 50 million)

is estimated importance of detection of SNP can be easily understood. Because of increased interest to SNPs, popularity of detection methods always increases.

There are a lot of techniques to detect SNPs and polymorphisms fluorescence resonance energy transfer (FRET) technology, restriction endonuclease analysis, restriction digestion, denaturing HPLC, chip technologies, denaturing gradient gel electrophoresis, denaturing capillary electrophoresis, automated constant denaturant capillary electrophoresis, temperature denaturant capillary electrophoresis, allele specific PCR, allele specific nucleotide hybridization, PCR-reverse dot blot, activated RAS-GTP specific biosensor, oligonucleotide hybridization assay, mass spectrometry, peptide nucleic acid PCR, array based techniques and single-strand conformation polymorphism analysis, are some of these detection technologies. All these methods have their advantages and disadvantages.

Reliability, sensitivity, specificity, fastness and cheapness of real-time PCR based High Resolution Melting (HRM) and KASP Analysis make them advantageous in the field. On the other hand, gel-electrophoresis, or DNA sequencing can still be performed after HRM analysis. These characteristics make HRM analysis ideal for use in routine diagnostic settings. Due to its numerous advantages, HRM analysis has been widely applied in diagnostic laboratories for screening for disease-associated mutations or DNA sequence variations such as polymorphisms and germline mutations. Besides HRM analysis is an attractive screening method for the detection of both known and unknown mutations as it is rapid to set up and inexpensive to operate (Do et al, 2008). Although its advantages, there is no standard method for usage of HRM. So the standardization of HRM analysis for detection of SNP improves originality of the study for the serotonin transporter (SERT) gene and clinically important SNPs and polymorphisms.

The neurotransmitter serotonin is involved in the regulation of a broad array of psychological, behavioral, and biological functions that when dysregulated influence the development and course of an equally broad range of psychiatric and medical disorders. Low central nervous system (CNS) serotonergic function has been implicated, for example, in the mediation of negative emotions, impulsive aggressive behavior, increased use of alcohol and nicotine, increased food consumption, increased sympathetic and decreased parasympathetic nervous system outflow, and altered neuroendocrine function. Peripherally, serotonin receptors and its transporter on platelets, immune system cells, cardiovascular system cells, and gastrointestinal system cells play an important role in regulating the functions of these

organs and systems. These broad-ranging effects of serotonin have led to the hypothesis that altered CNS serotonergic function is a driver of the clustering of negative moods, risky health behaviors, and altered biological functions that increase disease risks in certain individuals and groups for example, lower socioeconomic status (SES) (Williams et al, 2003).

SERT in brain and in many peripheral tissues is responsible for the active transport of serotonin into neurons, enterochromaffin cells, platelets, and other cells. SERT has a critical role in the homeostatic regulation of the magnitude, duration, and spatial distribution of signals reaching serotonin receptors. But at the structural level, SERT's function as a transporter protein in membranes remains incompletely understood (Murphy et al, 2004).

A recent explosion of molecular genetic studies and clinical findings has brought the serotonin (also termed 5HT) transporter into spotlight. Since the human serotonin transporter gene (SLC6A4, SERT, 5HTT) was sequenced and cloned a decade ago. The highly evolutionarily conserved SERT regulates the entire serotonergic system and its receptors via modulation of extracellular fluid serotonin concentrations. Differences in SERT expression and function produced by SERT genes and their variants show associations with multiple human disorders. (Murphy et al, 2004). Such as, depression, bipolar disorder, schizophrenia, obsessive-compulsive disorder, anxiety disorder, ADHD, alcoholism, social phobia, Alzheimer's disease, suicide behavior, eating disorder, migraine, irritable bowel sendrom, hypertension, myocardial infarction.

The research which is done by keywords in theses and studies in universities shows us that there are a few of studies with similar SNPs and polymorphisms and there is not any study with methods and polymorphisms which our interest. Studies in Turkey are also like this situation. The SNP and polymorphisms which our interests had been chosen by researches on universal literature and dbSNP database (The Single Nucleotide Polymorphism Database). Acording to the literature, these SNP and polymorphisms are related with some disease but there is no any evaluation study on case or HRM method which we aim to become prevalent. In this mean our study contributes to the native and foreign literature.

Genome-wide studies on serotonin neurotransmitters, particularly interesting on many issues related to the nervous system, and studies have shown that genotype - phenotype connections in the etiology of many diseases and related genes of this neurotransmitter SERT has been around. While international efforts to study at home, and this is a very limited group

studied genes and polymorphisms could not be expanded in a plane. Never before been studied in this work in our country, or try to keep certain groups and SNP polymorphisms studied and the method used in the study of polymorphisms in the world literature and to the best of our knowledge this is the first time SNPs being used extensively in both the impact and originality of the work is extremely important in terms of value added.

The method applied in this method, especially in the HRM and sequencing studies, the availability of this type of pairing and allelic discrimination features a very easy, cheap and reliable method that may be unknown, but the literature and this study, we will work to implement optimizations and then to work with these methods is currently will allow for the dissemination of all SNPs independently selected SNPs. This is also to advance the science of clinical trials and the impact and added value to be the most important common.

1.2 PURPOSE of STUDY

In this thesis, we aim to study clinically important polymorphisms of SERT gene by our opportunities, fast, reliable and cheapest way. The ultimate purpose of the thesis is contributing to epidemiological studies and adding value to polymorphisms of this important gene clinically and pharmacogenetically by data of the thesis.

CHAPTER 2

LITERATURE REVIEW

2.1 SNP (SINGLE NUCLEOTIDE POLYMORPHISM) and POLYMORPHISM

Since all polymorphism has a genetic basis, genetic polymorphism has a particular meaning: Genetic polymorphism is the simultaneous occurrence in the same locality of two or more discontinuous forms in such proportions that the rarest of them cannot be maintained just by recurrent mutation or immigration (Ford, 1940).

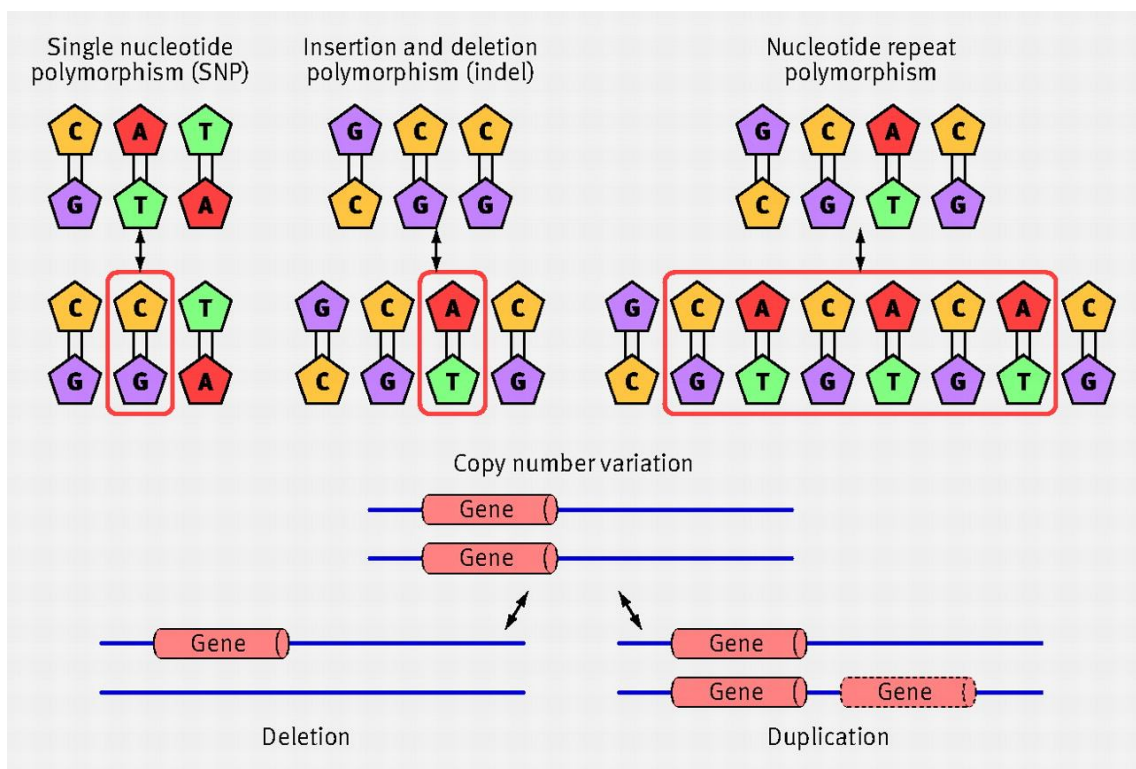


Figure 2.1 Polymorphism types (Hingorani et. al, 2010).

2.1.1 SNP

SNPs are the most common single base variations in the human population. A SNP is a variation where two alternative bases occur at appreciable frequency (the frequency of each base is above 1% in a population) (Lander et al., 1998). Most of these variants are neutral, but some are functional. One of the important goals of genetic analysis is to identify those SNPs and SNP variants (alleles), which are associated with a disease. For example, consider the following nucleotide sequence.

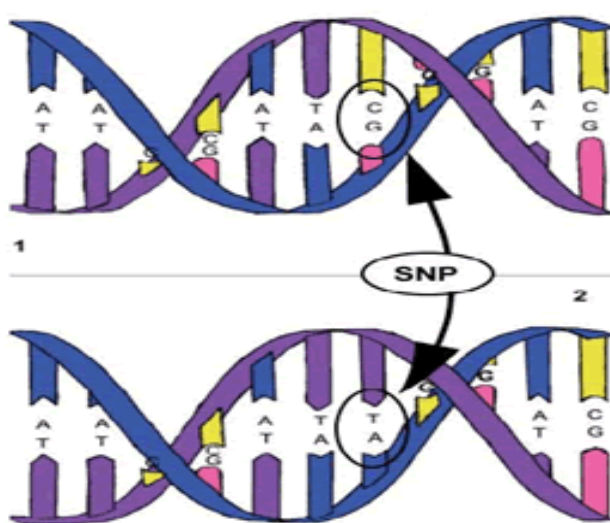


Figure 2.2 Single Nucleotide Polymorphism (SNP) <http://www.med.umich.edu/cancer/prevention/summer-11-three-pcgp.shtml> (University of Michigan Comprehensive Cancer Center, 2011).

The change from C to T is considered a SNP provided both alleles are present in more than 1 percent in the population. SNPs are often binary, i.e. they most often have only two alleles. They are less susceptible to mutations than microsatellite repeat markers. A microsatellite is a short sequence of repeated nucleotides in a genome e.g.: AATGAATGAATG-----, where AATG is repeated a variable number of times. Due to their stability, SNPs are very useful for studying human evolutionary history (Pungliya, 2001).

When any two human genomes are compared side by side, they are 99.9% identical (Cooper et al., 1985). However, with a 3.2 billion base pair genome, each person harbors some 3.2 million differences in his/her diploid genome. Most of the differences are due to single base substitution polymorphisms, popularly known as SNPs. While the majority of the SNPs are of no biological consequence, a fraction of the substitutions have functional significance and are the basis for the diversity found among humans (Collins et al., 1997). As genetic markers, SNPs can be used to follow the inheritance patterns of chromosomal regions from generation to generation and are powerful tools in the study of genetic factors associated with human diseases (Johnson and Todd, 2000; Risch, 2000).

