Gudep, Ravali, The Study of 5-HTTLPR and rs25531 to Identify Allele Variants of Offenders Charged with Driving While Intoxicated and Driving Under the Influence. Master of Science (Biomedical Sciences, Forensic and Investigative Genetics), July 2013, 42 pp., 12 tables, 6 figures, 19 reference titles.

Serotonin, also referred to as 5-hydroxytryptamine, is a neurotransmitter which plays a critical role in behavioral and genetics studies. The 5-hydroxytryptamine-transporter-linked polymorphic region (5-HTTLPR), present in the promoter region of the serotonin transporter gene, can have varying number of repeats. In addition, a single nucleotide polymorphism (SNP), rs25531, is present within the repeats. A novel assay was designed to genotype 5-HTTLPR and its SNP variant in a single capillary electrophoretic injection. A total of 100 samples received from the Center for Alcohol and Drug Studies at San Diego State University Research Foundation were successfully genotyped.

STUDY OF 5-HTTLPR AND rs25531 TO IDENTIFY ALLELE VARIANTS OF OFFENDERS CHARGED WITH DRIVING WHILE INTOXICATED AND DRIVING UNDER THE INFLUENCE

Ravali Gudep, B.Tech

APPROVED:

Rhonda Roby 07.22.13

Major Professor

Bobby LaRue PhD 19 July, 2013

Committee Member

Joseph E. Warren, Ph.D.

Committee Member

Committee Member

University Member

Chair, Department of Forensic and Investigative Genetics

Dean, Graduate School of Biomedical Sciences

STUDY OF 5-HTTLPR AND rs25531 TO IDENTIFY ALLELE VARIANTS OF OFFENDERS CHARGED WITH DRIVING WHILE INTOXICATED AND DRIVING UNDER THE INFLUENCE

THESIS

Presented to the Graduate Council of the

Graduate School of Biomedical Sciences

University of North Texas

Health Science Center at Fort Worth

Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Ravali Gudep, B.Tech

Fort Worth, Texas

July 2013

ACKNOWLEDGMENTS

Firstly, I would like to thank my family for their constant support and encouragement throughout the course of my Masters degree. I thank my committee members, Dr. Rhonda Roby, Dr. Joseph Warren, Dr. Bobby LaRue, Dr. Shaohua Yang and Dr. Jerry Simecka for guiding me through the project. I would like to specially thank Dr. Rhonda Roby, my major professor, without her none of this would be possible. She has been my teacher and inspiration. To Jessica Powers, thank you for helping me through the project. This project is a combined effort, without the work you did last year none of this would be possible. Thank you, Marc, Nicole and Alessandra for your words of encouragement when I was stressing out about the results. I would also like to thank Dr. Colin Hodgkinson for sending us reagents to test the assay. Finally I would like to thank Eric Ehli at the Avera Institute for Genetics.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

Serotonin is a neurotransmitter involved in carrying impulses between nerve cells. The serotonin transporter gene, located on chromosome 17, is a cell membrane protein that modulates the uptake and balance of serotonin in the brain (1). Serotonergic transmission in the brain impacts behaviors such as cognition and emotion. Length polymorphism repeats (LPR) in the promoter region of the serotonin transporter gene (SERT), also called 5-hydroxytryptamine or 5-HTT (Figure 1), are studied for their associations with different behaviors (2).

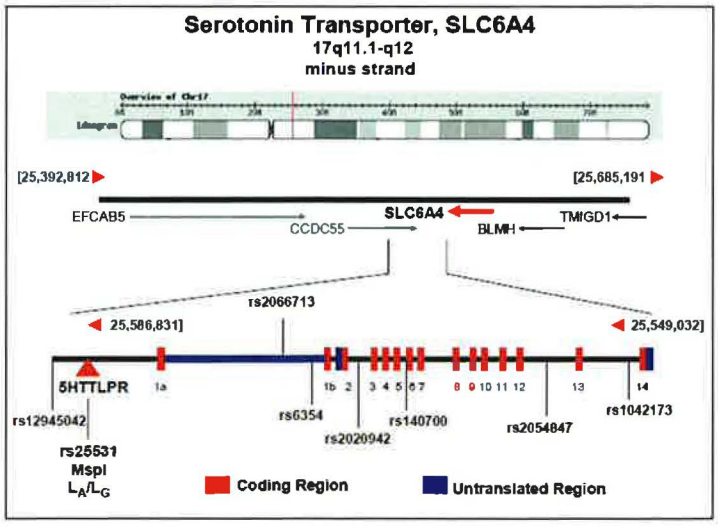


Figure 1: Serotonin Transporter Gene. Depiction of the serotonin transporter gene located within chromosome 17. The 5-HTTLPR region and identified SNPs are displayed including position rs25531 which is characterized with this assay (figure from Heils *et al.*, 3).

Repeat regions within 5-HTTLPR

5-HTTLPR has 20 to 23 base pair (bp) long GC-rich repeats in the 5' regulatory region of the gene. A repeat is said to be GC-rich if the occurrence of guanine (G) and cytosine (C) bases is higher than that of the adenine and thymine bases. More than 60% of the repeat regions within the 5-HTTLPR region have G and C bases. GC-rich regions, due to their high melting temperatures, can be hard to amplify. 7-deaza-2'deoxyguanosine triphosphate (deaza-dGTP) is a molecule that is more readily denatured than deoxyguanosine triphosphate (dGTP). Deaza-dGTP reduces the melting temperature and aids in amplification. (4). Length repeat polymorphisms create two main variants of the gene, the short (S) and the long (L) allele variants, which are 14 repeats and 16 repeats, respectively. A 43-bp insertion/deletion (indel) region exists within the gene, such that, this region is inserted in the L allele and deleted in the S allele (5). Each of these 16 repeat regions is assigned a Greek symbol (Figure 2). Studies performed by Nakamura *et al.* (6) report that, apart from the 16 repeat regions represented in Figure 2, variations exist within the sequence of these regions.

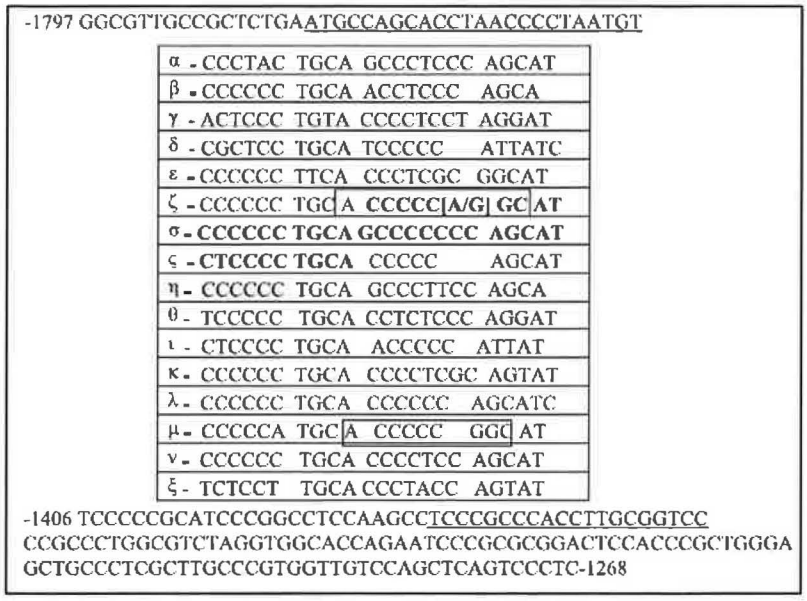


Figure 2: Repeat regions of 5-HTTLPR. Greek symbol representation of the different repeats within the 5-HTTLPR. The ζ, σ, and ς repeats have a 43bp indel, present in the L variant and absent in the S variant. The A to G SNP that leads to the L_A and L_G variants is present in the ζ repeat sequence. The indel region and the SNP position are in bold (figure from Eli *et al.*, 5).

A to G SNP rs25531

An A to G SNP, rs25531, occurs at the sixth position, 1,629bp upstream of the promoter region in the L allele. This SNP leads to the occurrence of two variants of the L allele, L_A and L_G. The G substitution on the L_G variant creates a functional activator protein-2 (AP-2) transcription factor binding site that decreases the messenger RNA expression (7). Messenger RNA (mRNA) plays a crucial role in carrying genetic information from the DNA to the ribosome for gene expression. The mRNA expression of the L_G variant is similar or the same to the S allele (Figure 3) (8).

