UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

The prevalence of eight single nucleotide variations in overweight and obese participants

By

Bianca Sansom

Dissertation submitted in fulfilment of the requirements for the degree

Magister Scientiae

in the Department of

Pharmacology

In

the Faculty of Health Sciences

at the

University of Pretoria

Supervisor

Dr Andre Marais

Co-supervisor

Dr Christa North

November 2018

PLAGIARISM DECLARATION

I, Bianca Sansom,

Student number: 04417577

Subject of the work: The prevalence of eight single nucleotide variations in overweight and obese participants

Declaration

1. I understand what plagiarism entails and am aware of the University's policy in this regard.
2. I declare that this project (dissertation) is my own, original work. Where someone else's work was used (whether from a printed source, the internet or any other source) due acknowledgement was given and reference was made according to departmental requirements.
3. I did not make use of another student's previous work and submitted it as my own.
4. I did not allow and will not allow anyone to copy my work with the intention of presenting it as his or her own work.

Signature: 

ACKNOWLEDGMENTS

When thinking of everyone I need to give recognition to for this achievement the list is long. However, at the top of the list is my husband, Andre, thank you for putting up with me through the late nights and early mornings, through the ups and downs and everything in between. You have always been by my side and have given me the support I needed.

To my family, Mom, Dad, Uncle Shaun and both my grannies, a very big thank you, without your continuous support and belief in my abilities, I could not have achieved this. You are always there to give me the push I need to achieve my goals – Thank you.

To my supervisor's Dr Andre Marais and Dr Christa North, thank for your patience. I know I can be a bit pushy sometimes, but your effort and guidance have always been truly appreciated. Without your knowledge and guidance this would not have been possible. To Prof Bekker, thank you for all of your assistance with the statistical analysis as well as your input and article support which has assisted with the write up of the dissertation.

Thank you to Prof Steenkamp and the Department of Pharmacology for affording me this opportunity. Thank you for your flexibility and assistance which has enabled me to complete this degree.

Finally, to GENEWAY™ and Riaan Jacobs, for your support through this journey, affording me the time to attend classes and prepare for tests and examinations. And for enabling me to conduct my research in your biotechnologically advanced laboratory. Thank you for your willingness to promote your employees.

TABLE OF CONTENTS

ABSTRACT

Introduction

Obesity is a growing epidemic not just nationally but worldwide and is responsible for a substantial economic burden in both developed and developing countries. Obesity is a major risk factor for type 2 diabetes, cardiovascular disease, some types of cancer and premature death. It has long been known that there is a genetic link to the complex nature of obesity that has both an environmental and psychological link to it. Neurotransmitters in the brain's reward cascade regulate the feeling of satiety and food cravings, which are associated with behaviours such as overeating and binge eating. The aim of the study was to investigate the prevalence of eight single nucleotide variations (SNVs) associated with the regulation of the brain reward system and have been linked to addictive behaviour and food cravings in an attempt to assess a causal relationship with overweight and obese individuals.

Methods

A total of 247 DNA buccal samples were collected from willing participants. Of the 247 DNA samples collected 223 were analysed, based on the inclusion and exclusion criteria. Of the 223 samples collected 107 participants were of normal weight and 116 were either overweight or obese (class I-III). The TaqMan® OpenArray™ Genotyping platform was utilised to genotype the 223 samples across eight SNVs, namely; *SLC6A4* (rs25531) which encodes for a monoamine transporter protein that transports serotonin from the synaptic cleft to the presynaptic neuron, *HTR2C* (rs3813926) encodes the G-protein coupled receptor that regulates excitatory neurotransmitters, *OPRM1* (rs1799971) provides the instruction for making the mu opioid receptor protein which regulates pain, reward and addictive behaviours, *GABRA6* (rs3219151) encodes for Gamma-aminobutyric acid receptor subunit alpha-6 that functions as an inhibitory neurotransmitter, *DRD2* (rs1800497) encodes the D2 subtype of the dopamine receptor which is a G-protein coupled receptor that inhibits adenylyl cyclase activity, *DRD4* (rs1800955) similarly encodes the D4 subtype of the dopamine receptor, *COMT* (rs4680) provides the instruction for making catechol-O-methyltransferase, an enzyme that controls the levels of certain hormones and *LEPR* (rs1137101) encodes for the Leptin receptor protein, which is involved in the regulation of body weight.

Results

The AG genotype for the Leptin Receptor (LEPR) gene was found to be more prevent in overweight and obese individuals, odds ratio 2.63 [95% CI = 1.13; 6.13]. None of the other SNVs showed any significant association. Upon more stringent review of the study, it was discovered that bias was introduced in several ways, including design, sampling, statistical, procedure and measurement bias.

Conclusion

Although the sample size was statistically determined, it was too small to draw any cause and effect relationship. There was also limited ethnic diversity in the samples collected. The sample size thus decreased the statistical power of the analysis. Obesity is a complex disease, with both genetic and environmental factors need to be taken into account. The lack of environmental and/or lifestyle information of participants narrowed the interpretation of the results. Obtaining the medical history and lifestyle information of participants could have been beneficial at presenting daily challenges or stressor that participant may experience that could affect their weight management. This essentially could have shed light on the genotype discrepancies observed. The fact that an individual's current BMI score is regarded as normal does not eliminate the potential that they struggle with their weight, or have struggled with their weight during their life. The lifestyle information could have changed the classification of the participant's risk, in terms of obesity risk.

Keywords: Obesity, Addiction, Depression, Single Nucleotide Variations, Genotyping, Brain Reward Cascade, Neurotransmitters, Leptin Receptor, *LEPR* and TaqMan.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AGRP	Agouti-related protein
AFR	African
AMR	Ad Mixed American
BMI	Body Mass Index
BP / bp	Base Pairs
CI	Confidence Interval
dbSNP	Single Nucleotide Polymorphism database
DNA	Deoxyribonucleic Acid
DRI	Dopamine Reuptake Inhibitor
DSM	Diagnostic and Statistical Manual of Mental Disorders
EAS	East Asian
EPIC	European Prospective Investigation into Cancer
EUR	European
Gln27	Glutamine at amino acid 27
GWAS	Genome-Wide Associate Studies
HapMap	Haplotype Map
MDD	Major Depressive Disorder
MGB	Minor Groove Binder
NAc	Nucleus Accumbens
NCBI	National Centre for Biotechnology
NCD	Non-communicable Diseases
NHGRI	National Human Genome Research Institution
OR	Odds Ratio
QC	Quality Control
RDS	Reward Deficiency Syndrome
Rs# / Ref SNP	Reference SNP ID Number
SAS	South Asian
SNRIs	Serotonin and Noradrenalin Reuptake Inhibitors

SSRIs	Selective Serotonin Reuptake Inhibitors
SNV(s)	Single Nucleotide Variation(s)
TCA	Tricyclic antidepressants
TE Buffer	10 mM Tris-HCl containing 1 mM EDTA
VNTR	Variable number tandem repeat
WHO	World Health Organization
wt_ NnN	Weight Not Normal (Overweight and Obese Class 1-3)

LIST OF GENE ABBREVIATIONS

5-HT	5-Hydroxytroptophan
5-HT2C	5-Hydroxytroptophan subtype 2C
5-HTTLPR	Serotonin-transporter-linked polymorphic region
ADRβ2	Beta-2 adrenergic receptor (β2 adrenoreceptor)
ADRβ3	Beta-3 adrenergic receptor (β3 adrenoreceptor)
ANKK1	Ankyrin repeat and kinase domain containing 1
AGRP	Agouti-related protein
BDNF	Brain-derived neurotrophic factor
COMT	Catechol-O-methyltransferase
CYP2D6	Cytochrome P450 family 2 subfamily D member 6
DRD2	Dopamine receptor subtype D2
DRD4	Dopamine receptor subtype D4
GABA	Gamma-aminobutyric acid
GABRA6	GABAergic gene
LEPR	Leptin receptor
MAO-I	Monoamine oxidase inhibitors
MC4R	Melanocortin-4 receptor
MLOGIT	Multinomial Logical Regression
OPRM1	Opioid receptor 1
PPAR-γ / PPARG	Peroxisome proliferator-activated receptor gamma
PROP	Propylthiouracil (6-n-Propylthiouracil)
SCL6A4	Solute carrier family 6 member 4
TAS2R38	Bitter receptor gene

CHAPTER 1: BACKGROUND AND LITERATURE REVIEW

1. LITERATURE OVERVIEW AND MOTIVATION

1.1. The Epidemic of Obesity and Depression

South Africa has approximately 14 million people retiring to bed hungry each day. However, on the other end of the scale South Africa has also been labelled the fattest nation in sub-Saharan Africa, with one in three persons being classified as obese, according to a press release in The South African and the statistics released by the Lancet[1, 2].