The main issue with global SNP discovery is that on average, there is one SNP in every 1.000 bp of DNA when two human genomes are compared to each other (The International SNP Map Working Group, 2001). To maximize the chance of finding SNPs, one must be able to scan 1.000 bp pieces of DNA in a generic way. The first attempt to identify SNPs randomly in the human genome was to scan for alterations in restriction sites in the genome (Botstein et al., 1980). Although the actual sequence variation was not determined, restriction fragment length polymorphisms (RFLPs) were the first SNPs found in a random, global approach (Donis-Keller et al., 1987). Developed before the polymerase chain reaction (PCR) was even conceived, RFLP analysis was a powerful, but very laborious, strategy. High quality genomic DNA from multiple individuals were cut with a number of restriction enzymes, separated by gel electrophoresis and transferred to nylon filters. These Southern Blots were then probed with random genomic clones to identify variations in the restriction fragment lengths. Because only small amounts of DNA were found on the filter, radioactive labels were used with the probes. The situation did not improve with the advent of PCR (Saiki et al., 1985). This was due to the fact that in order for PCR to work, DNA sequence data had to be obtained to design loci-specific PCR primers. To complicate things further, DNA sequencing and oligonucleotide synthesis were both costly activities until quite recently. Global scanning of SNPs was not seriously attempted in the early 1990's because simple sequence repeat polymorphisms (SSRPs), the highly informative microsatellite markers, were much more easily developed by probing genomic libraries with synthetic oligonucleotides bearing the simple sequence repeat motifs and sequencing the positive clones (Weber and May, 1989; Litt and Luty, 1989).

Nowadays, The International SNP Map Working Group describes a map of 1.42 million SNPs distributed throughout the human genome, providing an average density on available sequence of one SNP every 1.9 kilobases. These SNPs were primarily discovered by two projects: The SNP Consortium and the analysis of clone overlaps by the International Human Genome Sequencing Consortium. The map integrates all publicly available SNPs with described genes and other genomic features. We estimate that 60,000 SNPs fall within exon (coding and untranslated regions), and 85% of exons are within 5 kb of the nearest SNP. Nucleotide diversity varies greatly across the genome, in a manner broadly consistent with a standard population genetic model of human history. This high-density SNP map provides a public resource for defining haplotype variation across the genome, and should help to identify biomedically important genes for diagnosis and therapy (The International SNP Map Working Group, 2001).

In contrast to studies of disease that can take advantage of familial inheritance, homogeneous populations, and relatively straightforward ascertainment of affected individuals, options for pharmacogenomic analyses are limited. As drug response is a trait whose expression is mediated through the administration of a therapeutic compound, “responders” can only be identified after they receive the drug. This makes ascertainment of individuals from a general population setting difficult and precludes the use of families, except in the rare instance where multiple family members are given a drug. Whereas studies of disease can take advantage of homogeneous, isolated populations to increase linkage disequilibrium (LD) and thus the likelihood of finding the disease gene²⁰, clinical trials are likely to be the main source of patients for pharmacogenomic studies. Although drugs on the market may be distributed worldwide, most clinical trials of new therapies are performed in Caucasian Americans or Europeans. Pharmacogenomic studies will most likely be carried out in these genetically heterogeneous clinical trial populations using case-control or cohort association study designs employing either candidate gene or LD mapping approaches (McCarthy and Hilfiker, 2000).

SNPs can be followed from generation to generation as an inheritance pattern (Kwok and Chen, 2003) so they can be used as a genetic marker like other molecular markers. Because of their abundance and ease of detection, they are studied much more

especially in association studies and LD mapping (Dawson et al., 2002). Most of SNPs have no biological consequences, a part of them causes functional differences (Kwok and Chen, 2003). When they are found in a functional gene region, they can affect gene expression and protein 3D structures. These effects can be given rise to diseases or phenotypic differences. As molecular markers, they can be followed for phenotypes of interest. And finally, they are beneficial for mutation rate detection and evolution studies (Salisbury et al., 2003). Disease susceptibility and drug responses are the most common reasons of SNP usage as molecular markers in studies and projects (Dawson et al., 2002).

2.1.2 Variable Number of Tandem Repeats (VNTR)

A tandem repeat is a short sequence of DNA that is repeated in a head-to-tail fashion at a specific chromosomal locus. Tandem repeats are interspersed throughout the human genome. Some sequences are found at only one site -- a single locus -- in the human genome. For many tandem repeats, the number of repeated units vary between individuals. Such loci are termed VNTRs (Calderon, 2009).

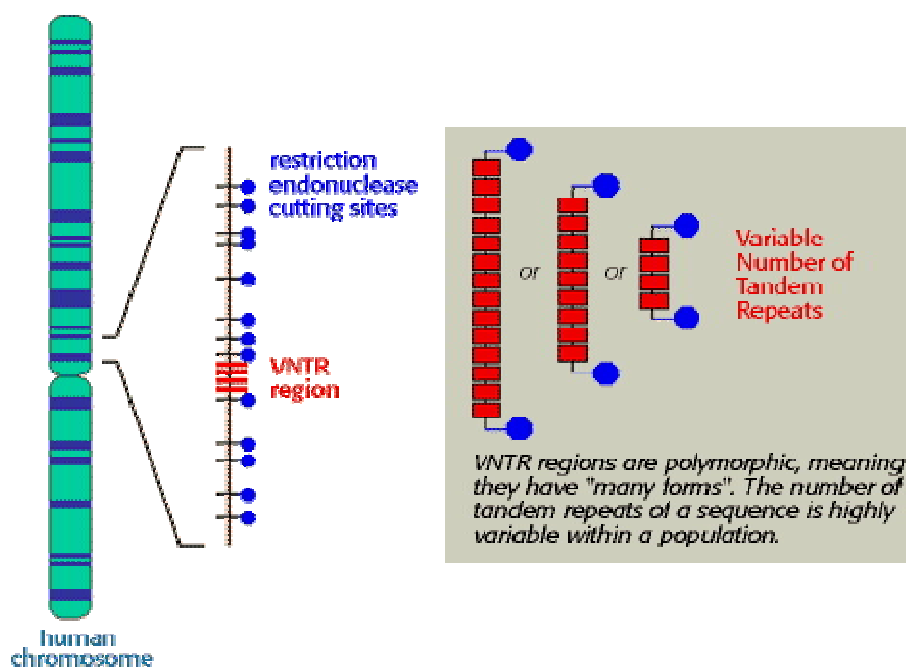


Figure 2.3 Variable Number of Tandem Repeat (VNTR) [http://www.biology.arizona.edu/human bio/problem sets/dna forensics 1/05t.html](http://www.biology.arizona.edu/human_bio/problem_sets/dna_forensics_1/05t.html) (University of Arizona, 1996).

2.2 SERT (SEROTONIN TRANSPORTER) GENE

Human SERT spans 37.8 kb on chromosome 17q11.2, and is composed of fourteen exons encoding a protein of 630 amino acids (see Figure 2.2) (21, 22). Alternate promoters in combination with differential splicing involving exon 1A, B, and C in specific tissues, and alternate polyadenylation site usage resulting in multiple mRNA species are likely participants in the regulation of SERT expression in humans (see Figure 2.4) (Murphy, 2004).

In addition to several regulatory domains controlling selective expression in serotonergic neurons, transcriptional activity of human SERT is modulated by a repetitive element of varying length in the 5' flanking region located ~1.4kb upstream of the transcription start site, termed the 5HTTLPR gene-linked polymorphic region (5HTTLPR) (Murphy, 2004).

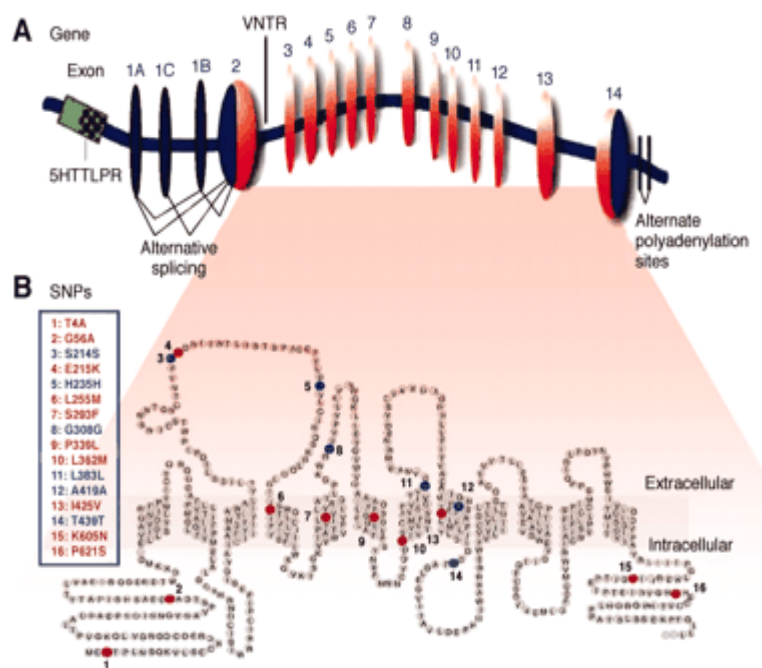


Figure 2.4 Serotonin Transporter (SERT) gene: (A) SERT gene organization showing coding (orange) exons and non-coding and intronic areas (blue), locations of the 5HTTLPR polymorphism ~1.4 kb upstream and location of the intronic variable number tandem repeat (VNTR) near exon 2, (B). Translated (red) and non-translated (blue) SNP sites are annotated in the 630 amino acids of the 12-transmembrane (TM)-containing SERT protein (Murphy et. al, 2004).

2.2.1 Clinically Important SNPs and Polymorphisms in SERT

Three additional variations have been described in untranslated regions: 1) a ~380bp deletion between the 5HTTLPR and the transcription start site which may be unstable in vivo, conferring mosaicism with cell or tissue differences in expression of SERT, and possibly involving 5HTTLPR-dependent mechanisms; 2) a variable number tandem repeat (VNTR) within intron 2; and 3) a G to T transversion in a putative adenylation site in the 3' Untranslated Region (3' UTR) region (Murphy, 2004).

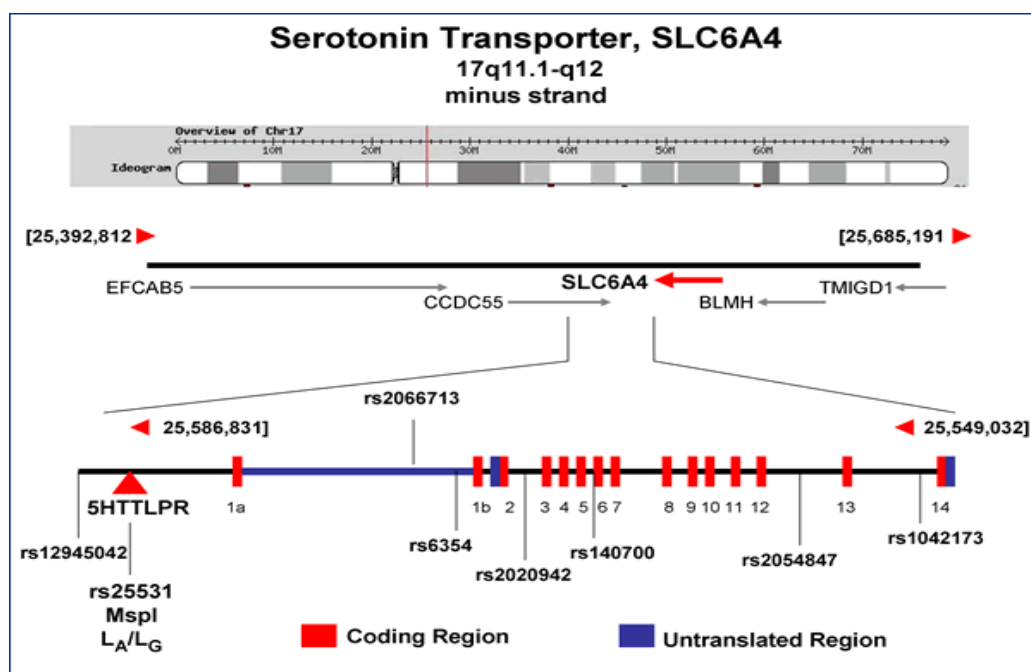


Figure 2.5 Serotonin Transporter (SERT) gene SNPs http://ibgwww.colorado.edu/genotyping_lab/pdf/5HTTLPR.pdf (University of Colorado Boulder, 2007).