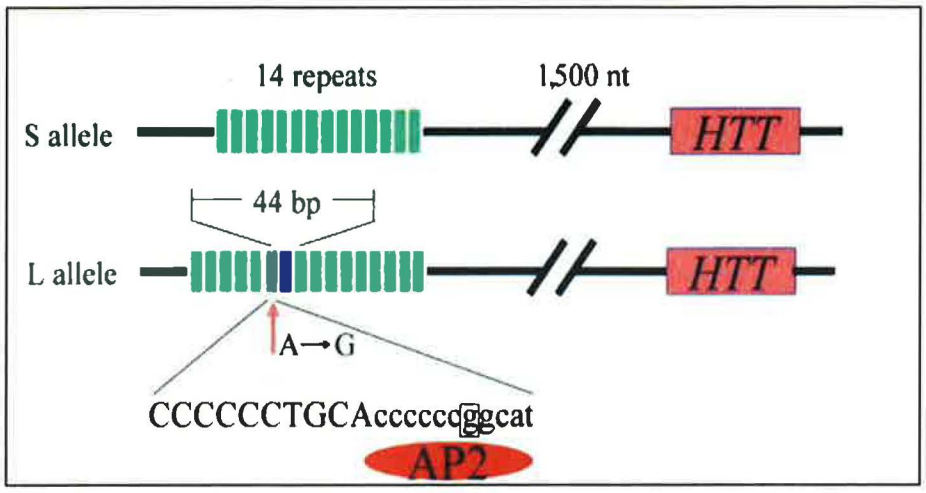


Figure 3: S and L allele variants in 5-HTTLPR. This figure focuses on the allele variants of 5-HTTLPR. 1500 nt (nucleotides) as seen in figure 1, have been removed. The L allele has a 43 to 44 bp insertion which makes it longer than the S allele. The A to G SNP within the L allele leads to the occurrence of the L_A and L_G variants. The G SNP produces an AP-2 binding sites which decreases mRNA expression (figure modified from Hu *et al.*, 7).

The three predominant variants are S, L_A and L_G. The S variant has lower transcriptional activity in comparison to the L_A variant which leads to reduced mRNA levels and serotonin reuptake (Figure 4) (7) .

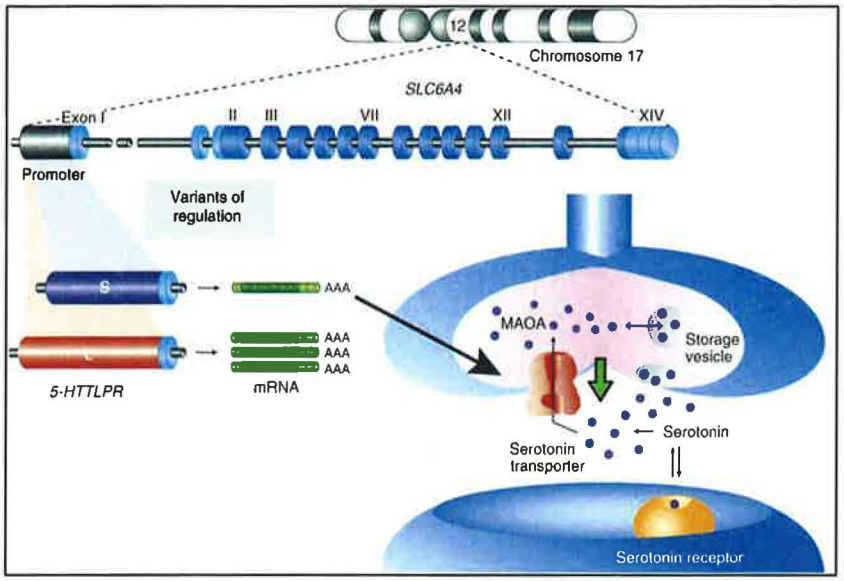


Figure 4: Gene expression of S and L alleles. Comparison of the S and L allele variants of 5-HTTLPR and how they differ in expression and regulation. The S allele (purple) has less transcriptional activity than the L allele (red). This leads to decreased mRNA expression as shown by the green bars. Decreased mRNA expression in turn leads to a decrease in the reuptake of serotonin through the serotonin transporter. This causes an increase in the serotonin levels in the synaptic cleft. The exons are numbered in roman numerals I through XIV; in Figure 1 the same exons are labeled in Arabic numerals 1 through 14 (figure from Canli *et al.*, 9).

Population studies and other less common variants of the 5-HTTLPR gene

Studies have shown that the highest frequency of the L allele is observed in the South African population, 78.36 %, while the S allele is observed most commonly in the Asian population, approximately 72% (10). Among Caucasians, the L allele is observed in 61% and the S allele in 38% of the population (5).

Extra long (XL) and extra short (XS) alleles have been reported. XL variants with repeats ranging from 17 to greater than 20 and XS variants with repeats as low as 11 have been detected. The XL, L, S and XS alleles differ in the repeat regions present (Figure 5). The XS_{11} variant exhibits mRNA expression similar to the S and the L_G variants. The XL_{17} and XL_{18} variants that

have been characterized have the 'A' SNP at rs25531 resulting in XL_{A17} and XL_{A18}, respectively. The XL_{A17} variant exhibited a decreased level of transcriptional efficiency when compared to the L_A allele (5).

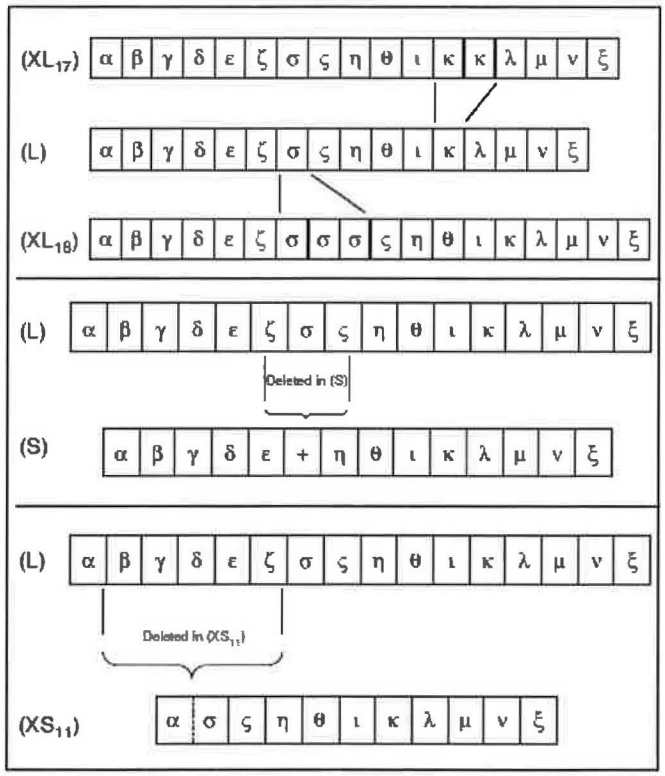


Figure 5: Repeats observed in XL_{17}, L, XL_{18}, S and XS_{11} compared to L. In this figure, XL_{17} has a duplicate of the κ repeat and XL_{18} has a triplicate of the σ repeat. The S allele lacks the ζ, σ, and ς repeats. XS_{11} lacks the β, γ, δ, ε and ζ repeats (figure modified from Ehli *et al.*, 5).

5-HTTLPR and its association with alcohol consumption

Consumption of alcohol and the risky behaviors that follow have been associated to variation within the 5-HTT gene (11). A study was performed in which subjects described their own alcohol use patterns. The subjects were interviewed as they were leaving a bar and reported on their intent to drive a motor vehicle. These subjects submitted saliva samples for genotyping of 5-HTTLPR. Using a distinction of only the S versus L alleles, it was observed that S homozygous individuals drank more than S, L heterozygous individuals and L homozygous individuals drank the least amount of alcohol. S homozygous individuals were three times more likely to intend to drive than L homozygous individuals, indicating a significant increase in risky behavior. Individuals with at least one S allele consumed more drinks at the bar and had higher blood alcohol levels than L homozygous individuals. Young adults with a high alcohol tolerance frequently carry at least one S allele (12). These analyses indicate that the presence of an S allele can be linked to higher levels of intoxication and possible risky behaviors after intake of alcohol.

Based on these studies, the Center for Alcohol and Drug Studies at the San Diego State University Research Foundation requested that DUI/DWI offenders with one arrest and two or more arrests be genotyped for 5-HTTLPR and SNP rs25531. An assay was designed that could characterize alleles as follows: XS, S, L_A, L_G and XL. The predominant alleles in all studies have been S, L_A and L_G.

5-HTTLPR and its association with other risky behaviors

The variations in the repeat regions of the 5-HTT gene lead to phenotypes that affect expression and have been associated with risky behaviors, diseases and disorders including anxiety, depression, and substance abuse (13). The S variant is associated with anxiety or

depression (6). The XL and XS genotypes, although rare, have been associated with attention deficit hyperactivity disorder (5). A thorough understanding of the 5-HTTLPR variants and their association with risky behaviors could help understand and treat several disorders.

Hypothesis: Subjects that have exhibited risky behavior such as driving while intoxicated or driving under the influence of alcohol will have a minimum of one S allele and/or one allele with the L_G phenotype.

Specific Aims:

Aim 1: Develop an assay to characterize 5-HTTLPR and its SNP variant at position rs25531 in a single assay.

Aim 2: Determine the genotypes of 100 known DWI and DUI offenders using the optimized assay.

CHAPTER II

MATERIALS AND METHODS

Subjects

Buccal samples were collected from 100 subjects participating in a rehabilitation program at the Center for Alcohol & Drug Studies at San Diego State University Research Foundation. All subjects had been arrested one or more times for drinking and driving.

Sample collection

Participants were instructed to self-collect two buccal swabs, one from each inner cheek. Each cotton swab was rubbed against the inner cheek while the opposite hand was pressing against the outside of the cheek. Both inner cheeks were swabbed with a different cotton applicator. The two buccal swabs were then placed inside a collection envelope and shipped to the University Of North Texas Health Science Center (Appendix A). A research coordinator labeled the outside of the package with a de-identified number.