The World Health Organization (WHO) defines overweight and obesity as the abnormal or excessive fat accumulation that may impair health[3]. Obesity is, among other reasons, the result of overeating and associated eating disorders. It is an ever-increasing worldwide epidemic and a major public health concern. Approximately 52% of global deaths are caused by non-communicable diseases (NCD), many of which are associated with the co-morbidities of obesity[4]. Cardiovascular disease accounted for 37% of deaths, followed by cancers (27%), respiratory diseases (8%) and diabetes mellitus (4%)[4]. Figures 1 and 2 show the proportion of deaths caused by NCD and their associated medical conditions which have a link to obesity. The Body Mass Index (BMI) is considered the universal measure of defining a body mass as healthy or unhealthy. BMI is calculated by taking the weight in kilograms divided by the square of height in meters. A normal weight is classified as a BMI between 18.5 kg/m^2 and 24.9 kg/m^2, while a person is classified as overweight if their BMI is between 25.0 kg/m^2 and 29.9 kg/m^2. There are three classes of obesity: class 1 (low risk) with a BMI between 30.0 kg/m^2 and 34.9 kg/m^2, class 2 (moderate risk) BMI between 35.0 kg/m^2 and 39.9 kg/m^2 and class 3 (high risk) BMI greater than 40.0 kg/m^2 (Table 1). Individuals with a BMI greater than 40.0 kg/m^2, are considered morbidly obese[3, 5].

Table 1: World Health Organization Adult Body Mass Index values & Classification

Classification	BMI Range
Underweight	< 18.5 kg/m^2
Normal weight	18.5 kg/m^2 - 24.9 kg/m^2
Overweight	25.0 kg/m^2 - 29.9 kg/m^2
Obese Class I	30.0 kg/m$^{2\,-}$ 34.9 kg/m^2
Obese Class II	35.0 kg/m^2 - 39.9 kg/m^2
Obese Class III	> 40.0 kg/m^2

Associated medical conditions of concern in obese individuals include; insulin resistance, type 2 diabetes mellitus, hypertension, cardiovascular disease, stroke, sleep apnoea, asthma, certain types of cancer and physical disability[6, 7]. Figure 2 indicates medical related complications associated with obesity[8]. According to the Heart Foundation's statistics released in 2016, 70% of women in South Africa are overweight or obese, and approximately one third of men in South Africa are overweight or obese. This is in agreement with statistics released by the South African Medical Research Council, stating 61% of South Africans being overweight or obese[9, 10]. Obesity is a complex condition, with both a social and psychological dimension that affects all ages and socio-economic groups, in both developed and developing countries[11].

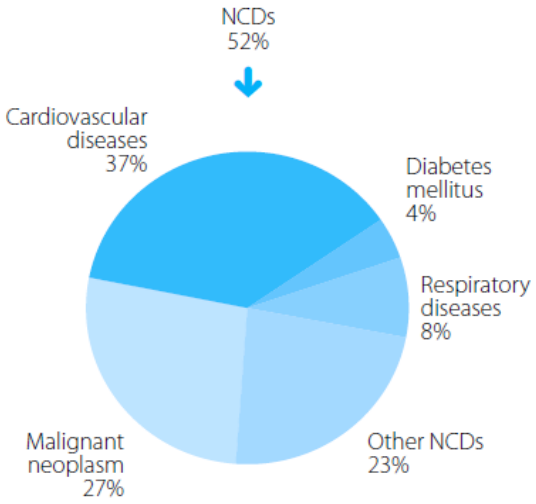


Figure 1: Proportion of non-communicable disease (NCD) related deaths that occurred among people under the age of 70 years, Global Status Report on noncommunicable diseases, WHO, 2014[4].

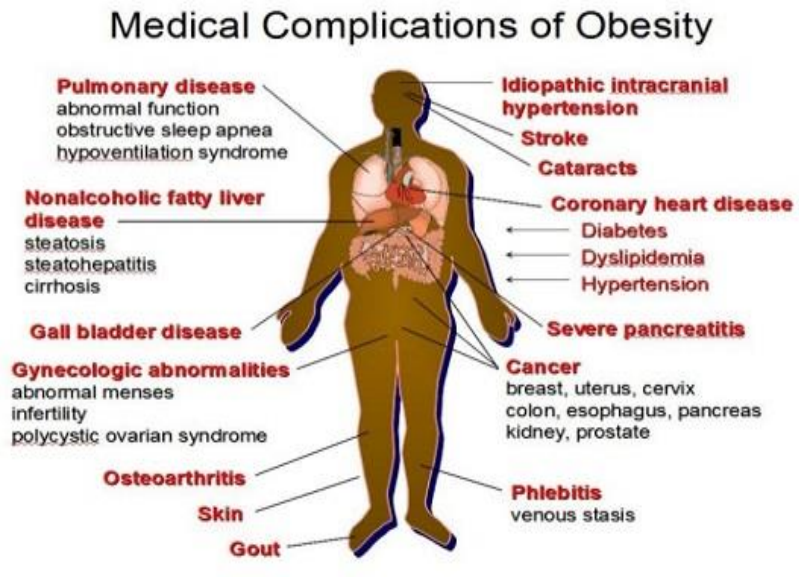


Figure 2: Medical complications associated with obesity, MHA News 2014[8].

Mental health disorders are also a major contributor to the local and global disease burden[12]. Mental disorders are linked to various health conditions, including overweight and obesity. In addition, mental disorders are a significant contributor to the healthcare expenditure, as treatment is often costly. Approximately 350 million people worldwide are affected by depression, according to the statistics released by the WHO[13, 14]. The Diagnostic and Statistical Manual of Mental Disorders (DSM) is used by Health Care

Professionals worldwide, as a guide to diagnose mental disorders. The DSM contains descriptions, symptoms and other criteria for diagnosing mental illnesses. Major Depressive Disorder (MDD) is classified according to the DSM-5 criteria, which states at least three of the manic/hypomanic symptoms need to be present during the majority of days of a major depressive episode[15, 16].

Major depression and obesity share co-morbidities, such as cardiovascular disease[17]. Symptoms of depression may include altered appetite, weight gain and reduced physical activity whereby increasing the risk of obesity[17]. This phenomenon illustrates the close relationship between these two distinct diseases. Both conditions typically show altered eating habits and cravings for certain foods containing high amounts of carbohydrates and trans-fatty acids (unsaturated fatty acids). The link between depression and obesity is bidirectional: depressed individuals may become obese when they resort to food as a source of comfort, while obese people may become depressed due to feelings of hopelessness. Many obese individuals identify specific food groups that foster continued consumption, whereby there is a lack of willpower and an inability to control said urges[18]. The effects of obesity leading to depression, may leave individuals experiencing feelings of discomfort, dissatisfaction, worthlessness and guilt[19].

Obesity has also been shown to have a genetic component linked to it[20-22]. Genes and environmental factors interact with each other to regulate body weight. The overall heritability of obesity has been estimated to be between 40% and 70%[23]. By understanding the genetic and epigenetic involvement in the phenotype of obesity, better intervention strategies can be implemented, to improve therapies for obesity and preventing future occurrences through early detection.

1.2. Genetics of Obesity

Research has shown that overweight and the different forms of obesity tended to be centred within a family. Individuals with a family history of obesity, have an increased risk of becoming obese, two to eight times higher than an individual with no family history[7]. This risk is even higher in cases of severe obesity. Heritability of obesity appears to be phenotype dependant. Phenotypes associated with adipose tissue distribution, visceral fat (abdominal fat located inside the abdominal cavity) and excess

visceral fat (also known as “belly fat”) in which the abdomen protrudes excessively, show a 40-50% and 5-41% higher heritability risk, respectively[7]. The genotype of an individual refers to the pair of genes inherited for a particular trait. The phenotype is the physical observable presentation of the trait[24].