Another functional polymorphism of this gene (STin2 VNTR) a 17 bp VNTR with four different alleles that correspond to the number of tandem repeats has been described. SERT polymorphisms characterised was the VNTRs in intron 2 of the gene and consists of 17 base pair (bp) repeats (Yeo et al., 2004). This region repeats 7, 9, 10, 12 times. The genotypes of this polymorphism are 12/12, 12/10, 10/10, 12/9, 10/9, 9/9,

12/7, 10/7, 9/7, 7/7 (Erdal et al., 2000). This polymorphic VNTR region acts as a transcriptional regulator (Florez et al., 2008).

The 5'LPR and a second VNTR within intron 2 termed Stin2 have been hypothesized as predisposing genetic factors for several neurological disorders (Haddley et al., 2008), although conflicting evidence has been reported in the literature (Lotrich and Pollock 2004). Both the 5' LPR and Stin2 VNTRs are transcriptional regulatory domains, and therefore, modulation of their function could affect the levels of SLC6A4 expression in a tissue-specific and stimulus inducible manner (Haddley et al., 2008).

2.3 METHODS for SNP DETECTION and POLYMORPHISM DETECTION in SERT

In the genotyping methods, there are at least two steps. Firstly PCR is done because PCR amplification is required for the continuation of the process. Then, other post-PCR processes like gel electrophoresis, matrix-gel separation are applied. Restriction fragment length polymorphisms (RFLP), temperature gradient gel electrophoresis (TGGE), denaturing gradient gel electrophoresis (DGGE), reverse dot blot hybridization, amplification refractory mutation system (ARMS) and direct sequencing can be given as examples for traditional polymerase chain reaction (PCR)-based methods. With developing technology, both PCR amplification and genotyping can be done in the same tube as closed-tube homogenous assays (Marashi et al.,2012).

2.3.1 Real Time PCR

Real-time PCR and RT-PCR are highly sensitive techniques enabling amplification and quantification of a specific nucleic acid sequence with detection of the PCR product in real time. Quantification of DNA, cDNA, or RNA targets can be easily achieved by determination of the cycle when the PCR product can first be detected. This is in contrast with endpoint detection in conventional PCR, which does not enable accurate quantification of nucleic acids. Real-time PCR is highly suited for a wide range of applications, such as gene expression analysis, determination of viral

load, detection of genetically modified organisms (GMOs), SNP genotyping, and allelic discrimination.

In one-step RT-PCR also referred to as one-tube RT-PCR both reverse transcription and amplification take place in the same tube, with reverse transcription preceding PCR. This is possible due to specialized reaction chemistries and cycling protocols. The fast procedure enables rapid processing of multiple samples and is easy to automate. The reduced number of handling steps results in high reproducibility from sample to sample and minimizes the risk of contamination since less manipulation is required. The advantages of one step RT-PCR method are easy handling, fast procedure, high reproducibility, low contamination risk (Reuter et. al, 2005).

Especially for SNP genotyping, real-time PCR is excellent providing specificity, sensitivity and cost effectiveness. High resolution melting analysis, hybridization probes based on melting curve analysis, hydrolysis probes, molecular beacons, scorpions are some of the techniques used in real-time PCR.(Karabulut, 2012) Mass spectrometry, RFLP, SSCP and real-time PCR is used in SNP detection, forensic, pharmacogenetics, clinic microbiology (Daştan, 2006).

2.3.2 High Resolution Melting Analysis (HRM Analysis)

High Resolution Melting Analysis (HRM or HRMA) is a recently developed technique for fast, high-throughput post-PCR analysis of genetic mutations or variance in nucleic acid sequences. It enables researchers to rapidly detect and categorize genetic mutations (e.g. single nucleotide polymorphisms (SNPs)), identify new genetic variants without sequencing (gene scanning) or determine the genetic variation in a population (e.g. viral diversity) prior to sequencing (KAPA Biosystems).

The first step of the HRM protocol is the amplification of the region of interest, using standard PCR techniques, in the presence of a specialized double-stranded DNA (dsDNA) binding dye. This specialized dye is highly fluorescent when bound to dsDNA and poorly fluorescent in the unbound state. This change allows the user to monitor the DNA amplification during PCR (as in quantitative PCR) (KAPA Biosystems).

After completion of the PCR step, the amplified target is gradually denatured by increasing the temperature in small increments, in order to produce a characteristic

melting profile; this is termed melting analysis. The amplified target denatures gradually, releasing the dye, which results in a drop in fluorescence. When set up correctly, HRM is sensitive enough to allow the detection of a single base change between otherwise identical nucleotide sequences (KAPA Biosystems).

HRM uses low-cost dyes and requires less optimization than similar systems based on TaqMan® and fluorescence resonance energy transfer (FRET) probes. Compared to these methods HRM is a simpler and more cost-effective way to characterize multiple samples (KAPA Biosystems).

2.3.3 KASP Analysis

The KBioscience Competitive Allele Specific PCR genotyping system (KASP) is a homogeneous, fluorescent, endpoint genotyping technology. The technology was initially developed for use at the KBio Science in house genotyping facility but gradually evolved into a global benchmark technology. KASP offers the simplest, most cost-effective and flexible way to determine both SNP and insertion / deletion genotypes. Analysis can be carried out in 48, 96, 384 and 1536 well plate formats (KBio Science).

The KASP genotyping system is comprised of two components (1) The SNP specific assay (a combination of three unlabelled primers) (2) The universal Reaction Mix, which contains all other required components including the universal fluorescent reporting system and a specially-developed Taq polymerase (KBio Science).

The KASP genotyping system has been used successfully in a wide variety of organisms, achieving around 90% SNP to assay conversion rate.

The mechanism of action behind KASP is novel; to elucidate, first it is necessary to consider the constituent oligonucleotides:

- Two allele specific primers (one for each SNP allele). Each primer contains a unique unlabelled tail sequence at the 5' end.
- One common (reverse) primer.

- Two 5' fluor labelled oligos, one labelled with FAM, one with CAL Fluor Orange 560. These oligo sequences are designed to interact with the sequences of the tails of the allele specific primers.
- Two oligos, with quenchers bound at the 3' ends. These oligo sequences are complementary to those of the fluorlabelled oligos (and therefore also complementary to the tails of the allele specific primers). These quenched oligos therefore bind their fluor labelled complements and all fluorescent signal is quenched until required (KBio Science).

In the initial stage of PCR, the appropriate allele specific primer binds to its complementary region directly upstream of the SNP (with the 3' end of the primer positioned at the SNP nucleotide). The common reverse primer also binds and PCR proceeds, with the allele specific primer becoming incorporated into the template. During this phase, the fluor labelled oligos remain bound to their quencher bound complementary oligos, and no fluorescent signal is generated (KBio Science).

As PCR proceeds further, one of the fluor labelled oligos, corresponding to the amplified allele, also gets incorporated into the template, and is hence no longer bound to its quencher bound complement. As the fluor is no longer quenched, the appropriate fluorescent signal is generated and detected by the usual means. If the genotype at a given SNP is homozygous, only one or other of the possible fluorescent signals will be generated. If the individual is heterozygous, the result will be a mixed fluorescent signal (KBio Science).

CHAPTER 3

METHODOLOGY PROCESS

3.1 MATERIALS

3.1.1 Samples

In this study, DNA samples were obtained by using whole blood and buccal cells of volunteers, after sign on of informed consent form as shown in Figure 2.1. For DNA isolation, the subjects used in the work were 20 (11 female and 9 male) by using whole blood.

3.1.2 Chemicals

3.1.2.1 HRM and KASP Chemicals

Bio-Rad SsoFast EvaGreen supermix is a 2x concentrated, ready-to-use reaction cocktail containing all components, except primers and template, HRM. This supermix uses technology to deliver performance in a wide range of PCR applications, by combining a novel engineered hot-start fusion polymerase with optimized buffer and EvaGreen dye.

The KASP SNP Genotyping version 4.0 kit includes KASP 2X Reaction Mix (containing Taq polymerase enzyme and the passive reference dye, 5-carboxy-X-rhodamine, succinimidyl ester (ROX)), MgCl₂ (50mM; for A/T-rich DNA regions), DMSO (for G/C-rich DNA regions), SNP-specific assay (a combination of three unlabelled primers), 3 dyes (FAM channel dye, VIC channel dye, FAM-VIC channel dyes). The used reagents and volume of these reagents were referred in Table 2.7

3.1.2.2 Primers

In this thesis, previously presented primers were choised and used. The prepared primers which were referred in Table 3.3 were synthesized at Metabion International AG (Deutschland).

Table 3.1 The designed primers used in multiplex HRM analysis.

Genes	Primer Sequence	Length	T _m	GC % content	amplicon length (bp)
5HTT-T3287C-F	5'-CACCCCGCCCCGTAG-3'	15	57°C	80,0	59
5HTT-T3287C-R	5'-GCGCTCCCGCTGGAT-3'	15	54°C	73,3	
5HTT-VNTR-F	5'-GTCAGTATCACAGGCTGCGAG-3'	21	63°C	57,1	299
5HTT-VNTR-R	5'-TGTTCCCTAGTCTTACGCCAGTG-3'	22	62°C	50,0	

3.1.2.3 Softwares

The below softwares were used in this study.

- Real-time PCR: Rotor-Gene Q 1.7.94
- DNA sequence viewer: Finch TV
- Bioinformatic studies: GeneTool Lite

3.1.3 Instruments

The used laboratory equipments in this work were given as a list in Table 3.2.

Table 3.2 List of instruments.

Equipments	Brand Name
Real-time PCR Cycler	Corbet Rotor-Gene 6000
Refrigerator	Philips, (+4°C, -20°C)
Freezer	BEKO, (-20°C)
Freezer	New Brunswick Scientific Ultra Low Temperature Freezer U570-premium
Water Purification System	Millipore
Microbalance	SCALTEC SPB54
Micropipettes	Thermo Scientific, GLP Digital Kits 1 & 2
Filter tips sterile	Neptune BT brand barrier tips
Filter tips sterile	Greiner Bio-one FT10 RNase/DNase free guaranteed

3.2 METHODS

3.2.1 HRM Analysis

HRM analysis were done by using BioRadSsOFAST kit. BioRadSsoFast reaction components and thermal cycling conditions were given in Table 3.3, Table 3.4 and Table 3.5.

Table 3.3 The reagents and amounts for HRM kit used in study.

Components	BioRadSsoFast (2x)
	Vol (Final Concentration)
Master Mix	10 μ l (1x)
Forward Primer (12.5μM)	0.8 μ l (500 nM)
Reverse Primer (12.5μM)	0.8 μ l (500 nM)
Template DNA	1-5 μ l (1-50 ng/ μ l)
ddH₂O	up to 20 μ l
Total Volume	20 μ l

Table 3.4 HRM conditions for rs25533.

Cycling Steps	rs25533
Enzyme Activation	98°C – 2 min
Denaturation	98°C – 5 sec
Annealing	59,1°C – 30 sec
Extension	
Cycle #	50
HRM	75–95°C
	0.1 °C / 2 sec

Table 3.5 HRM conditions for 5HTT-VNTR.

Cycling Steps	5HTT-VNTR
Enzyme Activation	98°C – 2 min
Denaturation	98°C – 5 sec
Annealing	62,1°C – 30 sec
Extension	
Cycle #	50
HRM	75–95°C
	0.1 °C / 2 sec

3.2.2 KASP Analysis

The KASP SNP Genotyping version 4.0 kit includes KASP 2X Reaction Mix (containing Taq polymerase enzyme and the passive reference dye, 5-carboxy-X-rhodamine, succinimidyl ester (ROX)), MgCl₂ (50mM; for A/T-rich DNA regions), DMSO (for G/C-rich DNA regions), SNP-specific assay (a combination of three unlabelled primers), 3 dyes (FAM channel dye, VIC channel dye, FAM-VIC channel dyes).