DNA extraction

DNA was extracted from 94 samples using the DNA IQ™ extraction with the Freedom EVO® 100 (TECAN, San Jose, CA) protocol. The Freedom EVO® 100 allows for rapid and consistent isolation of DNA. The extracted DNA was normalized to approximately 1ng/μL (14). DNA was extracted from six samples using the DNA IQ™ Reference sample kit for Maxwell ®

16 (Promega, Madison, WI) (15). The Maxwell® 16 allows for rapid and consistent isolation of DNA. The extracted DNA was normalized to approximately 1ng/μL.

Amplification

The 100 extracted samples were amplified for 5-HTTLPR. The protocol to amplify the DNA from the samples was a modification of the 2012 UNTHSC procedure (16). The modifications were a different reaction master mix, a different procedure for digesting the amplified product to characterize the rs25531 SNP, a different electrophoresis protocol and different data analyses.

The primers used for the amplification, the rs25531 SNP and the 43bp indel are represented in Figure 6. Table 1 represents the primers and the fluorophore attached to the 5' end of the forward primer.

```
  1 ggcgttgccg ctctgaatgc cagcacctaa cccctaatgt ccctactgca gccctcccgg
 61 catccccct gcaacctccc agcaactccc tgtaccctc ctaggatcgc tcctgcatcc
121 cccattatcc ccccttcac cctcgcggca tccccctgc acccccRgc atccccctg
181 cagccccccc agcatctccc ctgcacccc agcatccccc ctgcagccct tccagcatcc
241 ccctgcacct ctcccaggat ctcccctgca acccccatta tccccctgc acccctcgca
301 gtatcccccc tgcacccccc agcatccccc catgcacccc cggcatcccc cctgcacccc
361 tccagcattc tccttgcacc ctaccagtat tcccccgcat cccggcctcc aagcctcccg
421 cccaccttgc ggtccccgcc ctggcgtcta ggtggcacca gaatcccgcg cggactccac
481 ccgctgggag ctgccctcgc ttgcccgtgg ttgtccagct cagtccctct agacgctagc
541 ccaaccggcc gcacagtttt cagggtcagt tcctccaagt acaaggggcg gtggcttctc
```

Figure 6: Depiction of the primers, indel region and SNP rs25531. Sequence of 5-HTT gene representing the forward primer (written in green), the reverse primer (written in red),the 43bp indel region (highlighted in blue) and the SNP. The rs25531 SNP is represented by **R** (a pu**R**ine, either A or G) at position 168 in the GenBank reference AF126506.

Primer	Sequence	Fluorophore
Forward	5'-ATGCCAGCACCTAACCCCTAATGT-3'	6-FAM™
Reverse	5'-GAGGGACTGAGCTGGACAACCAC-3'	Not applicable

Table 1: Primers used to amplify the 5-HTTLPR gene.

For the amplification setup, a ready-to-use master mix CleanAmp™ GC-Rich 2X PCR Master Mix (TriLink Biotechnologies, San Diego, CA) (further abbreviated in this paper as CleanAmp), was used. CleanAmp is a hot-start reaction pre-mix which combines deoxynucleotide triphosphates (dNTPs), CleanAmp™ 7-deaza-dGTP (TriLink Biotechnologies) and *Thermus Aquaticus* DNA polymerase (Taq). The amplification master mix includes the reagents described in Table 2.

Reagents	Concentration per reaction	Volume per reaction (μL)
CleanAmp GC-Rich 2X PCR Master mix	1X	12.5
Forward Primer (12μM)	0.38μM	0.8
Reverse Primer (12μM)	0.38μM	0.8
Template	3ng	3
Water*	-	8.9
	Total volume	**25**

Table 2: Reagents used for the amplification of 5-HTTLPR gene. *Amount of water added can be adjusted if more than 3μL template is added.

Touchdown polymerase chain reaction (TD PCR) cycling parameters were followed for amplification (Table 3). TD PCR helps optimize amplification for GC-rich regions. It uses a

gradual decrease in the annealing temperature to reach the optimal cycling parameters, this allows the primer to bind to the secondary structures and hairpins that could form in GC-rich regions.

5-HTTLPR Amplification		
Number of cycles	**Temperature**	**Time (min:sec)**
HOLD	95°C	10:00
5	95°C	0:30
	61°C	0:30
	72°C	1:30
5	95°C	0:30
	60°C	0:30
	72°C	1:30
30	95°C	0:30
	59°C	0:30
	72°C	1:30
HOLD	72°C	30:00
HOLD	4°C	∞

Table 3: Thermal cycling conditions followed for amplification.

*Msp*I restriction enzyme digestion

After amplification, the samples were prepared for digestion. *Msp*I (New England Biolabs, Ipswich, MA) restriction enzyme was used to digest the amplified product to characterize the A to G SNP, rs25531. The enzyme cuts after the first C in any CCGG motif. Hence, the enzyme will not cut the amplified product if the CCAG motif is present. This distinguishes the A and G rs25531 position. See Figures 2 and 3.

The amplified products were prepared for enzyme digestion by heat inactivation of any restriction endonucleases prior to digestion. The optimal temperature for heat inactivation for restriction endonucleases is 65°C (Table 4).

Digestion Preparation		
Number of cycles	**Temperature**	**Time (min:sec)**
HOLD	95°C	10:00
HOLD	65°C	30:00
HOLD	4°C	∞

Table 4: Parameters for preparation of amplified samples for *Msp*I digestion.

After performing the digestion preparation step, the samples were digested using *Msp*I and NEBuffer 4 (New England Biolabs), the optimal buffer for *Msp*I restriction enzyme activity. (Table 5).

Reagents	**Concentration per reaction**	**Volume per reaction (µL)**
Sterile double distilled water	-	6.8
*Msp*I (20,000U/mL)	4U	0.2
NEBuffer 4 (10X)	1X	1
Amplified product	-	2
	Total volume	**10**

Table 5: Reagents for *Msp*I digestion.

The samples were digested in a 96-well plate; this same plate is used for the capillary electrophoresis setup. By using the same plate, there is a reduction in the number of tube transfers which in turn reduces the chances of contamination, sample waste, and costs associated with the consumables. *Msp*I digestion conditions are optimal at 37°C enzyme activation and digestion for the CCGG recognition site. (Table 6)

***Msp*I Digestion**		
Number of cycles	**Temperature**	**Time (min:sec)**
HOLD	37°C	180:00
HOLD	65°C	20:00
HOLD	4°C	∞

Table 6: *Msp*I digestion setup.

Capillary electrophoresis

After the samples were digested, they were prepared for capillary electrophoresis (CE) on the ABI PRISM® 3130*xl* Genetic Analyzer (Applied Biosystems, Foster City, CA) (further abbreviated as 3130*xl*). The master mix for the CE was prepared as shown in Table 7

Reagent	**Volume per reaction (µL)**
Hi-Di™ Formamide	20
GeneScan™ 600 LIZ® Size Standard	0.8
Total	**20.8**

Table 7: Reagents for CE set up.

A total of 20.8µL of the CE master mix was added directly to the 10µL cut product in the 96-well plate.

The original amplified product was denatured at 95°C for 4 minutes and placed on ice for 4 minutes to snap cool the product. A total of 2µL of the amplified product was immediately added in the plate. This was a novel procedure and allowed characterization of the 5-HTTLPR and its SNP variant in a single assay.

After the CE plate was setup, it was placed on the 3130*xl*. The instrument protocol for 5-HTTLPR is shown in Table 8.

Parameter	**5-HTTLPR**
Oven Temp	60°C
Injection voltage	3 kVolts
Injection time	20 secs
Run voltage	15 kVolts
Run time	3900 secs

Table 8: Protocol parameters used for CE run for characterization of multiplexed 5-HTTLPR and rs25531.

Data analysis

After capillary electrophoresis, the samples were analyzed using the GeneMapper™ *ID* Software Version 3.2.1 (Applied Biosystems). The amplified product was mixed with the digested product; this allowed for characterization of 5-HTTLPR and its SNP, rs25531, in a single capillary electrophoretic injection. Table 9 lists each of the alleles and the approximate sizes.

S uncut	L uncut	S cut	L_A cut	L_G cut
424bp	462bp	252bp	291bp	132bp

Table 9: Alleles observed and their approximate sizes.

Table 10 represents the approximate sizes at which peaks were expected for each of the genotype.

Allele / Genotype	Uncut		Cut		
	S 424bp	L 462bp	S 252bp	L_A 291bp	L_G 132bp
S	Yes	No	Yes	No	No
L_A	No	Yes	No	Yes	No
L_G	No	Yes	No	No	Yes
S, L_A	Yes	Yes	Yes	Yes	No
S, L_G	Yes	Yes	Yes	No	Yes
L_A, L_G	No	Yes	No	Yes	Yes

Table 10: The expected peaks for each allele and genotype are labeled "Yes". No peak should be seen in the fields labeled "No".

5-HTTLPR is a GC-rich gene. During the analysis of the samples, non-specific amplification peaks were observed which made distinguishing the true peak from the non-specific peak difficult. Guidelines were developed to help genotype the samples. Appendix C displays a representative image for each of the genotypes observed. Genotype L_G was not observed; hence, it is not represented in the appendix.