Obesity is a complex, multifactorial disorder resulting from the interactions of genes, environment and lifestyle[11, 25]. It has long been known that there is a genetic component to obesity, but this is more than simply a person having inherited the “fat” gene. Genome wide associated studies (GWAS) are observational studies conducted on a genome wide scale to identify sequence variations that can be linked to genetic risk factors or traits that are common to a particular population[26]. Genome-wide association studies have revealed a variety of genetic loci, in the form of Single Nucleotide Variations (SNVs), have been linked to obesity. SNVs is the modern terms for genetic variations and are single base-pair changes of the DNA sequence in the human genome. SNVs are used as markers in genomic regions, with the majority having a minimal impact. Some variations are silent, meaning they do not affect the amino acid sequence of the gene product, or they do not occur in coding regions of the DNA. Others, result in amino acid changes which can change the protein structure and function[27], Figure 3. Our bodies are a network; therefore, one gene can have more than one function and one disease can be affected by more than one gene.

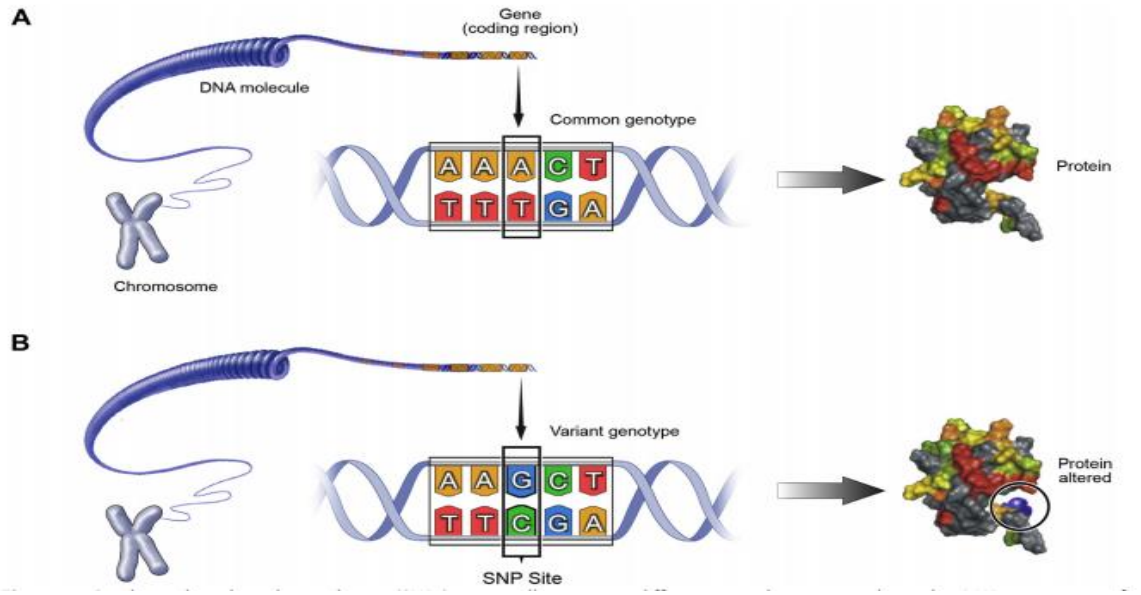


Figure 3: Single nucleotide variation that alter the protein structure, Camp & Trujillo 2014[27].

Single gene mutations have only been shown to be the cause of obesity in 1 to 5% of cases[25]. The complex nature of obesity suggests the involvement of multiple genes and gene variants[25]. In order to better understand the effects of environmental and lifestyle changes on the development of obesity, the need to investigate specific genes linked to obesity is of utmost importance. Genes associated with food intake and energy that influence factors such as hunger, appetite, eating behaviours, taste, satiety, basal metabolic rate and exercise motivation, are important to investigate. Currently, more than 400 genes and gene markers have been linked to obesity. Previous evidence of the role of genetics in weight regulation has come from numerous family studies. Stunkard *et. al.*[28] revealed that identical twins' BMI scores are highly correlated, even if they have been reared in different environments. Another study involving young Swedish male twins, showed that similar eating behaviours observed in fraternal and identical twins were due to genetic factors[29]. The retrospective cohort study from Whitaker *et. al.*[30] and a prospective observation study by Maffeis *et. al.*[31] showed that parental obesity was a strong indicator of childhood to adult obesity and more than doubled the child's risk irrespective of diet or physical activity. A study investigating the food and activity preferences in children between the ages of four and five years, of lean and obese parents, found that children of obese parents had a greater preference for fatty foods and a lesser liking of vegetables. In addition, these children demonstrated an increased tendency to overeating and decreased physical activity compared to children of normal weight parents[32].

One of the biggest current uncertainties pertaining to obesity research, is whether overeating and food craving is an addictive behaviour and shares a similar relationship to other addictive behaviours such as drug and alcohol abuse, obsession and pleasure seeking[11]. Some individuals may have an obsession with a particular food, given the way the food makes them feel, thereby seeking pleasure as a function of satiety and dysregulation of craving[11]. Although for many individuals this may seem like simply a bad habit, there is increasing evidence suggesting that genes implicated in other addictive behaviours are also associated with food cravings towards sweet and fatty foods[18]. Examples of these genes and their functions has previously been described and are summarized in Table 2[33-35].

Table 2: Examples of genes involved in obesity and their associated phenotypes.

Gene	Associated Phenotype (Characteristic), e.g. Function
Leptin	Satiation, metabolism
Melanocortin	Feeding behaviour, binge eating
Ghrelin	Appetite stimulation
Neutromedian β	Feeding behaviour
Bitter Receptor Gene (*TAS2R38*), 6-n-Propylthiouracil (*PROP*)	Taste preference - bitterness
Peroxisome proliferator-activated receptor (*PPAR*)	Fat metabolism
Mitochondrial upcoupling proteins *(UCP)*	Energy expenditure
Melanocortin and Melanocortin-4 receptor (*MC4R*)	Energy expenditure

Kenneth Blum and his team at the Department of Psychiatry, McKnight Brain Institute, University of Florida, USA, coined the term Reward Deficiency Syndrome (RDS). They defined RDS as the genetic and epigenetic phenomena leading to impairment of the brain reward circuitry resulting in a hypo-dopaminergic function[36]. RDS involves the interaction of powerful neurotransmitters and results in abnormal craving behaviours, which have been shown to be a predictor of impulsive and addictive behaviours[37, 38].

One of the major contributing factors to the obesity epidemic is the ever changing environment and lifestyle, which promotes high calorie intake in the form of quick meals and avert ion of physical activity[7]. Given the economic burden on South Africans, most households are experiencing increased financial stress and strain[39]. People are required to work harder than in the past to maintain their current lifestyles, thereby necessitating the need to manage increased stress. In doing so, individuals resort to various coping mechanisms to deal with their ever-increasing stress levels, which often includes overeating[40]. Stress has been identified as one of the leading causes of depression[41]. Genes also play a role in the way individuals handle and manage stress, which has an effect on cortisol the levels in the brain. Environmental changes such as lifestyle modifications in combination with genetic susceptibility, could likely be responsible for the increase in obesity prevalence in South Africa[40].

1.3. The Evolution of Obesity Theory

The theory of evolution and survival of the fittest is widely accepted. The survival of the Pepper Moth during the Industrial Revolution is the classroom example of how traits have altered to adapt to the environment. The pepper moth exhibited a directional colour change from being light-coloured to becoming dark-coloured as a consequence of air pollution during the Industrial Revolution. The frequency of dark-coloured moths increased during that time is an example of industrial melanism. Following a reduction in pollution, the light-coloured phenotype again predominated. Industrial melanism in the peppered moth was an early test of Charles Darwin's natural selection in action[42].

When food was scarce and the risk for famine and starvation high, vigorous physical activity in humans was required for gathering food and to fend off any potential threats. Over time, the genome adapted to the limited availability of food, by enabling phenotypic changes to sustain biological functions with efficiency, and to store excess energy in adipose tissue and triglycerides in non-adipose tissue. For example Beta-2 Adrenergic receptor (*ADRβ2*) also known as the “Thrifty gene” is responsible for the slow breakdown of fat and low energy expenditure[43].

The genes of modern humans, derived from the gene pool of their ancestors that had survival advantages during times of famine, have changed minimally over the last 200 years since the Industrial Revolution. However, our lifestyles have transformed rapidly over the past 10 000 years beyond recognition and measure, surpassing any possibility of genetic adaption. Our current state of conflict has little effect on our reproductive success, but rather acts as a promoter of chronic illnesses, such as an individual’s predisposition to obesity. Therefore, individuals have the genome of a hunter-gatherer, however, most live in a society where food is in abundance and easily accessible. It is this discordance between our genome (from our ancestors’ biology) and the current lifestyle that leads to an energy imbalance and eventually obesity[44]. Figure 4 illustrates the silhouette of the evolution of obesity. Not only are our genotypes programmed to store energy, our lifestyles do not allow for a sufficient amount of physical activity to metabolize the stored excess energy, especially when energy intake is high[45].