Although major optimization of KASP version 4.0 genotyping system were done in KBioscience laboratories, some further optimizations were required. MgCl₂ and high ROX containing dye additions were examples for reaction component optimizations of KASP, when increasing annealing temperatures and additional cycles were examples for thermal cycling conditions optimizations. The reaction components and thermal cycling conditions were outlined in Table 3.6 and Table 3.7.

Table 3.6 The reagents for KASP version 4.0 genotyping.

Component	KASP Genotyping mix assembly
DNA	1
KASP 2X Reaction Mix	10
Assay	0,275
H₂O	8,725
Total Reaction Volume	20 µl

Table 3.7 Thermal cycling conditions for KASP version 4.0 genotyping.

Cycling Steps	
Hot-start Enzyme Activation	94°C-15 min
Cycling 1	94°C-20 sec 65-57°C for 60 seconds (dropping 0,8°C per cycle)
Cycle #	10
Cycling 2	94°C-20 sec 57°C-60 sec
Cycle #	26
Reading	under 40 °C

3.2.3 Sequencing for Confirmation

For further confirmation of genotyping by KASP system, automated sequencing for selected PCR products were done in Refgen Biotechnology Company (METU, Ankara) (<http://www.refgen.com/index1.asp>). The results were obtained online as ABI extension documents and analyzed by the FinchTV 1.4.0 version, chromatograph analyzing program.

CHAPTER 4

RESULTS

Real-time PCR based approaches were used in this thesis. These approaches are:

- HRM genotyping technique based on sequence matching and auto-calling genotypes with reference to controls.
- KASP version 4.0 analysis of rs25533 which is based on FRET technology and
- The results are confirmed where necessary by sequencing.

Twenty samples were used in this study.

4.1 HRM RESULTS

For HRM analysis, two types of polymorphism were studied and so there were two primers. One of them was SNP polymorphism and the other one was VNTR polymorphism. SNP was rs25533 and VNTR was 5HTT-VNTR.

4.1.1 rs25533 Results

The worked amplicon of rs25533 which in this study included 59 bases. The results showed one type genotype according to rs25533 SNP.

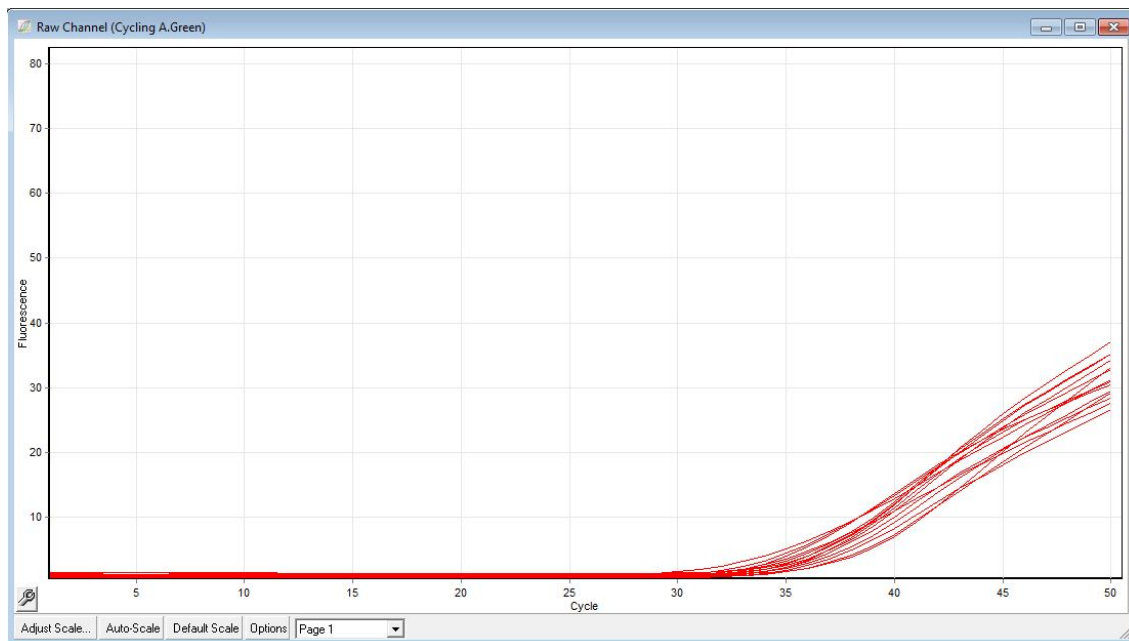


Figure 4.1 Amplification curves of all 20 samples for rs25533 analysis on HRM show that HRM has been done well.

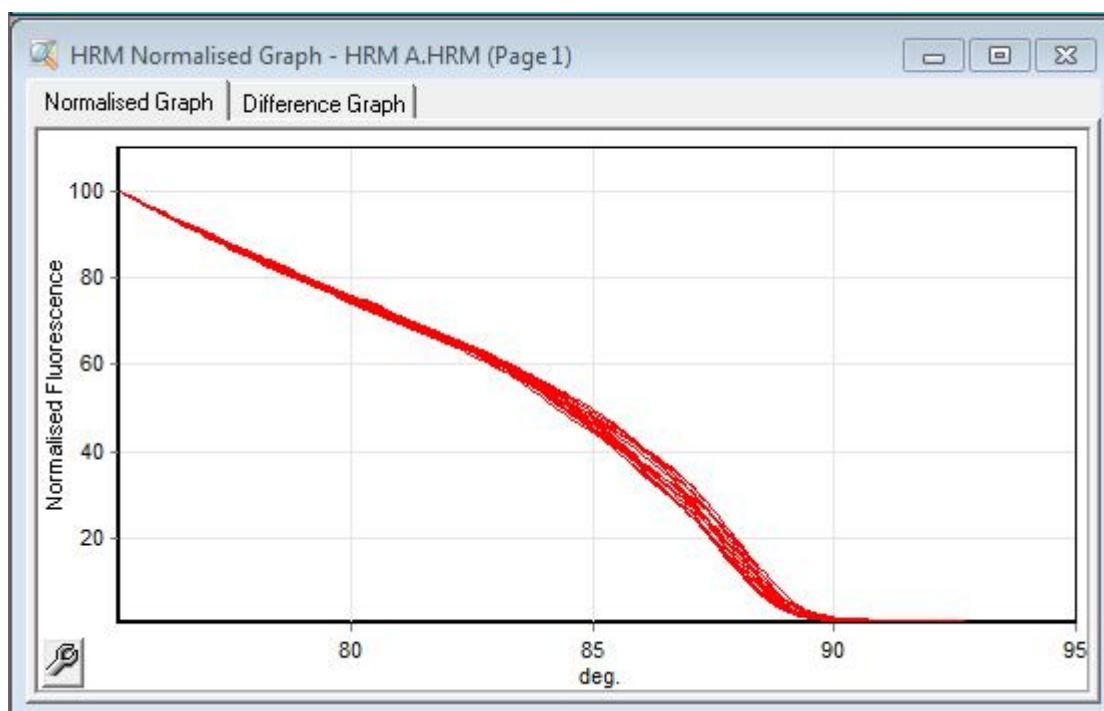


Figure 4.2 Normalisation curves of all 20 samples for analysis rs25533 on HRM show similarities of samples.

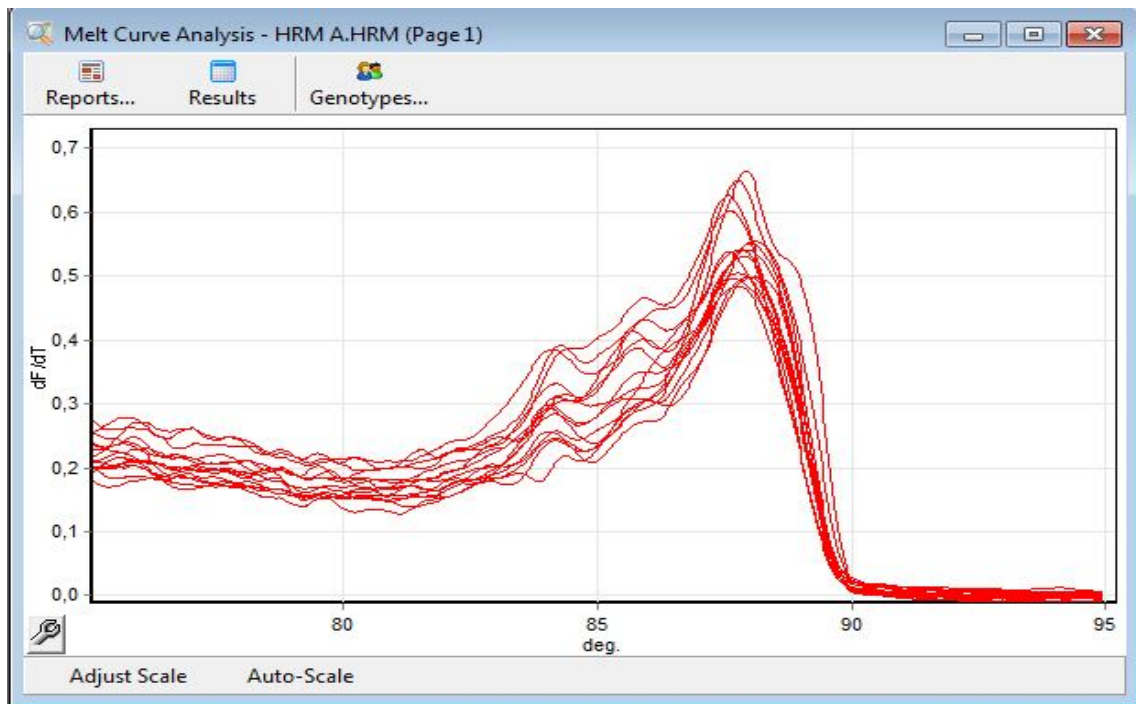


Figure 4.3 Melting curves of 20 samples for analysis rs25533 on HRM.

The figure shows a software window titled "Melt Curve Results - HRM...". It contains a table with the following data:

No.	Name	Genotype	Peak 1
1	1		88,02
2	2		87,92
3	3		88,08
4	4		88,08
5	6		87,95
6	7		87,78
7	8		87,83
8	9		87,60
9	10		87,95
10	11		85,92
11	12		87,72
12	13		87,87
13	14		87,77
14	15		87,65
15	16		87,95
16	17		88,25
17	18		87,92
18	19		87,75
19	20		88,00
20	21		85,92

Figure 4.4 T_m of all 20 samples for analysis rs25533 samples on HRM show proximity of melting temperatures.

No.	C	Name	Genotype	Confidence %
1	1		CC	98,98
2	2		CC	96,84
3	3		CC	100,00
4	4		CC	99,02
5	6		CC	96,87
6	7		CC	99,55
7	8		CC	94,83
8	9		CC	97,09
9	10		CC	95,39
10	11		CC	97,95
11	12		CC	94,48
12	13		CC	97,15
13	14		CC	86,24
14	15		CC	97,58
15	16		CC	98,82
16	17		CC	85,47
17	18		CC	95,26
18	19		CC	96,32
19	20		CC	96,68
20	21		CC	98,84

Figure 4.5 HRM genotyping results of all 20 samples for analysis rs25533 on HRM show affinity ratios.

4.1.2 5HTT-VNTR Results

The worked amplicon of 5HTT-VNTR which in this study included 299 bases. The results showed two different genotypes which were 7/7 repeats and 9/9 repeats.