Statistical Analysis

The offender status of each of the genotyped individuals was obtained from San Diego State University. The individuals were all enrolled in the First Offender Program or the Multiple Offender Program. Chi squared test for independence was performed using IBM® SPSS Software (Armonk, NY) to determine whether the individuals with moderate and high risk genotypes were more likely to be enrolled in the Multiple Offender Program instead of the First Offender Program.

CHAPTER IV

RESULTS

A total of 100 samples were received from the Center for Alcohol and Drug Studies at the San Diego State University Research Foundation from January 2012 to May 2013. These samples were extracted, amplified, digested, multiplexed, electrophoresed, and genotyped. The genotype of each of the sample is represented in Table 11.

Sample #	Genotype	Sample #	Genotype	Sample #	Genotype
10202	S, L_A	12771	S, L_A	12695	S, L_G
11996	S	12651	S	12656	S, L_A
11228	S	12432	S, L_A	12065	S, L_A
11308	S, L_A	12901	Inconclusive	12860	S, L_A
11080	S	12994	Inconclusive	12177	S, L_A
11961	S	12244	S, L_A	11799	S
11420	S	12667	S, L_A	12789	Inconclusive
10796	S	12339	L_A	12198	S
11176	S, L_G	12968	S	12061	L_A
11999	S, L_A	12936	L_A	12462	S
10127	S	10506	S	12604	S, L_A
10226	S	11317	L_A, L_G	12402	S
ABCDE	S	11520	S, L_A	12605	S
11363	S	10903	S	12582	S, L_A
10286	S	11207	S, L_A	12310	S, L_A
10620	L_A	10548	L_A	12797	L_A
11958	L_A	10642	S, L_G	12882	S, L_A
10718	S, L_A	10006	L_A	12245	S
10499	S	12000	L_A	12175	S
11570	S, L_A	11268	L_A	12775	S, L_A
11600	S, L_A	11434	S, L_A	12522	S, L_G
10300	S, L_G	10268	S, L_A	12581	L_A, L_G
11743	S	11461	L_A	12921	S
11657	S	11321	L_A, L_G	12020	L_A
10955	S, L_A	10522	Inconclusive	12896	L_A, L_G
11587	Inconclusive	10766	S	12281	S, L_A
11790	S, L_G	10183	S	12027	S, L_A
10213	L_A	11435	S, L_G	12384	S
10776	S, L_G	12984	L_A	12541	L_A
10056	S, L_G	12242	S	12633	S, L_G
10910	S, L_A	12236	S	12792	L_A
12769	XS, L_A	12884	S	12590	S, L_A
12554	Inconclusive	12858	L_A, L_G		
12366	S, L_A	12628	L_A		

Table 11: Genotypes of the 100 samples received from Center for Alcohol and Drug Studies. Each of the samples has a unique identifier.

The number of samples that have each of the genotypes is represented in Table 12.

Genotype	Number of samples	Risk Behavior
S	32	High
S, L_G	10	
L_G	0	
S, L_A	29	Moderate
L_A, L_G	5	
L_A	17	Low
XS, L_A	1	Unknown
Inconclusive	6	N/A
Total	**100**	

Table 12: Number of samples that have each of the genotypes of the 5-HTTLPR gene and the risk behavior associated with them. N/A= Not Applicable

The S and L_G alleles have been associated with high risk behavior, while the L_A allele has not been associated with high risk behavior (11, 17). Out of the 100 samples, 32 were S and 10 were S, L_G giving a total of 42 samples which had the high risk genotype. A total of 29 samples had the S, L_A genotype and 5 had the L_A, L_G genotype giving a total of 34 samples that had the moderate risk genotype. Only 17 samples had L_A, the low risk genotype.

A total of 76 samples had at least one of the risk alleles. This is consistent with the hypothesis that subjects that have exhibited risky behavior such as driving while intoxicated or driving under the influence of alcohol will have a minimum of one S allele and/or one allele with the L_G phenotype. One sample was genotyped as XS, L_A based on the sizes of the peaks observed. A total of 6 samples were called inconclusive. Excessive non-specific amplification of these samples made genotyping difficult. These samples should be re-amplified.

The offender status of each of the individuals was obtained from San Diego State University. A Chi-squared test for independence was performed using IBM® SPSS software to

determine whether subjects with moderate and high risk genotypes were more likely to be multiple DWI/DUI offenders versus first time offenders. A p-value of 0.16 was obtained which was greater than the alpha value of 0.05. Hence, the results were not significant; no conclusions could be drawn about the genotype, the risk behavior and its association with offender status. An interesting observation was made: Out of the 17 individuals with the low risk genotype, a higher percentage of them were first time offenders. Whereas out of the 42 individuals with the high risk genotype, a higher percentage of them were multiple offenders. This may indicate that the high risk genotype individuals would more likely be multiple offenders; however, more individuals should be genotyped to make a strong association.

Macros were written to create panels and bins for 5-HTTLPR using GeneMapper® *ID* software v 3.2.1 to automate genotyping. Three panels were created for Uncut S, Uncut L and HTT cut, respectively. Within the Uncut S panel, bins were created for XS and S alleles. Within Uncut L panel, a bin was created for the L allele and within the HTT cut panel, bins were created for XS cut, S cut, L_A and L_G alleles.

CHAPTER V

CONCLUSIONS

A total of 94 samples received from the Center for Alcohol and Drug Studies at the San Diego State University Research Foundation were successfully genotyped using the novel assay developed at UNTHSC. There were six samples that were called inconclusive. The method is the only known method that allows genotyping of 5-HTTLPR and the SNP at rs25531 in a single assay. The ready-to-use CleanAmp™ GC-Rich 2X PCR Master Mix was used to amplify the 5-HTTLPR gene. This reduced the number of reagents needed and saved time in amplification set up. The digestion of the amplified product and capillary electrophoresis setup was conducted in the same 96-well plate; this reduced the number of tube transfers.

Previously, sequencing was performed by Eric Ehli (5) in order to identify the repeats present in each of the alleles commonly observed for 5-HTTLPR. A total of ten extracted DNA samples with known genotypes were provided by Ehli to use as controls to identify the sizes of each of the alleles, after which the DWI/DUI offender samples were genotyped. Non-specific amplification was observed, which often made genotyping difficult. Guidelines were developed to make allele calls. Macros were developed using GeneMapper *ID* software to automate genotyping.

Risky behaviors such as the intention to drive after alcohol consumption and illegal substance abuse are known to be associated with 5-HTTLPR (11). The gene has also been associated with several personality traits like depression, anxiety, obsessive compulsive disorder and aggression (18). 5-HTTLPR plays a crucial role in the understanding of human behavior.

Further studies should be performed to develop an allelic ladder to automate genotyping. Previously, studies have been performed to identify the distribution of the S and L alleles in the different racial groups (19); however there have not been any studies on the distribution of the L_A and L_G alleles. Genotyping a control group consisting of at least 200 individuals from the general population belonging to different racial groups will allow the determination of the distribution of the L_A and L_G alleles. A second control group consisting of individuals who state that they drink but do not drive should be genotyped. Statistical analyses to compare the genotypes observed in the control groups and DWI/DUI offenders will allow the determination of whether there is a significant difference between the frequencies of individuals with a particular genotype.

Lastly, a standard naming convention should be developed for calling of the 5-HTTLPR genotypes. It has been observe that different authors use different naming conventions to make allele calls for 5-HTTLPR. Nakamura has referred to the alleles as 14-A, 14-B and so on depending upon the number of repeats present and the variations in the sequence of the repeats (6) whereas Ehli has referred to the alleles as S, L, XS and XL depending upon the size of each of the alleles (5).

APPENDIX A

INSTRUCTIONS FOR SAMPLE COLLECTION

INSTRUCTIONS/CHECKLIST FOR SAMPLE COLLECTION

SDSU/UNTHSC Collaboration

Only one envelope should be opened at a time.
Collect one person at a time.

1. Complete Consent Form.
2. One swab is used for each cheek. Apply firm (not severe) pressure, using a brushing twisting motion (approximately 20 strokes), in the middle of the right inner cheek. Repeat this step using the other swab on the left inner cheek.

   Note: If collection is performed by someone other than the subject, gloves must be worn.

3. Place both swabs together in the envelope provided. Blue envelopes are used for male subjects and pink envelopes are used for female subjects.
4. Complete subject number on the envelope. The research personnel must initial and date the label on the envelope at the time of collection.
5. Peel cover from adhesive and seal envelope.
6. Store sealed envelopes at room temperature until time of shipment.
7. Ship to UNTHSC DNA Laboratory at the address below:

   Rhonda Roby
   University of North Texas Health Science Center
   Department of Forensic & Investigative Genetics
   DNA Laboratory
   3500 Camp Bowie Blvd.
   Fort Worth, TX 76107
   817-735-2462

APPENDIX B

PROCEDURE FOR 5HTTLPR GENOTYPING AND SNP CHARACTERIZATION

Manual 5-HTTLPR Amplification using CleanAmp GC-rich 2X PCR Master Mix and Digestion Setup

Purpose: To manually prepare amplifications of the 5-HTTLPR region of chromosome 17, serotonin transporter gene. This amplification reaction is 25 µL. Amplified product is then digested with the MspI restriction endonuclease. The MspI restriction endonuclease cleaves dsDNA at 5'-C▼CGG-3'. Amplified and digested products are mixed for a single electrophoretic injection.