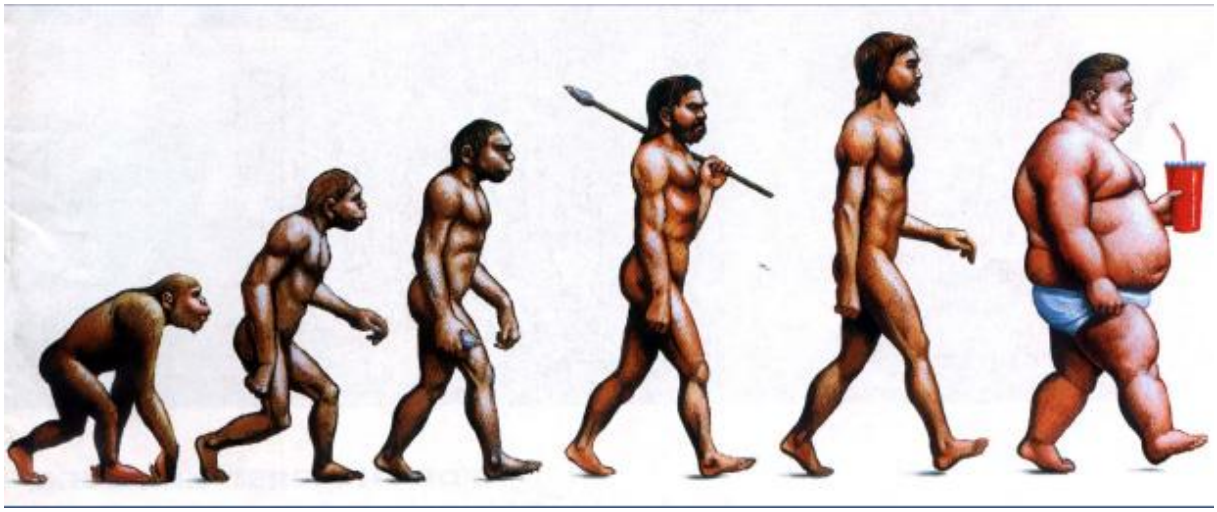


Figure 4: The shape of things to come: the theory behind the evolution of obesity, World of DTC Marketing 2015[45].

1.4. The Basic *basics* of Genes

Genes are the units of heredity and are the instructions that make up the body's blueprint. Genes code for the proteins that determine the entirety of a person's characteristics. The Human Genome Project has estimated that humans have approximately 20000 to 25000 genes[46]. Genetic disorders are caused by one or more changes, or mutations, in the instruction code of a particular gene(s), preventing the gene(s) from functioning properly. The physical location of a gene is its locus. Different versions of genes are called alleles. Point mutations alter the genetic code by changing the codons, these changes can alter the protein. Figure 5, illustrates the effect of a nucleotide change that on the expression of the Brain derived neurotrophic factor (BDNF) gene, and the effect environmental influences have on the observed phenotype[47].

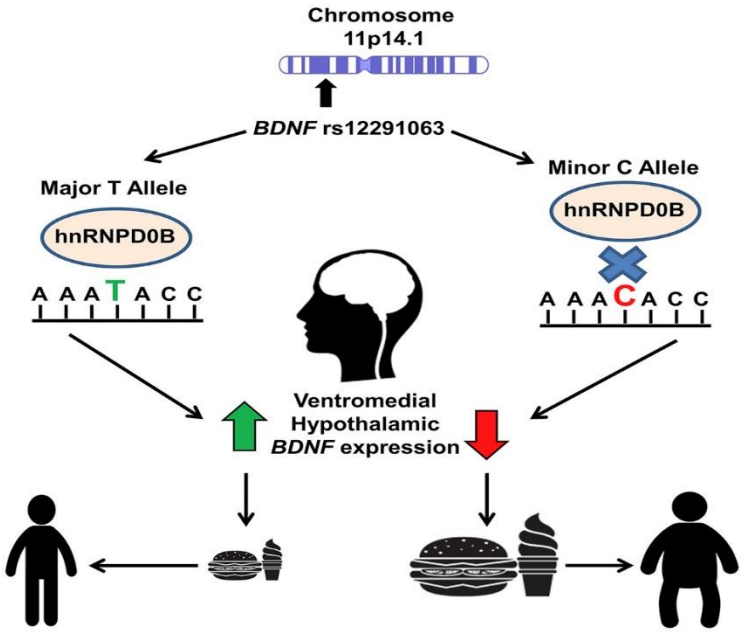


Figure 5: Single nucleotide variation in the *BDNF* genes (rs12291063), affecting the expression of BDNF and risk of obesity, Mou *et al*. 2015[47].

1.5. Epigenetics

Epigenetics is the study of potentially heritable changes in gene expression or cellular phenotypes, caused by mechanisms other than changes in the underlying Deoxyribonucleic Acid (DNA) sequence. This translates to a change in phenotype without a resulting change in genotype[48]. Epigenetics are therefore defined as "in addition to changes in the genetic sequence". This term has since expanded to include any process that alter gene expression without changing the DNA sequence.

Factors that influence epigenetic changes include; age, lifestyle and disease state. These changes include DNA methylation, acetylation, phosphorylation, ubiquitylation and histone modification. All these processes regulate gene expression without altering the DNA sequence, e.g. methylation adds a methyl group to cytosine, without changing the base. Epigenetic processes are naturally occurring and essential to the functioning of the organism, however, when these events occur unregulated, it may result in adverse health and behavioural effects[49].

1.5.1. Interactions

The relationship between gene variants; genes and age; and genes and the environment are complex. Below is a summary of some of the focus areas conducted to determine these relationships and how they relate to the obesity epidemic.

1.5.1.1. Gene-Gene Interactions

A case-controlled study revealed a strong association between two gene variants in the risk for obesity in children and adolescents, namely Beta-3 Adrenergic Receptor (*ADRβ3*) and Peroxisome Proliferator-Activated Receptor Gamma (*PPAR-γ* or *PPARG*). The study found that carriers of both gene variants were almost 20 times more likely to become obese than non-carriers (odds ratio [OR] 19.5: 95% confidence interval [CI], 2.4-146.8), showing the synergistic contribution of the two polymorphisms[50].

1.5.1.2. Gene-Age Interactions

The effect of an individual's age on gene expression was shown in the study by Argyropoulos *et. al*.[51]. The study revealed an association between a gene variant for Agouti-Related protein (*AGRP*, a powerful appetite effector) and obesity in older adults and their biological children. The gene variant was strongly associated with fat and abdominal adiposity in the parental group with a mean age of 53 years, but not in the study group or "children's group" with a mean age of 25 years.

1.5.1.3. Gene-Environmental-Lifestyle Interactions

Genetic factors only partially determine an individual's susceptibility to obesity. Lifestyle and environmental factors play an additional role in the phenotypic expression. For example, the severity of obesity is determined by lifestyle and environment conditions that an individual with the genetic predisposition of obesity is exposed to. When an individual move from a restricted environment or non-obesogenic environment to an obesogenic environment, they are likely to gain weight. However, individuals with a high genetic risk of obesity will gain more weight than those resistant to obesity[7]. Figure 6 illustrates the effects of an obesogenic environment on weight gain. Individuals with a high predisposition (left) to obesity will gain more weight and

have a higher BMI than an individual with a lower predisposition to obesity (right)[52]. The term “obesogenic environment” is defined as an environment that promotes weight gain and is not conductive to weight loss, contributing towards obesity.