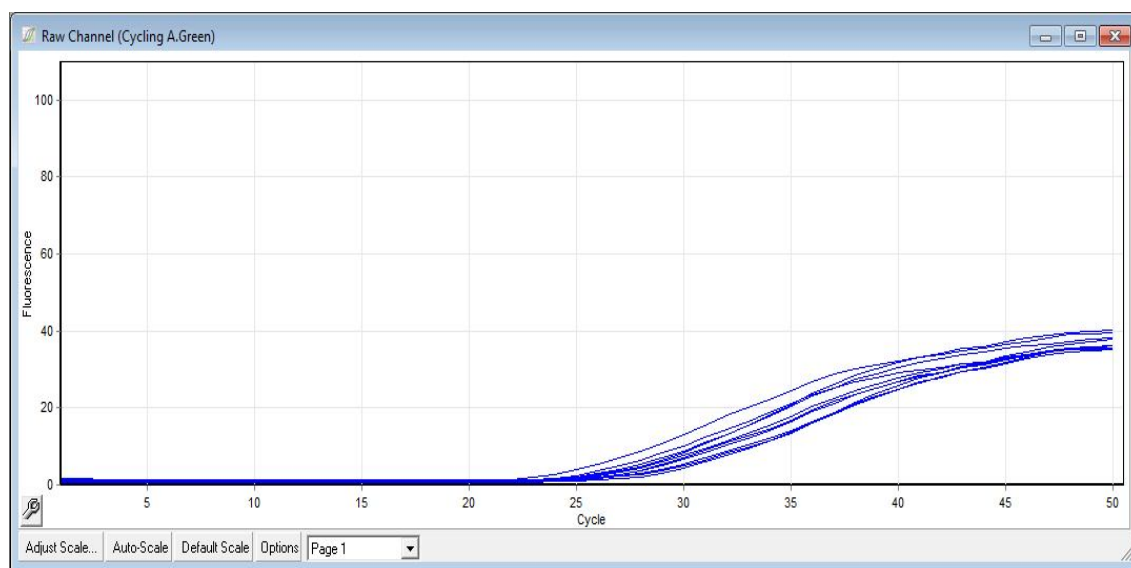


Figure 4.6 Amplification curves of same 12 samples for 5HTT-VNTR analysis on HRM show that HRM has been done well.

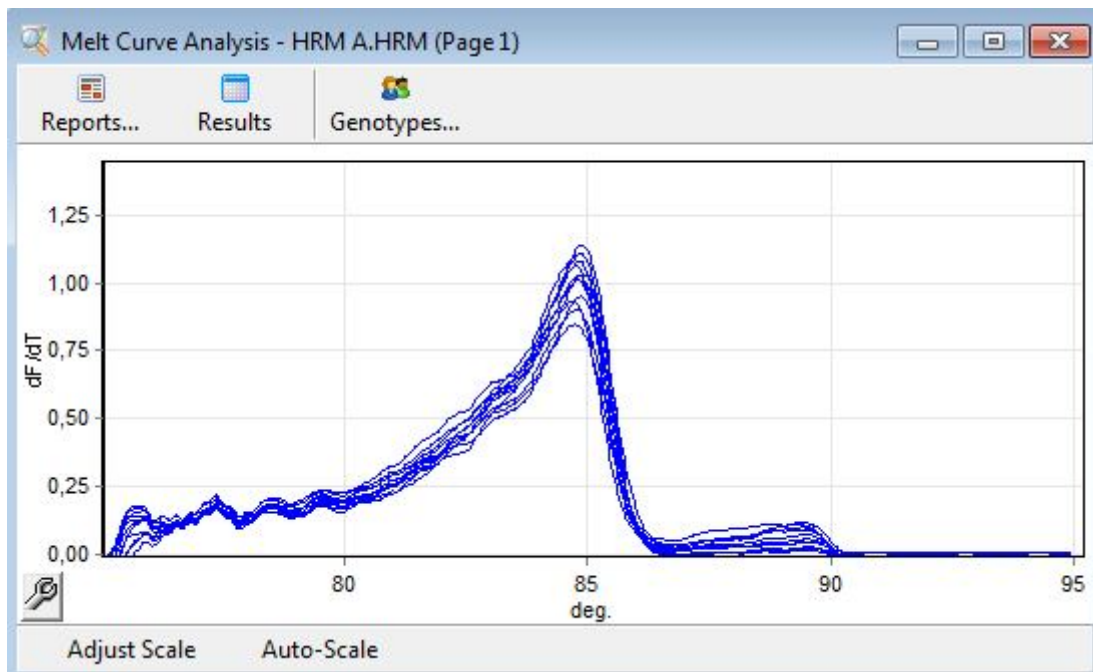


Figure 4.7 Melting curves of same 12 samples for analysis 5HTT-VNTR on HRM.

The figure shows a software window titled "Melt Curve Results - HRM...". It contains a table with the following data:

No.	Name	Genotype	Peak 1
1	1		84,85
2	2		84,88
3	3		84,85
4	4		84,82
5	9		84,77
6	10		84,73
7	12		84,80
8	13		84,90
9	18		84,68
11	19		84,68
12	20		84,75
13	21		84,83

Figure 4.8 T_m of same 12 samples for analysis rs25533 samples on HRM show proximity of melting temperatures.

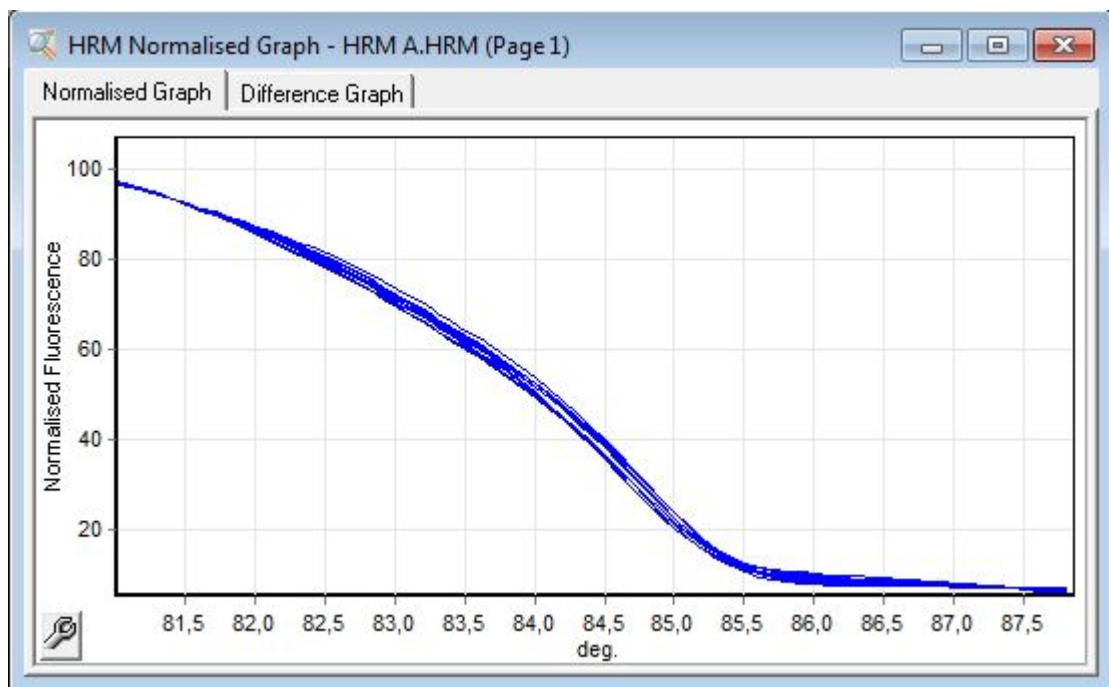


Figure 4.9 Normalisation curves of same 12 samples for analysis rs25533 on HRM show similarities of samples.

No.	C	Name	Genotype	Confidence %
1	1		7/7	96,29
2	2		7/7	96,48
3	3		7/7	98,60
4	4		7/7	99,41
5	9		7/7	99,57
6	10		7/7	98,04
7	12		7/7	100,00
8	13		7/7	96,53
9	18		7/7	95,40
11	19		7/7	98,07
12	20		7/7	98,68
13	21		7/7	99,61

Figure 4.10 HRM genotyping results of same 12 samples for analysis rs25533 on HRM show affinity ratios.

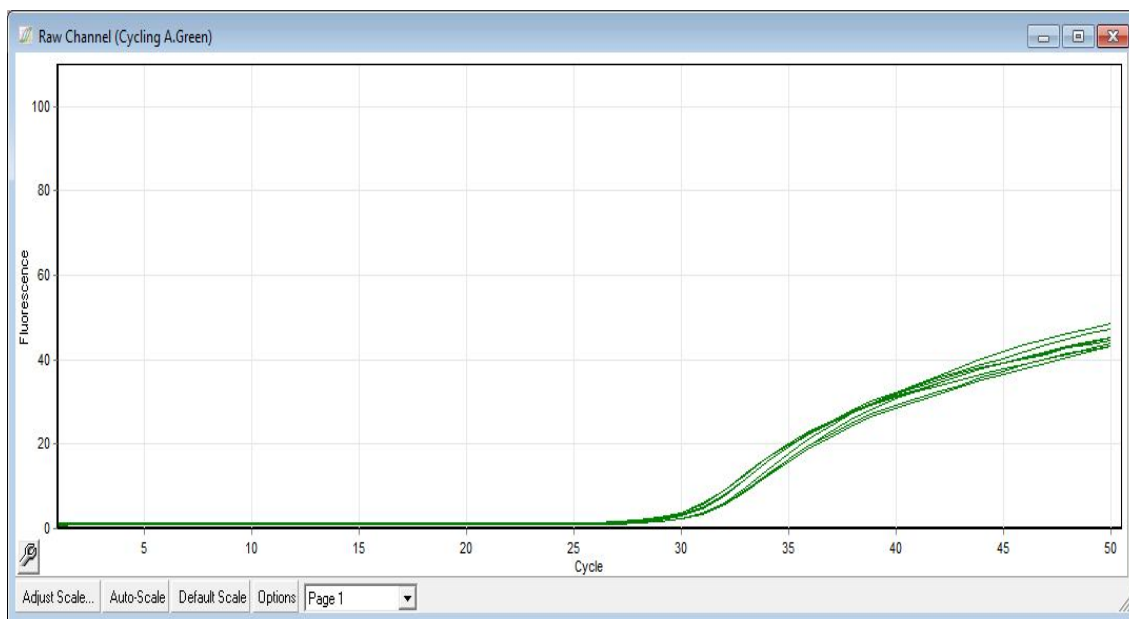


Figure 4.11 Amplification curves of 8 samples for rs25533 analysis on HRM show that HRM has been done well.

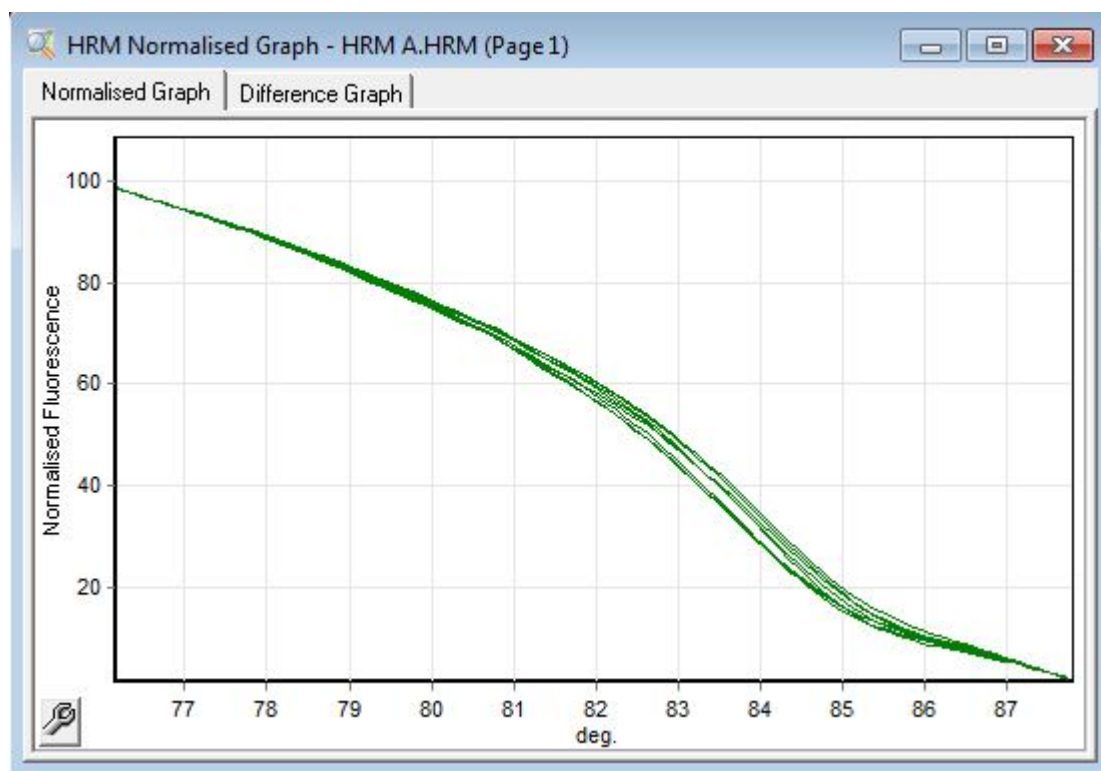


Figure 4.12 Normalisation curves of same 12 samples for analysis rs25533 on HRM show similarities of samples.