Equipment and Supplies

- Centrifuge, vortex
- Pipettors and pipette tips (aerosol barrier)
- 96-well plates/strip tubes and base supports
  - Aluminum foil plate seals or strip caps
  - 1.5mL or 2mL microcentrifuge tubes
  - GeneAmp® PCR System 9700

Safety

Gloves, lab coats and eye protection must be worn during this procedure.

Reagents

- UV irradiated molecular biology grade DNAse free H_2O (ddH_2O)
- CleanAmp GC-Rich 2X PCR Master Mix
- Primer HTT-F (12µM)
  5'- ATGCCAGCACCTAACCCCTAATGT -3'
- Primer RevHTT-R (12µM)
  5'- GAGGGACTGAGCTGGACAACCAC-3'
- NEBuffer 4 (1X)
- *Msp*I restriction endonuclease (20U/µL)
- Hi-Di™ Formamide
- GeneScan™ - 600 LIZ

Procedure

A. Amplification Master Mix Calculations and Preparation

1. Prepare the master mix as shown in the 5-HTTLPR amplification table
2. Calculate the volume needed for each component with a pipetting overage factor included.

Total volume of reagent	=	Volume of specified reagent needed per sample	x	Number of samples (*N*)	x	1.1 (Pipetting overage factor)

5-HTTLPR Amplification

Master Mix for **3µL** Template Addition
(Example, $N = 10$)

Reagent	*Vol. per Sample (µL)*	*Total Volume (µL)*
CleanAmp™ GC-Rich 2X PCR Master Mix	**12.5**	**125**
Forward Primer	**0.8**	**8**
Reverse Primer	**0.8**	**8**
ddH_2O	**7.9**	**79**

* the volume of ddH_2O can be adjusted if more than 3µL of DNA is needed.

3. In the no template hood, prepare the necessary master mixes for the amplifications. Add the reagents in order as shown. Dispense 22µL of master mix into each well or tube, including the wells designated for positive, negative and reagent blank controls.
4. In the template addition hood, add 3µL of ddH_2O to the negative control well, add 3µL template to each of the sample wells and the reagent blank, add 3µL of control DNA to the positive control well(s).

   Note: For optimal amplification, the template DNA is 2.5 to 3µL of 1ng/µL.
5. Seal the tubes/plate and briefly centrifuge.
6. Place the tubes in the retainer on the thermal cycler or the plate directly on the thermal cycler; record the thermal cycler number on the worksheet.

B. Thermal Cycling Parameters

1. Select the appropriate program.
2. Enter 25µL for the reaction volume.
3. Select START.
4. After the run is complete, samples can be electrophoresed or stored 14 days at 4°C. If digesting the PCR product, proceed to Section C.

5-HTTLPR Amplification		
Number of cycles	**Temperature**	**Time (min:sec)**
HOLD	95°C	10:00
5	95°C	0:30
	61°C	0:30
	72°C	1:30
5	95°C	0:30
	60°C	0:30
	72°C	1:30
30	95°C	0:30
	59°C	0:30
	72°C	1:30
HOLD	72°C	30:00
HOLD	4°C	∞

C. Sample Preparation for *Msp*I Digestion

1. Place the tubes in the retainer on the thermal cycler or the plate directly on the thermal cycler; record the thermal cycler number on the worksheet.
2. Select the appropriate program.
3. Enter 25µL for the reaction volume.
4. Select START.
5. After the run is complete, samples must be kept on ice and digested immediately.

Digestion Preparation		
Number of cycles	**Temperature**	**Time (min:sec)**
HOLD	95°C	10:00
HOLD	65°C	30:00
HOLD	4°C	∞

D. *Msp*I Digestion

1. Prepare the master mix as shown in *Msp*I digestion table.

*Msp*I Digestion

Master Mix for **2µL** Template Addition
(Example, $N = 10$)

Reagent	*Vol. per Sample (µL)*	*Total Volume (µL)*
sterile ddH_2O	**6.8**	**74.8**
1X NEBuffer 4	**1.0**	**11**
*Msp*I	**0.2**	**2.2**

1. In the no template area, prepare the necessary master mixes for the amplifications. Add the reagents in order as shown. Dispense 8µl of master mix into each well or tube.
2. In the template addition hood, add 2µL of PCR product to each of the sample wells.

Note: PCR product includes samples and all associated controls.

3. Seal the plate and briefly centrifuge.
4. Place the tubes in the retainer on the thermal cycler or the plate directly on the thermal cycler; record the thermal cycler number on the worksheet.
5. Select the appropriate program.
6. Enter 10µL for the reaction volume.
7. Select START.
8. After the digestion is complete, samples must be electrophoresed immediately. Proceed to Section E.

MspI Digestion		
Number of cycles	**Temperature**	**Time (min:sec)**
HOLD	37°C	180:00
HOLD	65°C	20:00
HOLD	4°C	∞

E. Preparation for Capillary Electrophoresis

1. Prepare the master mix as shown in the capillary electrophoresis table.

Capillary Electrophoresis

Master Mix for Template Addition
(Example, $N = 10$)

Reagent	*Vol. per Sample (µL)*	*Total Volume (µL)*
Hi-Di™ Formamide	20	220
GeneScan™ - 600 LIZ™	0.8	8

2. In the no template area, prepare the necessary master mixes for the amplifications. Add the reagents in order as shown. Dispense 20.8µL of master mix into each well of the 96-well plate which contains the cut product.
3. In the template addition hood, add 2µL uncut PCR product.

   Note: Denature the uncut products for 4 minutes at 95°C and snap cool on ice for another 4 minutes, prior to adding to the 96-well plate.
4. Briefly centrifuge plate to remove any residual air bubbles; visually inspect for bubbles. Prepare plate assembly by placing the sample plate into the black base plate and snapping the white plate retainer onto the plate and plate base.

   Note: Ensure that the holes of the white plate retainer align with the holes of grey plate septa. Otherwise, damage to capillary array tips will occur.

   Note: Refer to the Maintenance and Use of the 3130*xl* Genetic Analyzer protocol for steps to create a plate record and start the fragment analysis run. In the plate document, select 5-HTTLPR for both the Results Group and Instrument Protocol.

References

1. Anchordoquy, H., McGreary C., Lui, L., Krauter, K., and Smolen, A. Genotyping of Three Candidate Genes After Whole-Genome Preamplification of DNA Collected from Buccal Cells. *Behavior Genetics*; 2003 Jan; 33(1):73-8.

2. Ehli, E., Hu, Y., Lengyel-Nelson, T., Hudziak, J., and Davies, G. Identification and functional characterization of three novel alleles for the serotonin transporter-linked polymorphic region. *Molecular Psychiatry*; 2011 Feb; 17(2):185-92

3. Lemogne, C., Gorwood, P., Boni, C, Pessiglione, M., Lehericy, S., and Fossati, P. Cognitive Appraisal and Life Stress Moderate the Effects of the 5-HTTLPR Polymorphism on Amygdala Reactivity. *Human Brain Mapping*; 2011 Jan; 32:1856-67.

Revisions

Date of revision	Revised by	Description of changes made

APPENDIX C

LIBRARY OF GENOTYPES COMMONLY OBSERVED FOR 5-HTTLPR

Library of genotypes commonly observed for 5-HTTLPR

There are six commonly observed alleles for 5-HTTLPR.

S uncut	L uncut	S cut	L_A cut	L_G cut
424bp	462bp	252bp	291bp	132bp

Table 1: Commonly observed alleles and their approximate sizes.

Each of these alleles can occur in combination with another allele, yielding six different genotypes.

Allele / Genotype	Uncut		Cut		
	S 424bp	**L 462bp**	**S 252bp**	**L_A 291bp**	**L_G 132bp**
S	Yes	No	Yes	No	No
L_A	No	Yes	No	Yes	No
L_G	No	Yes	No	No	Yes
S, L_A	Yes	Yes	Yes	Yes	No
S, L_G	Yes	Yes	Yes	No	Yes
L_A, L_G	No	Yes	No	Yes	Yes