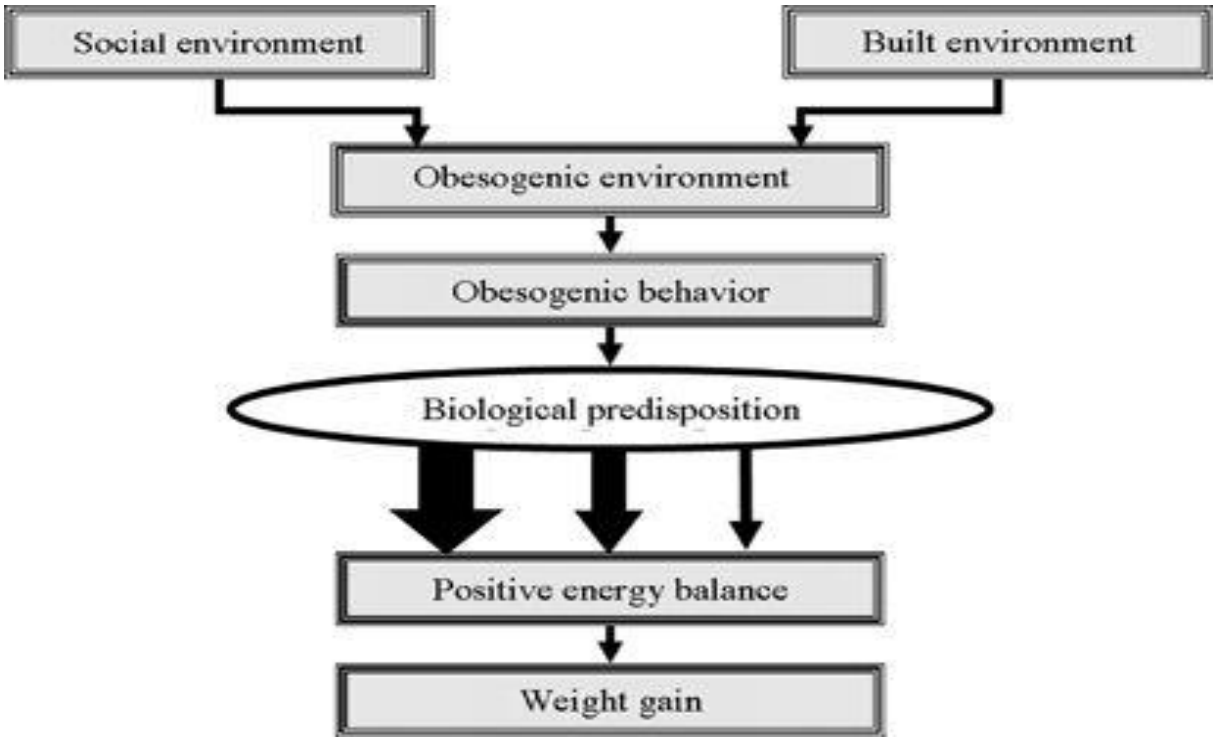


Figure 6: The effects of an obesogenic environment on weight gain, which is dependent upon an individual’s genetic predisposition to obesity, Bouchard 2007[52].

Several studies have investigated the gene-environment-lifestyle interactions; however, it was Martinez *et. al*.[53] that established the relationship between diet and specific genes that may affect the obesity risk. Their research found that women with the Gln27 variant (glutamine at amino acid 27) in the beta-2 adrenoceptor (*ADRβ2*) gene were at an increased risk of obesity, when following a diet with more than 49% of the calories derived from carbohydrates, compared to non-carriers.

Another example is observed in the Pima Indian population which are predisposed to obesity and type 2 diabetes. Pima Indians living in a non-obesogenic environment in a remote region of Mexico (Sierra Madre Mountains), have a lower prevalence of obesity and type 2 diabetes than those living in Arizona (south western United States) in an obesogenic environment[54].

1.5.1.4. Genes, Obesity and Metabolic Disorders

The manner in which our bodies store fat may be an important indicator of metabolic disorders associated with obesity, as opposed to the total body fat. A study by Cassano *et. al*.[55] showed that individuals with fat concentrated around their mid region were at a greater risk of developing diabetes compared to individuals with the same amount of total fat distributed throughout their bodies. The involvement of genetics in the abdominal fat deposition seen in postmenopausal women has also been identified[56].

1.6. Obesity and the relationship with addiction

Obesity has always been perceived as a behavioural disorder in society, with overweight individuals lacking the willpower and self-control when managing their food intake. Several similarities have been identified between obesity and addictive disorders, including co-morbidities, personality and behavioural traits and brain/ biological mechanisms[57].

The term addiction is generally applied to the excessive ingestion of substances or chemical entities which leads to a dependency characterised by tolerance and withdrawal symptoms[58]. Compulsive behaviours such as gambling, sex and eating were previously not considered to be true addictions because of the psychologically motivated component. The conceptual model of addiction is changing, where the chemical properties of the substances themselves is not the key element but rather the behaviour of substance, which causes a physiological change in the brain. Addiction is therefore currently seen as a syndrome that may express in a variety of specific behaviours, which include overeating[59]. Despite knowledge of the effects of overeating and the associated negative health implications, individuals continue to overeat. It has been proposed that obesity (or the cause of overeating) be included in The Diagnostic and Statistical Manual of Mental Disorders (DSM-V), due to the similarities the disease shares with addiction (Table 3 and Table 4). Binge eating disorder diagnosis is currently included in the DSM-IV, which is defined as the loss of control over food intake and consumption of large quantities of food over a short period of time[57].

Table 3: DSM-IV criteria for substance dependence diagnosis and parallel criteria for a possible disorder or overeating.

Substrate Dependency Criterion	Parallel Criterion for "Overeating Disorder"
1. Tolerance, including need for more of the substance to achieve the same effect or a diminished effect when using the same amount of the substance over time.	1. Physiological tolerance unlikely, but some individuals feel need for increased quantities for food in order to feel satisfied.
Example: Alcohol dependent individual does not feel intoxicated after consuming entire 6-pack in an evening.	Examples: Overweight or obese individual feels hungry after a large meal.
2. Withdrawal, including characteristic syndrome of withdrawal symptoms for specific substance or use of the substance or a similar one to relieve or prevent those symptoms.	2. Comparable withdrawal syndrome not yet identified, but dieters and other individuals deprived of food report psychological preoccupations with foods and some individual use substance such as nicotine or stimulants to suppress appetite.
Example: Heroin dependent individuals experiences dysphoria, nausea, sweating and insomnia when she can't obtain heroin, takes Oxycontin (oxycodone) to compensate.	Example: Dieter feels lethargic and depressed, smokes or drinks caffeinated beverages to compensate.
3. Individuals frequently takes more of a substance than intended or takes it over a longer period of time than planned.	3. Food is often consumed in larger amounts or over a longer time than was intended.
Example: Alcoholic plans to stop at the local bar for, one beer, ends up staying until closing having several drinks.	Example: Dieter plans to have one small serving of ice cream, but ends up having eating an entire pint.
4. Repeated unsuccessful efforts to reduce substance use or persistent desire to do so.	4. Obese individuals who overeat often have a persistent wish to reduce or control how much they eat or try repeatedly to eat less.
Examples: Cocaine dependent individual repeatedly vows to stop using at the start of the day, but ends up using by the end of the day.	Example: Repeated, unsuccessful diets or regaining weight after successful diet are the norm for most obese individuals.
5. Substantial amount of time spent obtaining, using or recovering from use of substance.	5. Overeaters can spend substantial time shopping for food, eating and snacking and recovering from physical and psychological effects of overeating (e.g., nausea, guilt about eating too much).
Examples: Cannabis dependent individual spends hours calling his various contact to locate available marijuana, travels 2 hours to get it, then smokes for most of the weekend.	Example: Obese individual snacks throughout the day in addition to or instead of eating regular meals.
6. Individual abandons or cut back on social activities, work or family responsibilities and recreational interests in order to use substances.	6. A range of activities may be abandoned or reduced because of consequences of overeating (i.e., obesity) and accompanying decreased mobility, increased social anxiety, etc.
Example: Drug user stops associating with non-drug using friends.	Example: Obese individual stops participating in sports or going to the beach because of embarrassment about weight.
7. Substance use continues in spite of associated physical and psychological problems.	7. Overeating continues in spite of associated physical and psychological problems.
Example: Alcohol dependent individuals continue to drink after being diagnosed with hypertensive and gastric ulcers.	Example: Obese individuals continues to eat candy after being diagnosed with type II diabetes mellitus.

Table adapted from Barry, Clarke & Petry 2009[57].

Although the treatment of obesity is theoretically uncomplicated (reduce food intake and increase physical activity), few people achieve significant weight loss and even fewer manage to maintain their weight loss. This suggests that the drive to consume food is beyond what is necessary for survival. Current studies show that foods rich in fat and sugar result in the brain's reward system to be overstimulated, thereby antagonizing the appetite-suppressing hormones (such as leptin) and the brain's ability to ensure negative feedback to stop eating[59].

Table 4: Characteristics common to individuals with overweight/ obesity and substance use disorder.