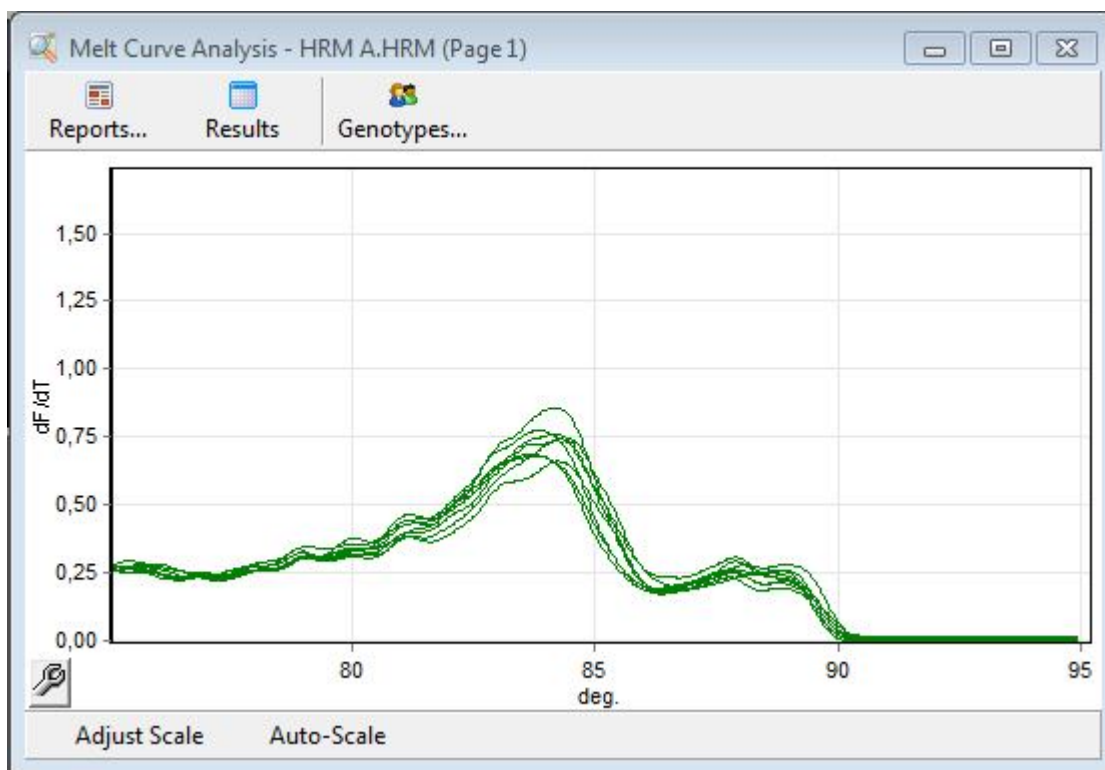
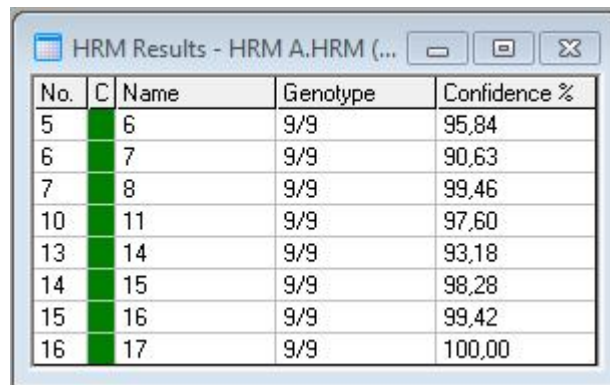


Figure 4.13 Melting curves of same 8 samples for analysis 5HTT-VNTR on HRM.

The figure shows a window titled "Melt Curve Results - HRM...". It contains a table with the following data:

No.	Name	Genotype	Peak 1
5	6		84,17
6	7		84,13
7	8		83,68
10	11		83,70
13	14		83,80
14	15		83,63
15	16		84,23
16	17		84,25

Figure 4.14 T_m of same 8 samples for analysis rs25533 samples on HRM show proximity of melting temperatures.



No.	C	Name	Genotype	Confidence %
5	6		9/9	95,84
6	7		9/9	90,63
7	8		9/9	99,46
10	11		9/9	97,60
13	14		9/9	93,18
14	15		9/9	98,28
15	16		9/9	99,42
16	17		9/9	100,00

Figure 4.15 HRM genotyping results of same 8 samples for analysis rs25533 on HRM show affinity ratios.

4.2 VALIDATION RESULTS

The KASP genotyping system which is based on FRET technology is a homogeneous, fluorescent, endpoint-genotyping assay. For SNP genotyping it is one of the simplest and cheapest way. KASP system is composed of two components as SNP-specific assay (includes three unlabelled primers) and universal reaction mix (includes universal fluorescent reporting system, specially-developed Taq polymerase and all required components).

For KASP analysis, SNP type of polymorphism were studied and so there was one designed primers. The SNP was rs25533.

The procedure is as follows: after HRM and melt analysis, amplification products, which was 59 bp in size of rs25533 samples and 5HTT-VNTR samples from all samples were sent to Refgen Biotechnology Company (METU, Ankara) (<http://www.refgen.com/index1.asp>) for sequencing. The results were as ABI extension documents which can be analyzed by Finch TV 1.4.0 versions, chromatograph analyzing program.

No.	Name	Genotype	Cycling A. Greer
1	1	Mutant	Reaction
2	2	Mutant	Reaction
3	3	Mutant	Reaction
4	4	Mutant	Reaction
5	6	Mutant	Reaction
6	7	Mutant	Reaction
7	8	Mutant	Reaction
8	9	Mutant	Reaction
9	10	Mutant	Reaction
10	11	Mutant	Reaction
11	12	Mutant	Reaction
13	13	Mutant	Reaction
14	14	Mutant	Reaction
15	15	Mutant	Reaction
16	16	Mutant	Reaction
17	17	Mutant	Reaction
18	18	Mutant	Reaction
19	19	Mutant	Reaction
20	20	Mutant	Reaction
21	21	Mutant	Reaction

Figure 4.16 Genotyping results of KASP analysis of rs25533 show all 20 samples have same genotypes.

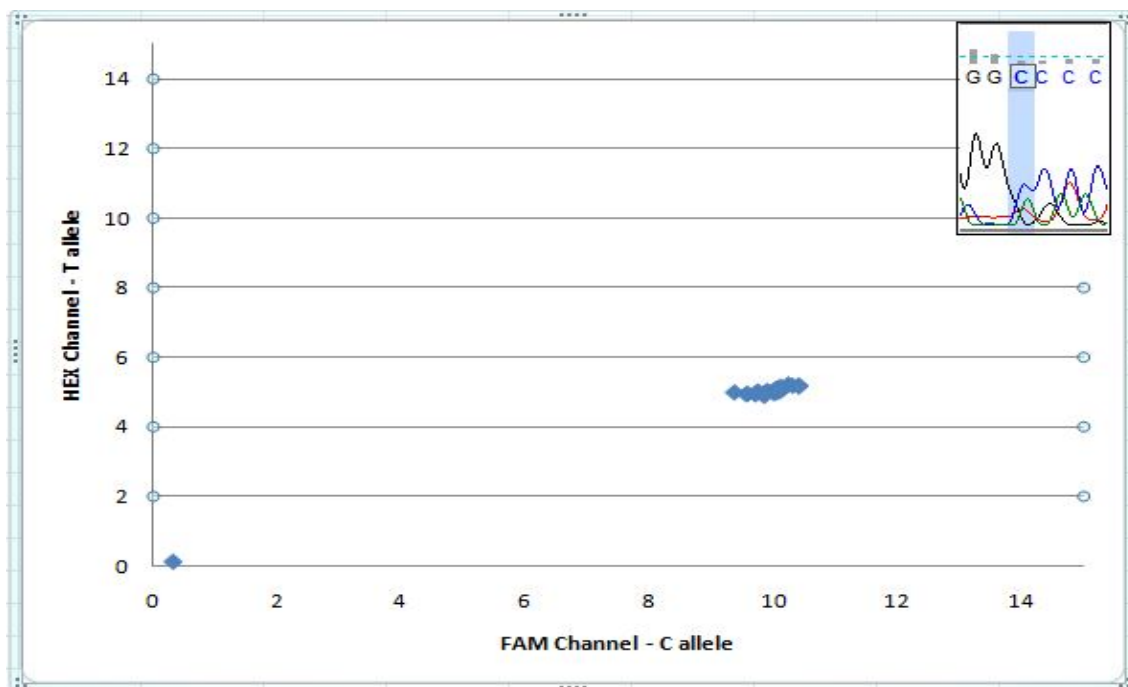


Figure 4.17 KASPar version 4.0 genotyping graph for rs25533 SNP shows all 20 samples have same genotype. Sequencing results of all 20 samples show all samples are same genotype.

cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	1
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	2
cgcgagggcg	acctatcca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	3
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	4
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	6
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	7
cgcgagggcg	acctatctca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	8
cgcgagggcg	acctacctca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	9
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	10
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	11
cgcgagggcg	acctatctca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	12
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	13
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	14
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	15
cgcgagggcg	acctatcca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	16
cgcgagggcg	acctatctca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	17
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	18
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	19
cgcgagggcg	acctacctca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	20
cgcgagggcg	acctaccca	acgcgagcgg	ccctccccg	gcgcgatgcc	ccgccccac	21

Figure 4.18 Sequencing results of all 20 samples which show all samples are CC genotype.