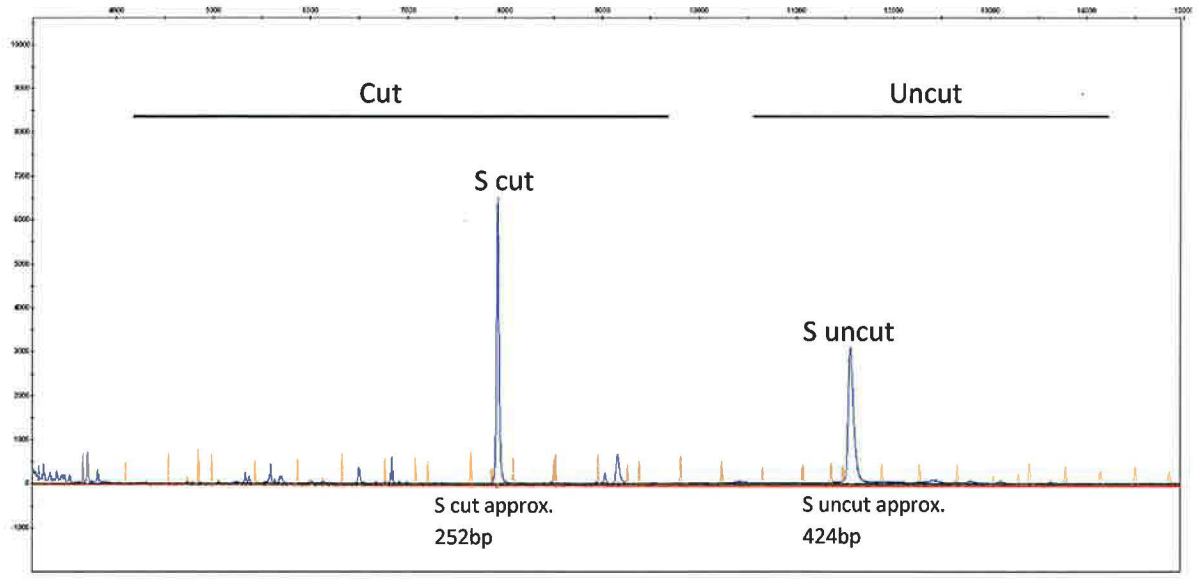
Table 2: The six commonly observed alleles and genotypes

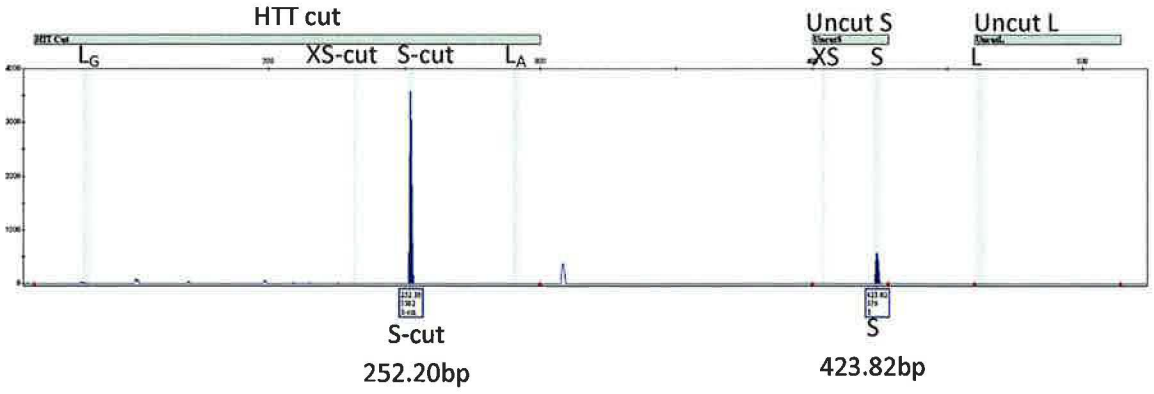
Note: Genotype L_G was not observed, hence it is not included in the library.

All the data is reviewed in the raw data view of GeneMapper ID software, after which it analyzed using the panels and bins developed.

S Homozygous

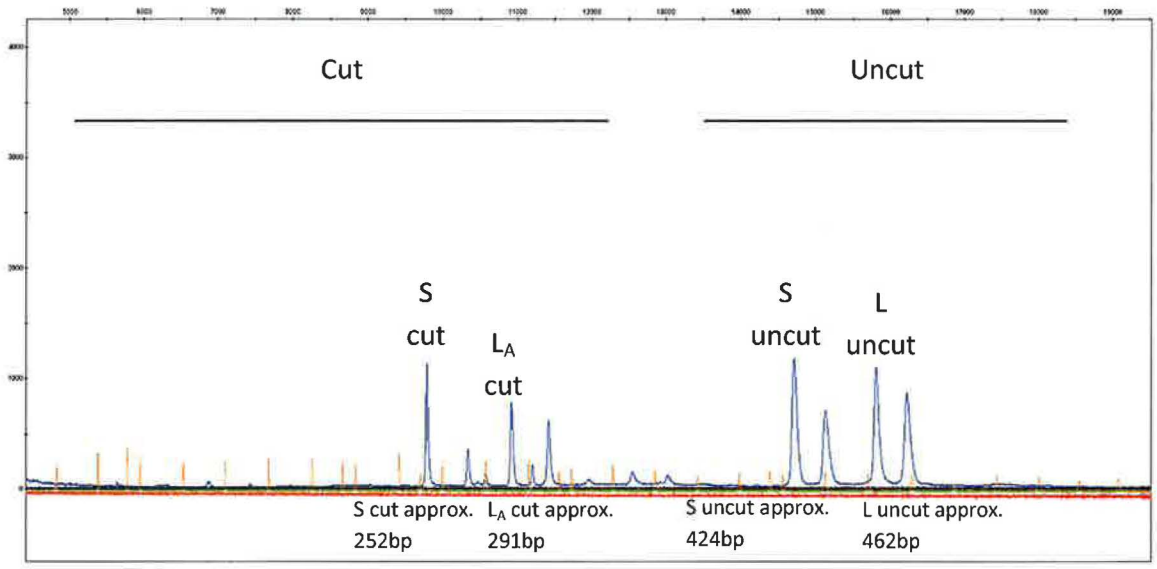
A peak is observed at approximately 424bp for the S allele in the uncut region and 252bp for the S allele in the cut region.

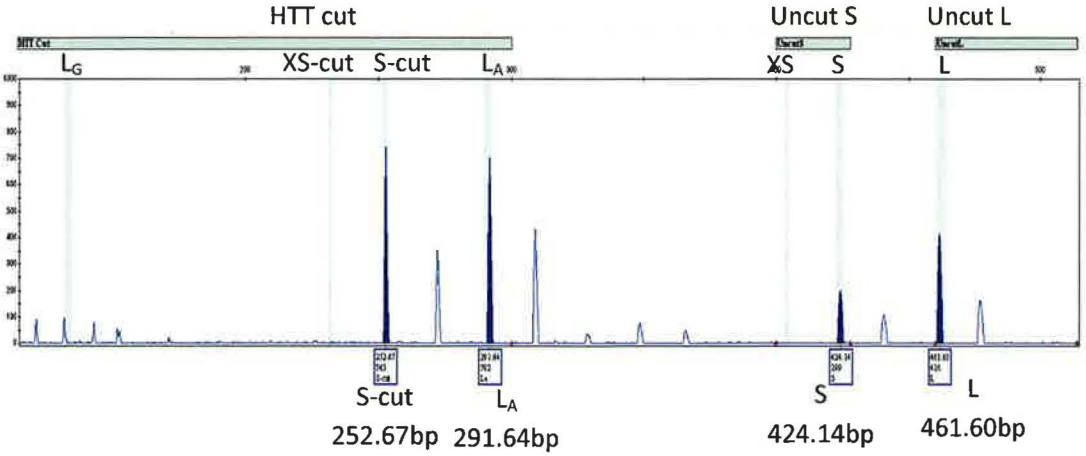




S, L_A

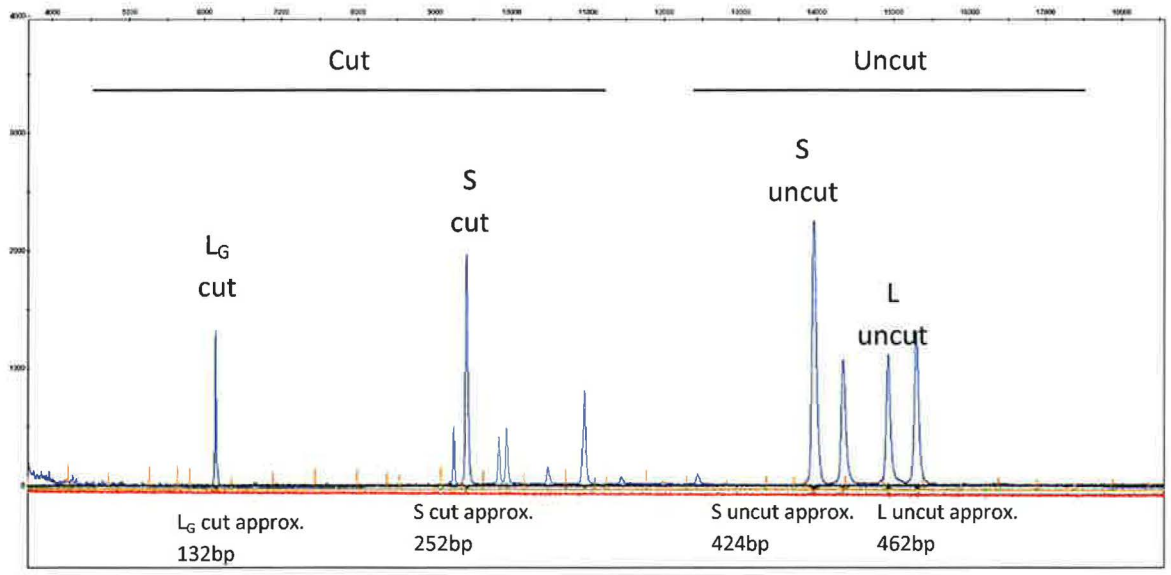
A pattern of four peaks is observed at approximately 424bp, 440bp, 462bp and 480bp, respectively, in the uncut region. The peaks at approximately 424bp and 462bp are the true peaks; the others are non-specific targets that amplify with every S, L_A genotype. In the cut region, peaks are observed at approximately 252bp for the S cut allele and 291bp for L_A cut allele.