Personality Characteristics	
1.	Elevated scores on novelty-seeking scale of the Temperament and Character Inventory (TCI)
2.	Low scores on the self-directedness scale of the TCI
3.	Higher scores on self-report measures of impulsivity
4.	Poorer scores on the Iowa Gambling Task
5.	Preference for smaller immediate vs larger delayed rewards on Delayed Discounting Task
Disruptive Behaviour Disorders	
1.	Higher rates of Attention Deficit Hyperactivity Disorder
2.	Higher rates of Conduct Disorder
3.	Deficits on tests of executive functions
Brain Mechanisms	
1.	Overeating and substance use stimulate mesocorticolimbic dopamine system activity
2.	Number of D2 dopamine receptors decreases from normal levels in brain of obese individuals and chronic substance users, suggesting downregulation of receptors with chronic stimulation of dopamine system

Table adapted from Barry, Clarke & Petry 2009[57].

Obesity is treated as a medical disease, however, the possibility that obesity in a sub set of individuals could be a psychiatric disorder, needs further investigation. It is as difficult for an obese person to limit their consumption, as it is for an individual with an alcohol or substance dependence to limit consumption of these substances. Identifying an addictive model for overeating may effectively form part of the prevention and treatment for obesity[57].

1.7. The Brain Reward Cascade

The brain reward system incorporates a group of neural structures responsible for incentive salience, associative learning and positive emotions. Neurotransmitters are chemicals that allow the transmission of signals from one neuron to the next through the synapses. A rewarding stimulus is defined as any stimulus, object, event, activity or situation that has the potential to trigger the need for an individual to seek the effects produced by these stimuli. Rewarding stimuli therefore function as positive reinforcers[60].

The mesolimbic system, also referred to as the reward centre, is the anatomical area of the brain where the feelings of pleasure and wellbeing are interpreted and experienced. The interrelationship of four important neurochemical messengers: serotonin, enkephalin, gamma-aminobutyric acid (GABA) and dopamine, provide for the release of dopamine into the Nucleus Accumbens (NAc). These messengers control the synthesis, vesicular storage, metabolism, receptor formation and neurochemical function essential for the feeling of wellbeing (Figure 7). The cascade starts with the stimulation of the serotonergic system in the hypothalamus. Serotonin stimulates the delta/mu receptors to release enkephalin. Enkelphalin stimulates the mu receptors, resulting in activation of the enkelphalinergic system at the GABA neurons, causing inhibition of the GABA transmission at the substania nigra, providing the normal release of dopamine in the NAc[61].

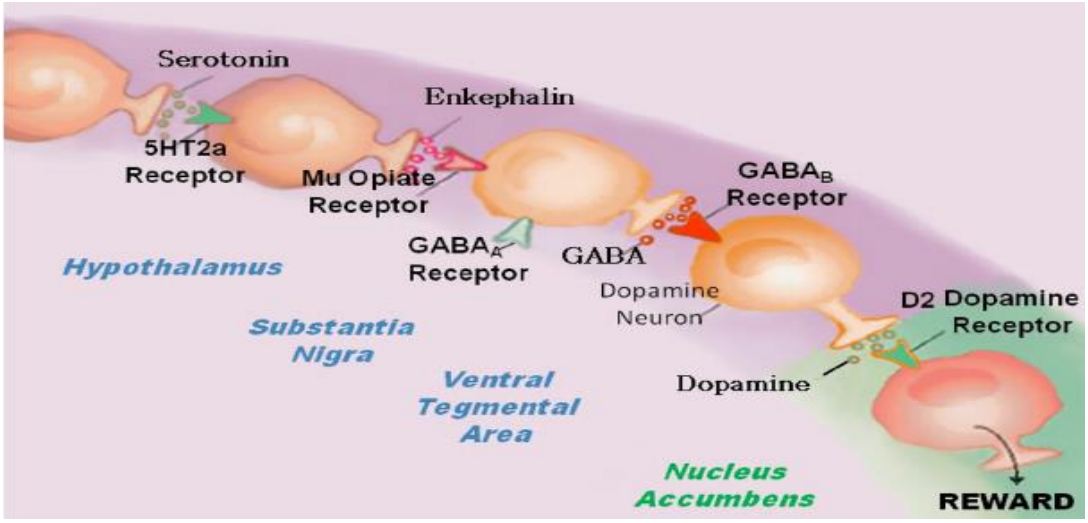


Figure 7: The Brain Reward Cascade, Blum *et al* 2014[61]**.**

1.7.1. Dopamine

Dopamine is one of the neurotransmitters involved in the reward cascade. It functions in three important pathways: the mesocortical pathway (involved in memory, motivation, emotion, reward, desire and addiction), nigrostriatal pathway (involved in motor control) and tuberoinfundibular pathway (involved in hormone regulation, nurturing behaviour and sensory processes). Dopamine controls the feelings of wellbeing as a result of its interaction with other neurotransmitters such as serotonin and opioids[38]. Dopamine is an unique neutrotransmitter as it displays both excitatory and an inhibitory characteristics, depending on the type of receptor present. When dopamine binds to its receptor, it either prevents the neuron from firing, or it may promote the firing of the neuron leading to a physiological effect. Dopamine is released into the synpase which stimulates a number of receptors and results in an increased feeling of wellbeing and a reduction in stress. Drugs such as cocaine, opium, heroin and alcohol increase the levels of dopamine and produce a similar response[18]. Mental diseases such as schizophrenia has been associated with excessive amounts of dopamine in the frontal lobe, while Parkinson's disease has been associated with reduced levels of dopamine in the motor areas of the brain[22]. Dysregulation or dysfunction of dopamine results in the breakdown of the brain reward system. Any reduction in the function of dopamine could lead to reward deficiency (Figure 8)[37]. Variations that occur within the genes that govern the function of the brain reward cascade, result in dysregulation in the mesolimbic system[61].

1.7.2 Serotonin

Serotonin, also known as 5-hydroxytryptophan (5-HT), is another neurotransmitter present in the reward cascade pathway. It is an inhibitory neurotransmitter with no stimulating effect on the brain. Serotonin is necessary for a stable mood and emotional state by suppressing excessive excitatory neurotransmitter release and firing from the almost 86 billion cells in the human brain. Serotonin plays a role in various other brain processes, including body temperature regulation, sleep cycle, appetite and pain control. Low levels of serotonin in the brain have been linked to obsessive-compulsive disorder, anxiety disorders, anger control, suicide, increased carbohydrate cravings and depression. Most pharmacological agents used to treat depression act by

increasing serotonin levels in the brain. A well known drug, fluoxetine, is an example of a selective serotonin re-uptake inhibitor (SSRI), whereby the serotonin concentration in the synpatic cleft is increased as a result of the re-uptake inhibition into the pre-synaptic neuron[61]. Serotonin's link to carbohydrate cravings, forms the basis for agents such as lorcaerin (a selective *5-HT2C* agonist), that have been introduced to aid in weight reduction by accelerating the onset of satiety[62].

1.7.3. Glutamate and Gamma-Aminobutyric Acid (GABA)

Glutamate and GABA are the central nervous system's most abundant neurotransmitters. GABA is inhibitory and glutamate is excitatory. These neurotransmitters act through positive and negative feedback mechanisms to control multiple processes of the brain, including the overall level of excitation. GABA inhibits the excitatory neurotransmitters which are responsible for anxiety. Individuals with low GABA levels are known to have anxiety disorders and insomnia. Alcohol and barbituates are both indirect agonists on GABA receptors, thereby inhibiting the neuronal signaling and the excitatory effect of glutamate. Individuals with addictions to alcohol, drugs, tobacco, caffeine and food tend to lack GABA[61].

1.7.4. Endorphins

Endorphins (endogenous morphine) are a category of neurotransmitters that the body uses as internal analgesics due to their inhibitory effect on pain signal transmission, increased ability to experience pleasure and the overall feeling of happiness. They are structually and functionally similar to opioids. Endorphin levels are greatly elevated after exercise[61].