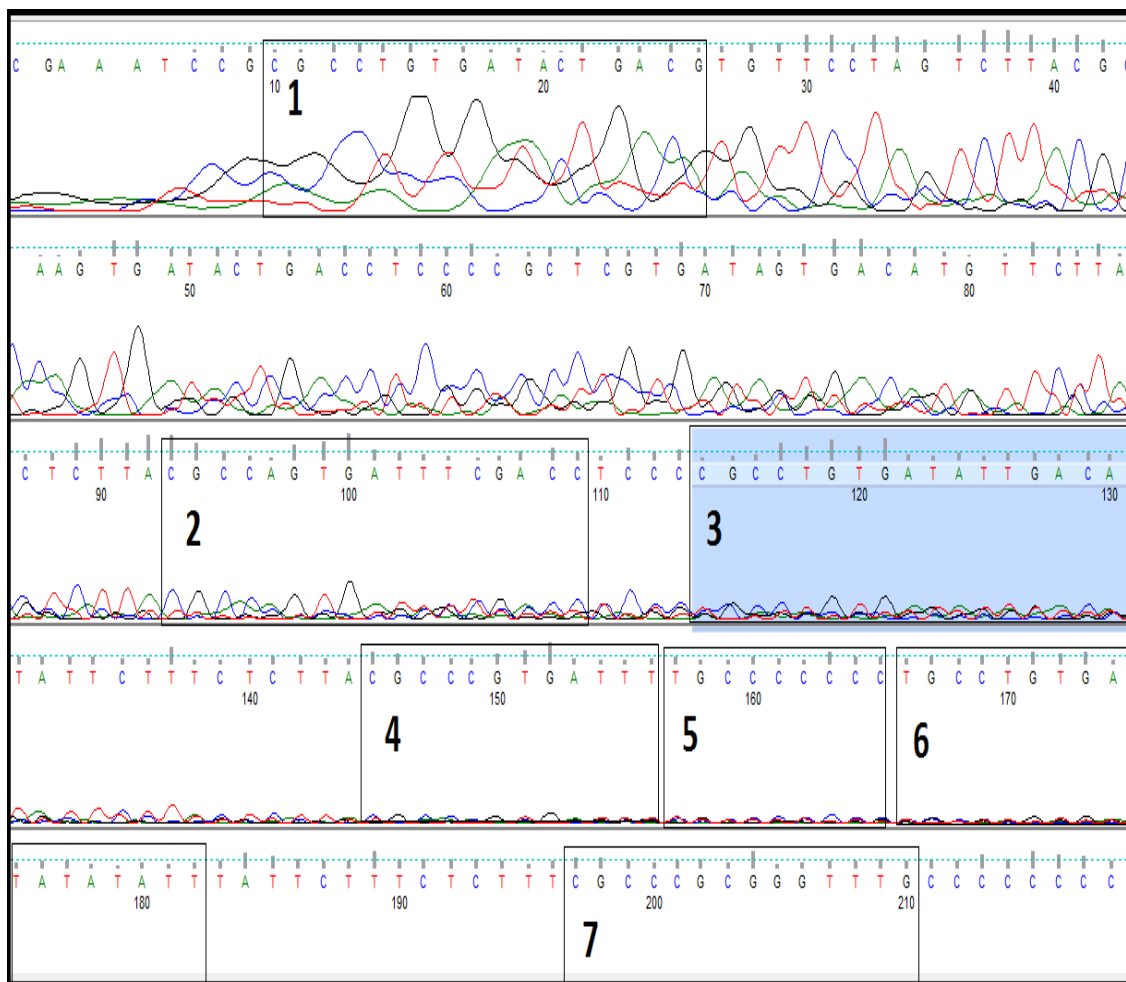


Figure 4.19 Sequencing results of same 12 samples of 5HTT-VNTR of chromatograph analyzing program (Finch TV 1.4.0). The results show that these 12 samples include 7 repeats and they are 7/7 genotype according to 5HTT-VNTR gene.

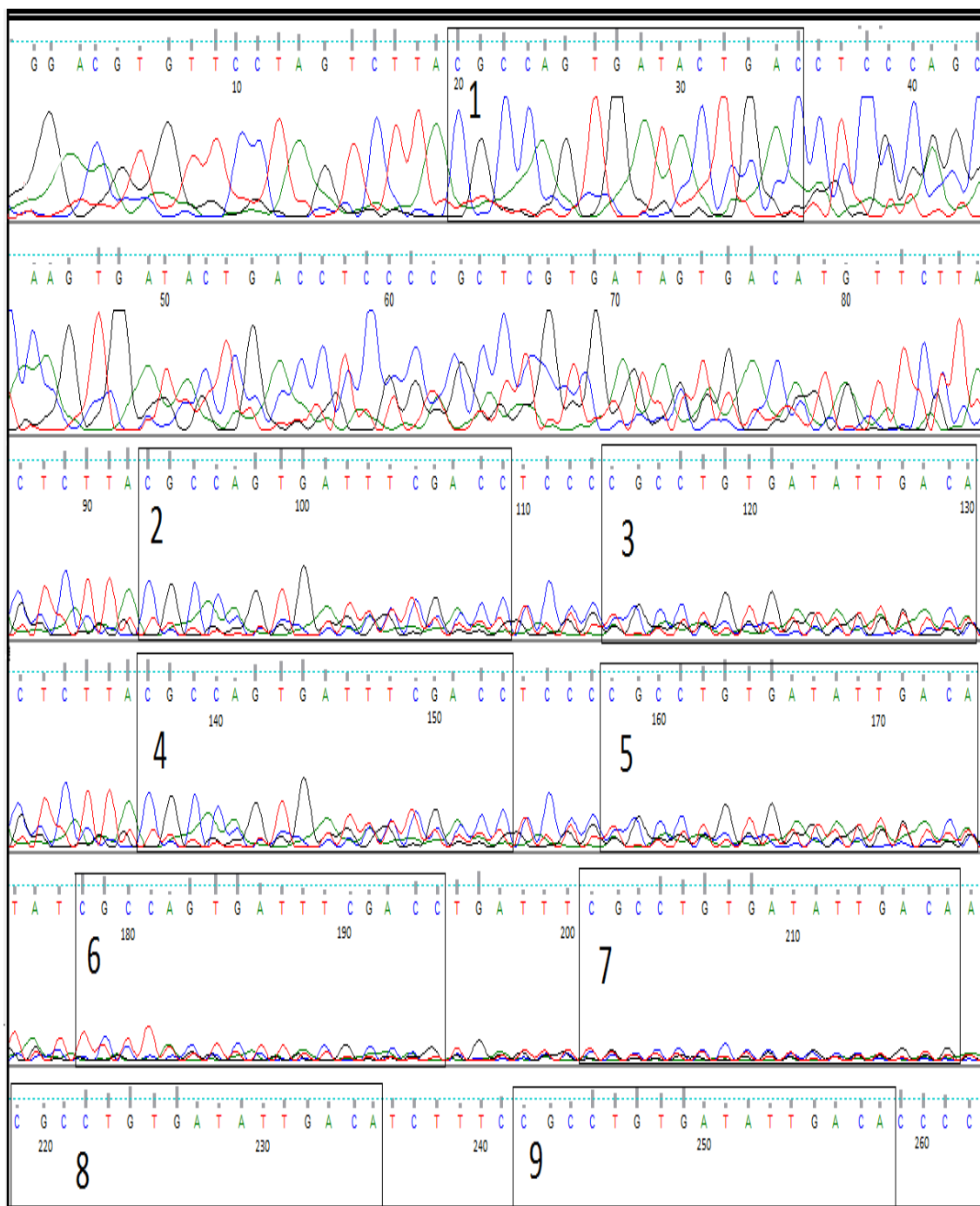


Figure 4.20 Sequencing results of same 12 samples of 5HTT-VNTR of chromatograph analyzing program (Finch TV 1.4.0). The results show that these 8 samples include 7 repeats and they are 9/9 genotype according to 5HTT-VNTR gene.

As a evaluation for 20 subjects, all the SNP and VNTR genotyping results for 5HTT-T3287C and 5HTT-VNTR gene SNPs and VNTR have been given in Table 4.1 and Table 4.2.

Table 4.1 SNP genotyping results for 20 samples.

Genes	wild type (homozygote)	mutant type (homozygote)	heterozygote
5HTT-T3287C (rs25533)	0 (C/C)	20 (T/T)	0 (C/T)

Table 4.2 VNTR genotyping results for 20 samples.

Genes	9/9	7/7	12/12 – 12/10 10/10 – 12/9 10/9 – 12/7 10/7 – 9/7
5HTT-VNTR	8	12	0

The results gave that our 20 samples are CC genotype according to rs25533 type SNP. Besides, 12 of 20 samples include 7/7 repeat genotypes and 8 of 20 samples include 9/9 repeat genotypes according to 5HTT-VNTR gene.

CHAPTER 5

DISCUSSION

As aforementioned in the introduction part of this thesis, the most of the variations are single nucleotide changes and variable number of tandem repeats in human genome. The contribution of SNPs to the complex diseases and drug responses is a very important and complex issue which cannot be disclaimed. Currently, SNPs are considered to be the most prevalent genetic marker replacing the others like microsatellites in genome wide association studies and linkage disequilibrium studies (Karabulut, 2012).

Prior studies have noted the importance of SNP detection technologies which are grouped in two categories; scanning DNA sequences for previously unknown polymorphisms and screening individuals for known polymorphisms (Kwok and Chen, 2003). With the completion of the Human Genome Project, SNP detection technologies had an evolution, thus new technologies are emerging day by day. But, the requirements change from large-scale projects to normal laboratory studies which requires laborsaving, cost-effectiveness (Karabulut, 2012). According to these requirement HRM analysis is a convenient method.

Known polymorphism detection methods, such as DNA sequencing has been considered the gold standard technique because it can identify the specific mutation that may be present. However it has the disadvantages of high relative cost, laborious, time-consuming and limited sensitivity. Also, pyro-sequencing has advantage of greater sensitivity and lower cost, but the relatively high cost of the required apparatus has limited its uptake (Krypuy et al., 2006). Many other methods for mutation scanning have been developed to screen for differences between two copies of DNA within an

individual. These techniques include single-strand conformational polymorphism analysis (SSCP), denaturing high performance liquid chromatography (DHPLC), temperature gradient capillary electrophoresis (TGCE) and mass spectroscopy (Bocker, 2007). All of these methods require separation of the sample on a gel or other matrix. Fluorescently labeled, probe-based methods, such as dual hybridization, exonuclease (TaqMan), or hairpin (Molecular Beacon) probes, may be used for polymorphism detection, but only for the bases covered by the probe. Hence, these methods are not amenable to mutational scanning as mutational scanning requires methods that can detect mutations over larger regions. Furthermore, some of the above methods are not automated and are therefore labor intensive while others are complex, costly and require specialized instrumentation (Li et al., 2011).

HRM analysis is a simple PCR-based method. In the presence of saturating concentration of DNA binding dyes, the specific sequence of the amplicon determines the melting behavior as the temperature of the solution increased. Fluorescence intensity decreases as the double stranded DNA becomes single stranded and the dye is released. The distinctive melting curve can be used to detect DNA sequence variation in the amplicon without the need for any post-PCR processing (Li et al., 2011). The method is easy to use, highly sensitive, specific, low cost and yields rapid sample turn-around, making HRM analysis an attractive choice for the detection of SNP or VNTR type polymorphism.

An other important advantage of HRM over many of the above methods is that it is an in-tube method in which the analysis is performed immediately after the amplification and is thus particularly suitable for medium to high-throughput applications. Methods that require amplified PCR product to be removed from the tube for analysis are inevitably more laborious and require stringent precautions to prevent crossover of PCR products. The HRM analysis has a major potential for the rapid and inexpensive detection of DNA sequence variations such as polymorphisms in DNA (Krypuy et al., 2006).

HRM method is similar to melting analysis by principle but it differs from this method with used chemistry, software and instrumentation. As mentioned in the literature, it is a closed-tube, fast and cost effective method. With these properties, HRM is separated from other SNP genotyping technologies. According to the samples GC content, sequence characteristics and heterozygosity (Pereyra et al 2012), difference in melt curves emerges.

As a discussion for cost effectiveness, prices of HRM analysis and KASP analysis have been given in Table 5.1.

Table 5.1 Comparisons for the different SNP detection methods' prices.

	Price for Master mix	Price for Primers
Bio-Rad SsoFastEvaGreenSupermix	686,03 TL for 200 samples	~14.4 \$
KASP V.4.0	210 \$ for 200 samples	189 \$ (design and validation included)
Sequencing	1 SNP → 65 TL/sample	

The present study was designed to evaluate the HRM analysis to detect clinically important SNP and polymorphisms on serotonin transporter gene and to try show easiness of HRM analysis and to make easier HRM analysis possibly. For these purposes, ready primers were chosen to not design any new primers. Also, using ready primers is an advantage for cost and time. On the other hand, ready master mix which includes EvaGreen was chosen as HRM mix, it is also easier decision to work HRM analysis. But for all that, primers and master mix may need optimizations. Given conditions cannot be enough. First step of study, annealing temperature, HRM conditions and reagent ratios should be tried and after these steps ideal conditions can be found.