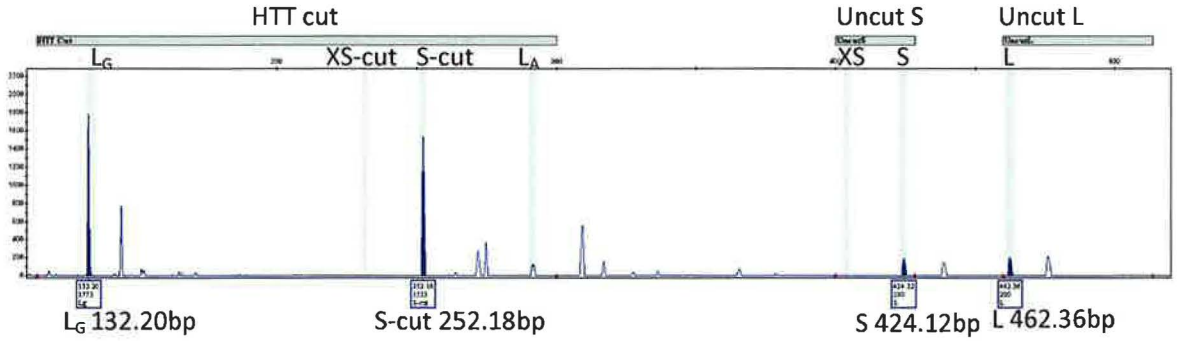




S, L_G

A pattern of four peaks is observed at approximately 424bp, 440bp, 462bp and 480bp, respectively, in the uncut region. The peaks at approximately 424bp and 462bp are the true peaks; the others are non-specific targets that amplify with every S, L_G genotype. In the cut region, peaks are observed at approximately 252bp for the S cut allele and 132bp for L_G cut allele.

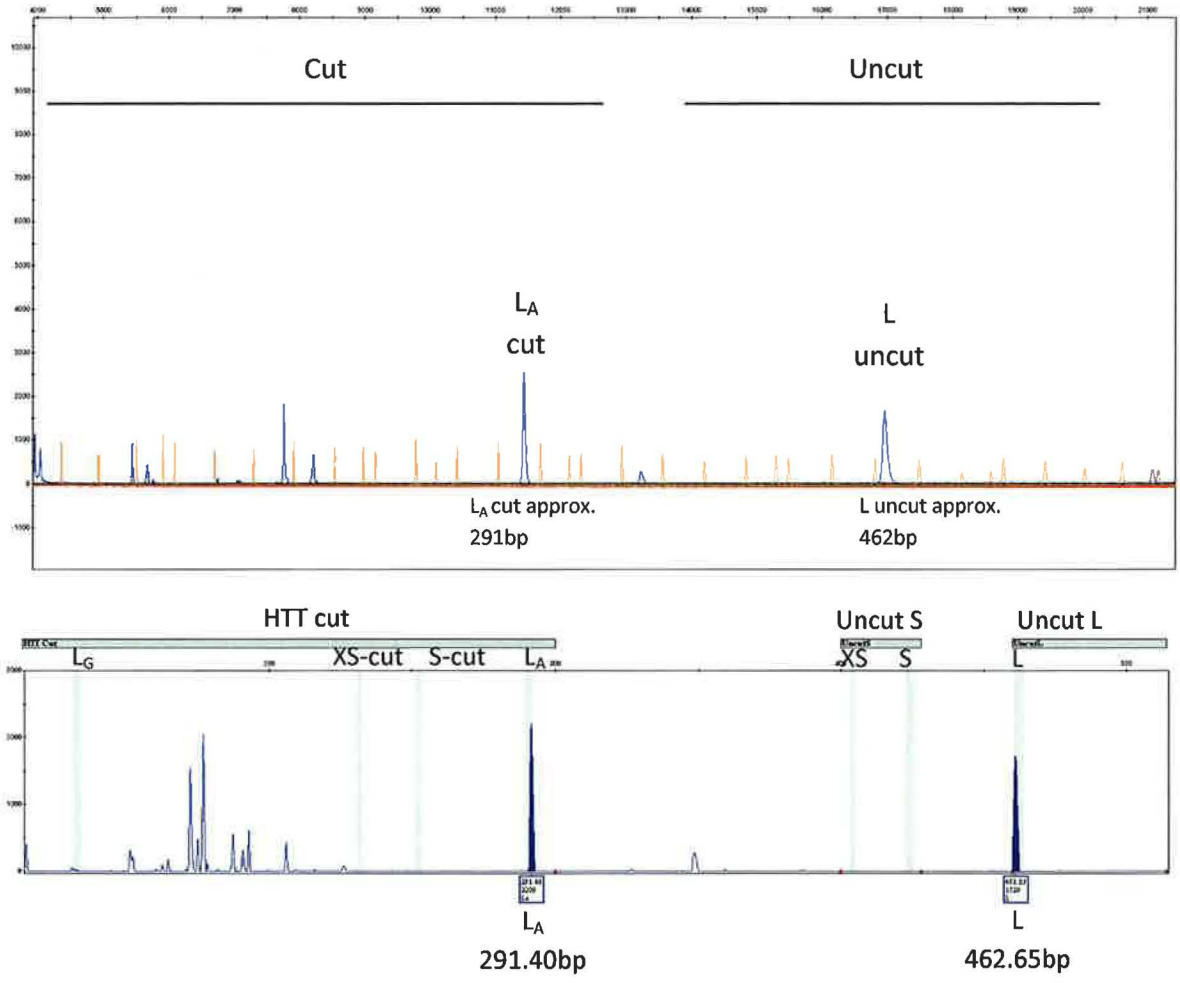




Note: If peaks are observed for both L_A and L_G in the HTT cut panel for S, L_A and S, L_G genotypes the peak heights should be compared. The allele with the greater peak height is the true cut allele.

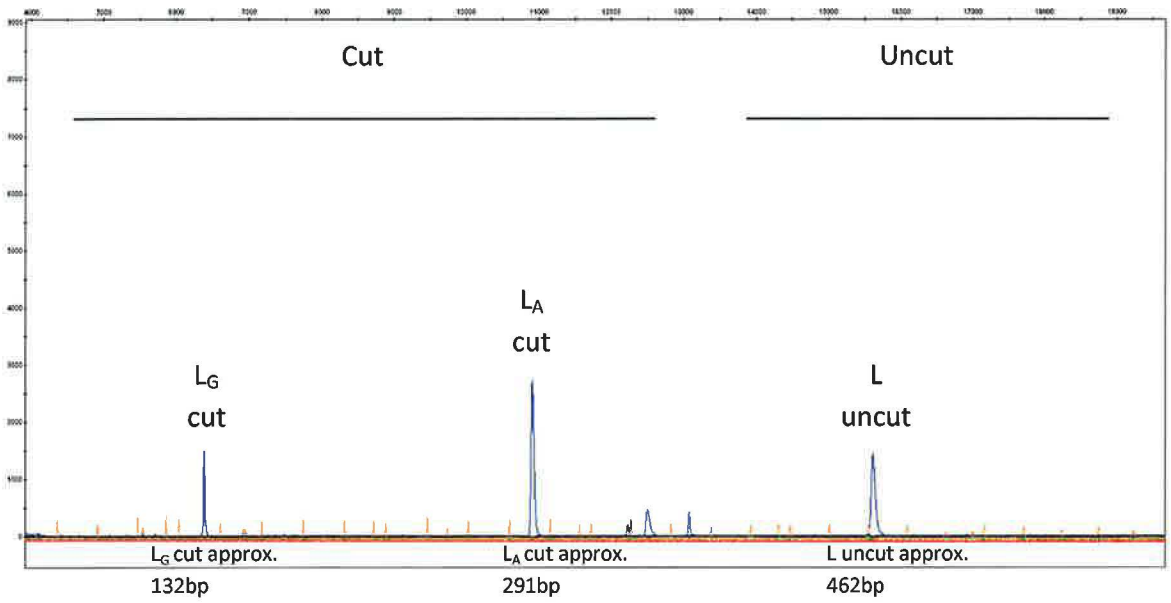
L_A Homozygous

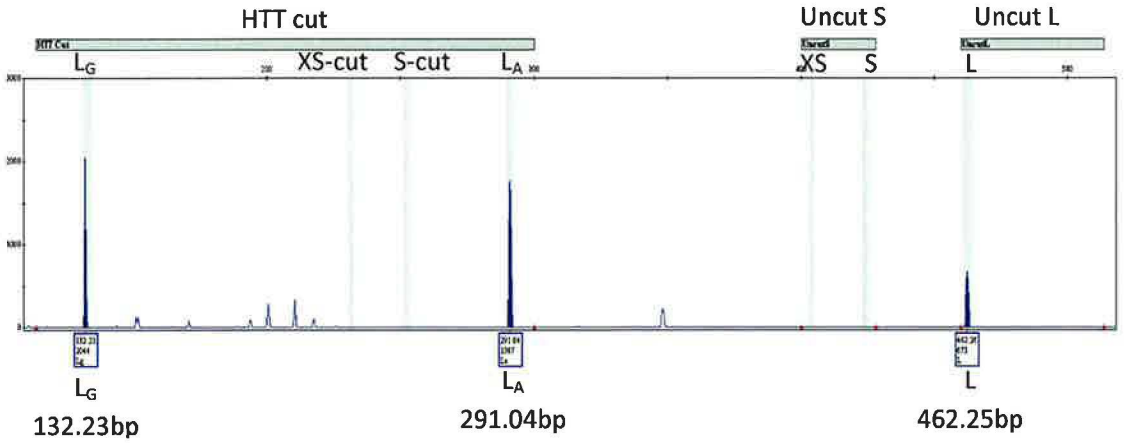
A peak is observed at approximately 464bp for the L allele in the uncut regions and at approximately 291bp for the L_A allele in the cut region.