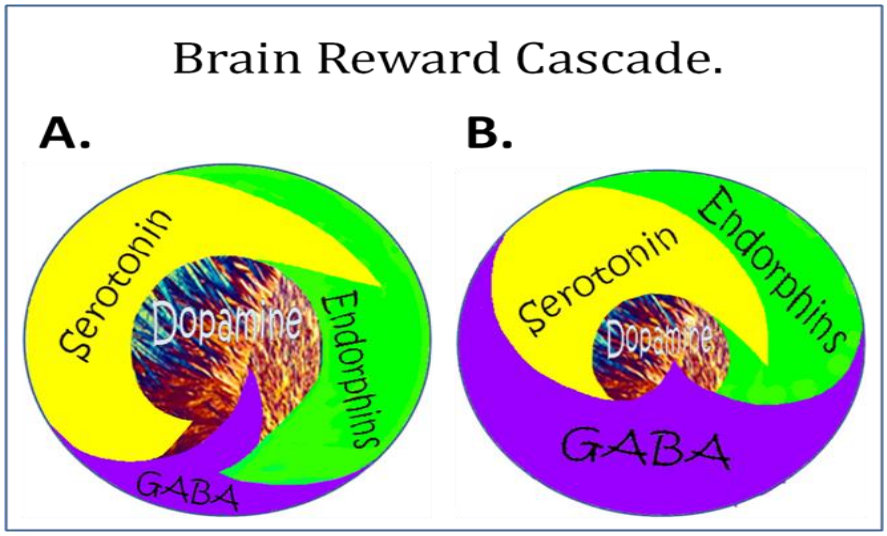


Figure 8: Neurotransmitters of the Brain Reward Cascade indicating the four major neurotransmitters (serotonin, dopamine, endorphins and GABA) in the mesolimbic region. a) Normal dopaminergic state, b) Hypodopaminergic state is the dysregulation or dysfunction of dopamine and results in the breakdown of this system, Blum *et al*. 2015[37].

1.8. Obesity and the Reward Cascade

Several studies have shown that overweight and obese individuals tend to behave differently to food stimuli and reward than that of normal weight individuals. Response to food related stimuli is affected by the dopamine and leptin signaling pathway[11]. As shown in Figure 8b, hypodopaminergic function may be the result of gene variations or environmental elements causing dysregulation or reduced dopamine receptor subtype D2 (*DRD2*) densities. The *DRD2* gene is referred to as the reward gene[11, 63-65]. In the presence of low *DRD2* receptor levels, individuals seek substances or behaviour that stimulate the dopaminergic system, including a desire for food. Studies indicate that obese individuals have a heightened sensitivity to the satisfaction provided by food. Once food is consumed, they experience a reduced reward, which results in overeating to compensate for the unfavourable reaction.

Leptin is a hormone produced by adipose cells and is partly responsible for the regulation of energy balance by inhibiting the feeling of hunger. It is also known as the “satiety hormone”. In obese individuals, there is a decreased sensitivity to leptin, resulting in an inability to detect satiety despite high energy stores[66].

Leptin regulates satiety by decreasing hypothalamic brain activation in response to food. Plasma leptin levels increase once food is consumed. In leptin deficient individuals, hypothalamic brain activation to food cues remain present even after eating. Such an individual hardly experiences the feeling of satiety, and therefore often overeats as a result. Figure 9 illustrates the leptin feedback loop[63-65].

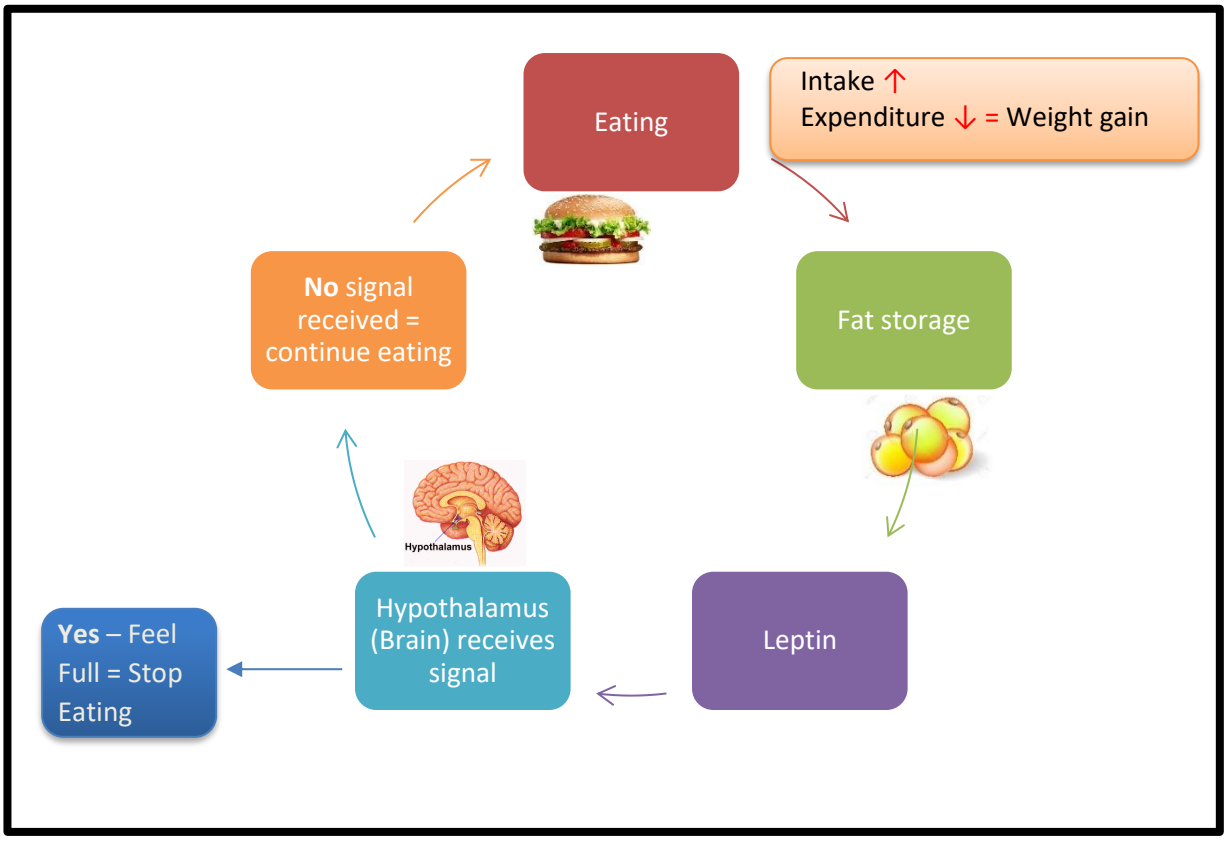


Figure 9: Leptin and satiety feedback loop.

Food consumption leads to the storage of fats in the adipose tissue. With an increase in adiposity there is an increase in the level of leptin released into the blood. Once leptin is released from the adipose tissue, it is transported in the plasma to the brain. When the brain receives the leptin signal, the hypothalamus responds with a negative feedback signal of satiety, which physiologically instructs an individual to stop eating, thereby maintaining constant stores of fat. When the brain does not receive the leptin signal (likely due to a genetic defect), the feeling of satiety is not initiated and an individual continues to eat. This leads to an increase in the uptake of fats, a decrease in the expenditure of energy and results in weight gain[67].

1.8.1 Drugs used in the treatment of Obesity

Drugs which increase the amount of serotonin and dopamine present within synapses could potentially cause weight reduction, due to a knock-on effect. These drugs act by displacing neurotransmitters from their storage vesicles (e.g. amphetamine), stimulate the release and inhibiting reuptake into presynaptic vesicles (e.g. phentermine, d-norpseudoephedrine, diethylpropion), or reduce appetite by activating the serotonin 5-hydroxytryotamine subtype 2C (*5-HT2C*) receptor (e.g. lorcaserin)[62]. These serotoninergic drugs act by accelerating the onset of satiety, enhancing basal metabolic rate by 100 calories per day and inhibiting carbohydrate cravings. Carbohydrate craving has been exhibited by people who are overweight, or are becoming so. This phenomenon is postulated to be interpreted by the brain as a serotonin deficiency and administering a serotoninergic drug might constitiute a specific therapy for the etiological process responsible for causing obesity. An association between mood disturbance, the inability to lose or to stop gaining weight and cravings for carbohydrates has been obeserved in a number of syndromes such as weight gain after exposure to stress, nicotine withdrawal, or premenstrual syndrome. Carbohydrate rich foods act via insulin secretion to increase the plasma tryptophan ratio, thereby accelerating the production and release of serotonin. Protein intake lacks this effect. The resulting increase in brain serotonin levels also ameliorates mood disturbances, sleep onset, pain sensitivity and blood pressure regulation. Eating foods rich in carbohydrates therefore results in a psycho-pharmacological effect whereby patients feel better after consuming these carbohydrate and fatty meals or snacks, leading to weight gain[62].