It is very important not to trust the curve similarities that much, the success of application is case dependent that sometimes samples exhibiting similar curves do not necessarily exhibit the same genetic variants. This phenomenon is inherent in any technique based on melting of amplicons. To overcome challenges intrinsic in the nature of this work, so it is very important to conduct experiments to optimize all factors in the HRM workflow — from the use of positive controls and amplicon design (the best results are obtained with amplicons <400 bp), to standardization of buffers and reagents. Identifying the ideal number of samples per amplicon in each experiment is also important. Software normalization calculations for each sample affect the shape of the melt curve generated; experiments with

fewer samples per amplicon make it difficult for HRM software to distinguish wild-types (Karabulut, 2012).

HRM analysis has been successfully applied for the two clinically important polymorphisms of SERT genes which were 5HTT-T3287C (rs25533) and 5HTT-VNTR. The results gave that our 20 samples are CC genotype and wild type according to rs25533 type SNP of population. Besides, 12 of 20 samples include 7/7 repeat genotypes and 8 of 20 samples include 9/9 repeat genotypes according to 5HTT-VNTR gene which shows VNTR type polymorphism. Analysing HRM results and observing SNP and polymorphism changes were easy because of sensitivity of HRM analysis method. But not having negative control was not good condition to analyse and compare results for our study. Evaluate 5HTT-VNTR gene region by HRM analysis gave successful results but analysing sequencing results was not easier as HRM analysis. 5HTT-VNTR gene is complex and random region according to repeats and sequences, nevertheless working HRM analysis with this gene was not difficult. 5HTT-VNTR part of this work gives a view about difference with HRM analysis and another method.

Traditional sequencing experiments should be conducted in parallel to compare results and verify accuracy of this new technique when used for new SNPs. Initially, all HRM analysis results should be confirmed by sequencing or KASP like validation methods; after results demonstrated high enough sensitivity and specificity of HRM analysis, with few false positives the confidence can be achieved for the SNP genotyping. After developing the confidence in HRM analysis, the diagnostic aspects of the SNP works which require a high degree of throughput should be evaluated as to process sample assays in single replicates and sequencing only the aberrant melting curves to confirm the presence of a genetic variant. For HRM analysis, everything that is assigned to an aberrant melting curve cluster (by the software) has to be sequenced (Karabulut, 2012).

Real-time PCR based technologies are among the most preferred choices in SNP genotyping currently, like probe-based genotyping, KASP genotyping, High Resolution Melting Analysis methods (Karabulut- 2012).

KASP v.4 SNP genotyping technology which was a FRET based, homogenous, end-point SNP detection method was patented in Mar 2011. We can say that it is a new technology although KASP v.3 was present in before it. KASP method is an extension of

allele-specific amplification which is highly specific for the SNP of interest but do not give any further details about the neighbouring SNPs. It is not the first but can be the second after HRM among the SNP detection technologies in cost-effectiveness and ease in application in Rotor Gene 6000 (Karabulut, 2012).

In this study, KASP analysis and sequencing methods were chosen to validate HRM analysis results, and validation results supported our HRM results successfully. Also in this step, KASP analysis method which is a new one was tried and KASP analysing results were parallel with HRM analysing and sequencing result. On the other hand, without negative control HRM analysis is an enough method to check SNP type or VNTR type change, it is possible but not enough to determinate fact SNP or polymorphism type.

REFERENCES

- Bocker, S., "Simulating multiplexed SNP discovery rates using base-specific cleavage and mass spectrometry", *Bioinformatics*, Vol. 23, pp.5-12, 2007.
- Botstein, D., White, R.L., Skolnick, M., and Davis, R.W., "Construction of a genetic linkage map in man using restriction fragment length polymorphisms", *Am. J. Hum. Genet.*, Vol. 32, pp. 314-331, 1980.
- Brown, K., "The human genome business today", *Scientific American*, pp. 50-55, July 2000.
- Calderon, B., "Tandem repeats in promoter regions of *S. Cerevisiae* generate variability in gene expression with phenotypic consequences", *THURJ*, 2009.
- Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A., Poulton, R., "Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene", *Science*, Vol. 301, pp. 386-389, 2003.
- Collins, F.S., Guyer, M.S., and Chakravarti, A., "Variations on a theme: cataloging human DNA sequence variation", *Science*, Vol. 278, pp. 1580-1581, 1997.
- Cooper, D.N., Smith, B.A., Cooke, H.J., Niemann, S., Schmidtke, J. 1985. "An estimate of unique DNA sequence heterozygosity in the human genome", *Hum. Genet.*, Vol. 69, pp. 201-205, 1985.
- Daştan, K., Mitokondriyal DNA Genomu Tek Nükleotid Polimorfizminin RealTime Pcr Cihazı İle İdentifikasyonu Ve Standardizasyonu, M.S. Thesis, İstanbul University, 2006.
- Do, H., Krypuy, M., Mitchell, P.L., Fox, S.B., Dobrovic, A., "High Resolution melting analysis for rapid and sensitive EGFR and KRAS mutation detection in formalin fixed paraffin embedded biopsies", *BMC Cancer*, Vol. 8(142), pp. 142, 2008.
- Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T.P., Bowden, D.W., Smith, D.R., Lander, E.S., et al., "A genetic linkage map of the human genome", *Cell*, Vol. 51, pp. 319-337, 1987.
- Dawson, E.G., Abecasis, R., et al. "A first-generation linkage disequilibrium map of human chromosome 22", *Nature*, Vol. 418, No. 6897, pp. 544-8, 2002.

- Erdal, M.E., Herken, H., Barlas, Ö., Erdal, N., “Polymorphisms in the serotonin transporter gene”, *Klinik Psikiyatri*, Vol. 3, pp. 192-196, 2000.
- Florez, G., Saiz, P., Garcia-Portilla, P., Alvarez, S., Nogueiras, L., Morales, B., Alvarez, V., Coto, E., Bobes, J., “Genetics and Cell Biology”, *Alcohol & Alcoholism*, Vol. 43, No. 5, pp. 516-522, 2008.
- Ford, E.B., *The New Systematics*, Clarendon Pr, Oxford, 1940.
- Haddley, K., Vasiliou, A.S., Ali, F.R., Paredes, U.M., Bubb, V.J., Quinn, J.P., “Molecular genetics of monoamine transporters: relevance to brain disorders”, *Neurochem Res*, Vol. 33, pp. 652-667, 2008.
- Hingorani, A.D., Shah, T., Kumari, M., Sofat, R., Smeeth, L., “Translating genomics into improved healthcare”, *BMJ*, Vol. 341, pp. c5945, 2010.
- Johnson, G.C. and Todd, J.A., “Strategies in complex disease mapping”, *Curr. Opin. Genet. Dev.*, Vol. 10, pp. 330-334, 2000.
- KAPA Biosystems, “Introduction to high resolution melting analysis”, Application guide.
- KBio Science, “The KASP SNP genotyping quick start guide”.
- Karabulut, D.S., Optimizations in SNP Detection Using Real-Time PCR High Resolution Melting Analysis, M.S. Thesis, Fatih University, 2012.
- Krypuy, M., Newnham, G.M., Thomas, D.M., Conron, M., Dobrovic, A., “High resolution melting analysis for the rapid and sensitive detection of mutation in clinical samples: KRS codon 12 and 13 mutations in non-small cell lung cancer”, *BMC Cancer*, Vol. 6, pp 295, 2006.
- Kwok, P.Y., and Chen, X. “Detection of Single Nucleotide Polymorphisms”, *Current Issues in Molecular Biology*, Vol. 5, pp. 43-60, 2003.
- Lander, E.S., Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., et al., “Large-scale identification, mapping, and genotyping of single nucleotide polymorphisms in the human genome”, *Science*, Vol. 280, No. 5366, pp. 1077-1082, 1998.
- Li, B.S., Wang, X.Y., Ma, F.L., Jiang, B., Song, X.X., Xu, A.G., “Is high resolution melting analysis (HRMA) accurate for detection of human disease-associated mutations? A meta analysis”, *PLoS ONE*, Vol. 6, pp. e28078, 2011.
- Litt, M. and Luty, J.A., “A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene”, *Am. J. Hum. Genet.*, Vol. 44, pp. 397-401, 1989.
- Lotrich, F.E., Pollock, B.G., “Meta-analysis of serotonin transporter polymorphisms and affective disorders”, *Cell Mol. Life Sci.*, Vol. 14, pp. 121-129, 2004.

- Marashi, S.J., S.A. Eshkoo, et al., "Detection of eight common β -globin gene mutation in thalassemia major patients using real time polymerase chain reaction (PCR)-high resolution melting and EvaGreen™ dye", *African Journal of Biotechnology*, Vol. 11(2), pp. 448-459, 2012.
- McCarthy, J.J., Hilfiker, R., "The use of single-nucleotide polymorphism maps in pharmacogenomics", *Nature Biotechnology*, Vol. 18, pp. 505-508, 2000.
- Murphy, D.L., Lerner, A., Rudnick, G., Lesch, K.P., "Serotonin Transporter: Gene, Genetic Disorders, and Pharmacogenetics", *Molecular Interventions*, Vol. 4(2), pp. 109-123, 2004
- Pereyra, S., T. Velazquez, et al., "Rapid multiplex high resolution melting method to analyze inflammatory related SNPs in preterm birth", *BMC Research Notes*, Vol. 5(69), 2012.
- Pungliya, M.S., *Single Nucleotide Polymorphism Analysis in Application to Fine Gene Mapping*, M.S. Thesis, Worcester Polytechnic Institute, 2001.
- Reuter, M., Küpper, Y., Schmitz, A., Breuer, J.P., Wend, U., Hennig, J., "Detection of single nucleotide polymorphisms by means of real time PCR", *Journal of Genetic*, Vol. 84, pp. 341-345, 2005.
- Risch, N.J., "Searching for genetic determinants in the new millennium", *Nature*, Vol. 405, pp. 847-856, 2000.
- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N., "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia", *Science*, Vol. 230, pp. 1350-1354, 1985.
- Salisbury, B.A., Pungliya, M., et al., "SNP and haplotype variation in the human genome", *Mutation Research*, Vol. 526, pp. 53-61, 2003.
- The International SNP Map Working Group. "A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms". *Nature*, Vol. 409, pp. 928-933, 2001.
- University of Arizona, Tucson, Arizona, United States, 1996, http://www.biology.arizona.edu/human_bio/problem_sets/dna_forensics_1/05t.html
- University of Colorado Boulder, Boulder, Colorado, United States, 2007, http://ibgwww.colorado.edu/genotyping_lab/pdf/5HTTLPR.pdf
- University of Michigan Comprehensive Cancer Center, Ann Arbor, United States, 2011, <http://www.med.umich.edu/cancer/prevention/summer-11-three-pcgp.shtml>

- Venter, C., et al., "The sequence of the human genome", *Science*, Vol. 291(5507), pp. 1304-1351, 2001.
- Weber, J.L. and May, P.E., "Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction", *Am. J. Hum. Genet.*, Vol. 44, pp. 388-396, 1989.
- Williams, R.B., Marchuk, D.A., Gadde, K.M., Barefoot, J.C., Grichnik, K., et al., "Serotonin-Related Gene Polymorphisms and Central Nervous System Serotonin Function", *American College of Neuropsychopharmacology*, Vol. 28(3), pp. 533-541, 2003.
- Yeo, A., Boyd, P., Lumsden, S., Saunders, T., Handley, A., Stubbins, M., Knaggs, A., Asquith, S., Taylor, I., Bahari, B., Crocker, N., Rallan, R., Varsani, D., Montgomery, D., Alpers, D.H., Dukes, G.E., Purvis, I., Hicks, G.A., "Association between a functional polymorphism in the serotonin transporter gene and diarrhea predominant irritable bowel syndrome in women", *Gut*, Vol. 53, pp., 1452-1458, 2004.