L_A, L_G

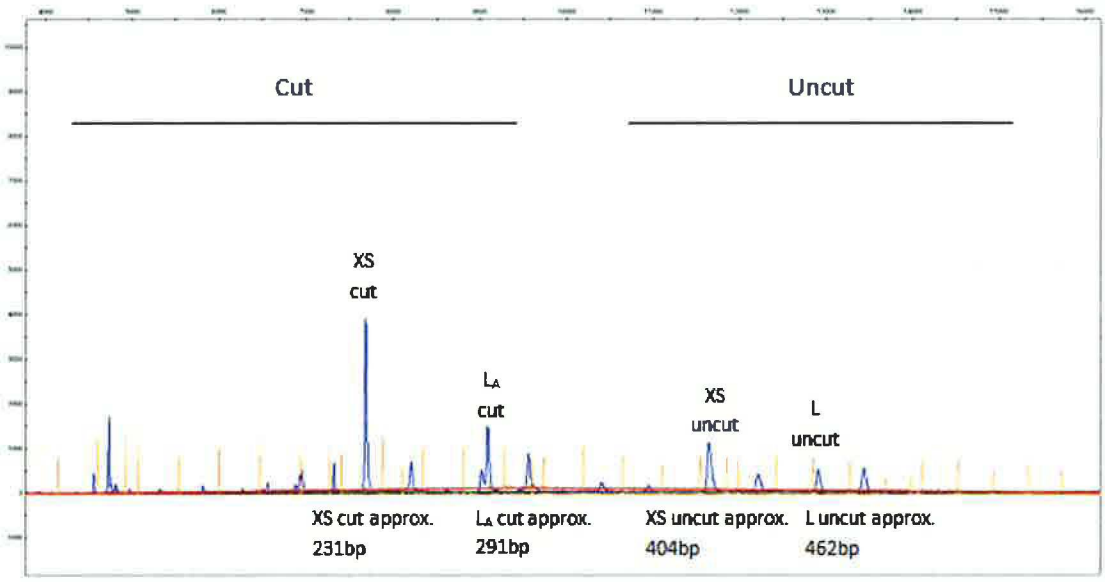
A peak is observed at approximately 462bp for the L uncut allele. Peaks are observed at approximately 291bp for L_A cut allele and 130bp for the L_G cut allele.

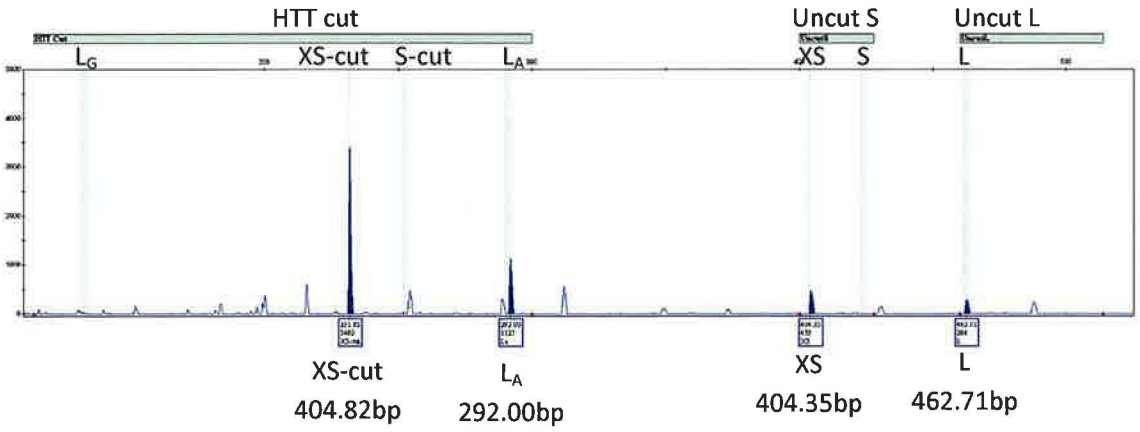




XS, L_A

A pattern of four peaks is observed at approximately 404bp, 430bp, 462bp and 480bp, respectively, in the uncut region. The peaks at approximately 404bp and 462bp are the true peaks. In the cut region, peaks are observed at approximately 231bp for the S cut allele and 132bp for L_G cut allele.





REFERENCES

1. Lewis G, Mulligan J, Wiles N, Cowen P, Craddock N, Ikeda M, et al. Polymorphism of the 5-HT transporter and response to antidepressants: Randomised controlled trial. Br J Psychiatry. 2011 Jun;198(6):464-71.

2. Wray NR, James MR, Gordon SD, Dumenil T, Ryan L, Coventry WL, et al. Accurate, large-scale genotyping of 5HTTLPR and flanking single nucleotide polymorphisms in an association study of depression, anxiety, and personality measures. Biol Psychiatry. 2009 Sept;66(5):468-76.

3. Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D, et al. Allelic variation of human serotonin transporter gene expression. J Neurochem. 1996 Jun;66(6):2621-4.

4. Frey, UH. Bachmann, HS. Peters . Siffert,W. PCR-amplification of GC-rich regions: 'slowdown PCR'. Nature. 2008;3(8):1312-7.

5. Ehli EA, Hu Y, Lengyel-Nelson T, Hudziak JJ, Davies GE. Identification and functional characterization of three novel alleles for the serotonin transporter-linked polymorphic region. Mol Psychiatry. 2011 Jan;17:185-192.

6. Nakamura MM. The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. Mol Psychiatry. 2000 Jan;5(1):32-8.

7. Hu XZ, Lipsky RH, Zhu G, Akhtar LA, Taubman J, Greenberg BD, et al. Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. Am J Hum Genet. 2006 May;78(5):815-26.

8. Hu XZ. An expanded evaluation of the relationship of four alleles to the level of response to alcohol and the alcoholism risk. Alcohol Clin Exp Res. 2005 May;29(1):8-16.

9. Canli T, Lesch K. Long story short: The serotonin transporter in emotion regulation and social cognition. Nat Neurosci. 2007 Sept;10(9):1103-9.

10. Nasserddine S, Hamzi K, Diakite B, Serbati N, Nadifi F. Distribution of 5-HTTLPR serotonini transporter polymorphism in morrocan population. Int J LifeSc Bt & Pharm. 2012 April;1(2):278-81.

11. Thombs DL, O'Mara RJ, Hou W, Wagenaar AC, Dong HJ, Merves ML, et al. 5-HTTLPR genotype and associations with intoxication and intention to drive: Results from a field study of bar patrons. Addict Biol. 2011 Jan;16(1):133-41.

12. Türker T, Sodmann R, Goebel U, Jatzke S, Knapp M, Lesch K, et al. High ethanol tolerance in young adults is associated with the low-activity variant of the promoter of the human serotonin transporter gene. Neurosci Lett. 1998 Jun;248(3):147-50.

13. Lemogne C, Gorwood P, Boni C, Pessiglione M, Lehericy S, Fossati P. Cognitive appraisal and life stress moderate the effects of the 5-HTTLPR polymorphism on amygdala reactivity. Hum Brain Mapp. 2011 Nov;32(11):1856-67.

14. Plopper F, Roby R, Planz J, Eisenberg A. In: High throughput processing of family reference samples for missing persons programs: The use of robotics in extraction and amplification setup or STR and mtDNA analysis. Proceedings of the 17th international symposium on human identification; Nashville, TN.; 2006.

15. Krnajski, Z. Geering, S. Steadman,S. Performance verification of the Maxwell 16 instrument and DNA IQ reference sample kit for automated DNA extraction of known samples. Forensic Sci Med Pathol. 2007;3(3):264- 269.

16. Powers J. The design of a novel 5-HTTLPR assay for association studies of subjects charged with driving while intoxicated and driving under the influence [dissertation]. University of North Texas Health Science Center; 2012.

17. Lau JY, Goldman D, Buzas B. Amygdala function and 5-HTT gene variants in adolescent anxiety and major depressive disorder. Biol Psychiatry. 2009;65(4):349-55.

18. Irwin L. The spectrum of behaviors influenced by serotonin. Biol Psychiatry. Aug 1998; 44(3):151-62.

19. Noskova T, Pivac N, Nedic G. Ethnic differences in the serotonin transporter polymorphism (5-HTTLPR) in several European populations. Prog Neuropsychopharmacol Biol Psychiatry. 2008 Oct; 32(7): 1735-9.