1.8.1.1 Appetite Suppressants

Appetite suppressants are β-phenylamine by-products and are structurally similar to noradrenaline, dopamine and amphetamine. Appetite suppresants increase the levels of noradrenaline and dopamine in the synapse in a three phase process; Firstly, it displaces the neurotransmitters from the storage vesicles, secondly it stimulates the release and prevents the re-uptake into the presynaptic vesicles (phentermine, diethylpropion, phendimetrazine, d-norpseudoephedrine and sibutramine), and thirdly it agonises adrenergic receptors (phenylpropanolamine). Early stimulation of the hypothalamic satiety centre suppresses the feeling of hunger which results in reduced

food intake[68]. Common appetite suppressants in South Africa include Duromine®, Obesan X®, Relislim®, Obex LA® and Tenuate Dospan® [69].

1.8.1.2 Pancreatic lipase inhibitors

Pancreatic and gastrointestinal lipases are enzymes found in the lumen of the stomach and small intestine and play an important role in the digestion of dietary fats. Its primary function is to breakdown triglycerides into monoglycerides and free fatty acids which are easily absorbable. Lipstatin and its derivatives are selective inhibitors of these enzymes and bind to the active site of the gastric and pancreatic lipases. These inhibitors reduce the activity of lipases by inhibiting the hydrolysis of triglycerides, which are removed from the system undigested and reduces the absorption of monoacylglycerides and free fatty acids from the diet. Therefore reducing the calorie intake also reduce the amount of energy available for absorption[68]. Orlistat (XeneCal®) is a common lipase inhibitor prescribed in South Africa[69].

1.8.2 Drugs used in the Treatment of Depression that have an effect on weight gain and loss

Almost all antidepressants have the possible side-effect of weight gain. However, many individuals respond to antidepressants differently as a possible result of their genetic profile[70].

1.8.2.1 Selective Serotonin Reuptake Inhibitors (SSRIs)

Selective serotonin reuptake inhibitors (SSRIs) are first line pharmacological agents that are used to treat depression, mood, eating, panic and anxiety disorders. They have also been known to have a temporary effect on weight loss[70]. They act by increasing the serotonin levels in the synaptic cleft by preventing the re-uptake into the pre-synaptic nerve terminals. Various generic formulations are available as a cost-effective alternative to the originator drug. Some commonly available SSRIs in South Africa include: Fluoxetine (Prozac®, Deprozan®, Lorien®, ProHexal®, Ranflocs® and Sandoz-Fluoxetine®), Fluvoxamine (Luvox® and Faverin®), Escitalopram (Cipralex®, Citraz®, Dolin®, Lexamil®, Marprem®, Zitolex® and Zytomil®), Paroxetine (ARopax®, Deparoc®, Lenio®, Paxil®, Paroxetine Unicorn®, Serapress®, Texine® and XET20®), Citalopram (Arrow Citalopram®, Austell Citalopram®, Bio Clitalopram®, Cilate®,

Cilift®, Ciloram®, Cipramil®, Depramil®, Recita® and Talomil®) and Setraline (Aspen Sertraline®, Austell Sertraline®, Dyna Sertraline®, Serdep®, Serlife®, Sertra®, Sertaline Winthrop®, Zolid® and Zoloft®)[69].

Liver cytochrome P450 enzymes metabolise SSRIs. The majority of SSRIs undergo hepatic oxidative metabolism before being eliminated from the body. Genetic differences in the oxidative metabolism (pharmacokinetic and pharmacodynamic properties) can affect the levels of active drugs circulating in the system[71]. This could result in other drugs using this pathway not being metabolised, leading to toxic levels of SSRIs[72].

1.8.2.2 Tricyclic antidepressants (TCAs)

Tricyclic antidepressants (TCAs) primarily inhibit noradrenaline and serotonin reuptake, in addition to blocking histaminic, α-adrenergic and muscarinic receptors[72]. Although TCAs have been effective in the management of depressive disorders, their adverse anticholinergic side-effects and toxicity in overdose have rendered them unfavourable. In general, they are only prescribed as an alternative to patients not responding to first line therapy. TCAs currently available in South Africa include; Amitriptyline (Trepiline® and Amitriptyline®), Imipramine (Tofranil® and Ethipramine®), Clomipramine (Anafranil®, Clomidep® and Equinorm®), Dothiepin (Thaden®) and Trimipramine (Tydamine®)[69]. TCAs are metabolised by Cytochrome P450 family 2 subfamily D member 6 (*CYP2D6)* and Cytochrome P450 Family 1 Subfamily A Member 2 (*CYP1A2)* enzymes[72].

1.8.2.3 Monoamine Oxidase Inhibitors (MAO-Is)

Monoamine oxidase enzymes are responsible for the breakdown of catecholamines in the synaptic cleft. Monoamine Oxidase Inhibitors (MAO-Is) inhibit these enzymes, thereby increasing the noradrenaline, serotonin and dopamine concentrations. MAO-Is have been effective in the treatment of depressive disorders associated with hypersomnia, increased appetitie and weight gain[72]. These drugs are usually second line drugs and used when people do not respond to other treatment. Certain foods and beverages (e.g. beer and wine) can interact with MAOIs, therefore a strict diet needs to be followed and certain medications avoided to prevent blood pressure problems[73].

Examples of MAO-Is include Moclobemide (Depnil®) and Tranycypromine (Parnate®)[69]. MAO-Is are metabolised by cytochrome P450 isoenzymes.

1.8.2.4 Serotonin and noradrenaline reuptake inhibitors (SNRIs)

These drugs increase the availability of serotonin and noradrenaline in the brain, and elevate the mood through the selective inhibition of the presynaptic reuptake of serotonin and noradrenaline. Serotonin and noradrenaline reuptake inhibitors (SNRIs) are used in the treatment of depression, anxiety, pain syndromes and panic disorders[72, 74]. They are mostly used as second-line treatment. Examples of SNRIs include Venlafaxine (Efegen XR®, Efexor XR®, Illovex SR®, Odiven®, Sandoz-Venlafaxine®, Venlafaxine Unicorn XR®, Venlafaxine XR Adco®, Venlor XR®), Desvenlafaxine (Exsira®) and Duloxetine (Cymbalta®, Cymgen® and Yelate®)[69]. *CYP2D6* is the isoenzyme that metabolises SNRIs in the liver.

1.8.2.5 Serotonin receptor antagonist and reuptake inhibitors (SARIs)

Serotonin receptor antagonist and reuptake inhibitors (SARIs) drugs block neurotransmitter chemicals in the brain, such as serotonin or dopamine, thereby increasing serotonin levels [75]. Trazodone (Aspen Trazodone®, Biotech Trazodone® and Molipaxin®) is a 5-HT2A receptor antagonist. It is used in the treatment of insomnia and not as an antidepressant[72].

1.8.2.6 Dopamine reuptake inhibitors (DRIs)

Dopamine reuptake inhibitors (DRIs) inhibit the reuptake of dopamine by blocking the action of the dopamine transporter, resulting in an increased concentration of extracellular dopamine and dopaminergic neurotransmission. The increase in extracellular dopamine has been linked to increased susceptibiltiy to addicitve behaviour, therefore DRIs use has been restricted due to the high potential of abuse. They have also been used for the treatment of obesity and binge eating due to the appetite suppressant effects. Buproprion (Wellbutrin®) is the only approved DRI in South Africa for the management of MDD[72].

1.8.3 Drug Metabolism

CYP2D6 is a member of the cytochrome P450 family and is involved in the oxidative metabolism of drugs. It is expressed in the liver and central nervous system and is involved in the metabolism of xenobiotics. Approximately 20% of all drugs pass through this pathway, in particular antidepressants and antipsychotics. Genetic variations in *CYP2D6*, cause variability between individuals in terms of the activity of oxidative metabolism. There are four categories of *CYP2D6* activity: ultrarapid metaboliser (UM), extensive or normal metaboliser (EM), intermediate metaboliser (IM) and poor metaboliser (PM), Figure 10. The impact of the CYP2D6 activity differs from drug to drug and depends on whether CYP2D6 is an activator or inactivator of the drug. The *CYP2D6* allele is located on chromosome 22. The G allele is the normal form, while the A form causes a disruption to mRNA formation, resulting in a non-functional *CYP2D6* protein, *CYP2D6*4* (non-functioning variant). Carriers of two copies of the A allele (homozygotes) are classified as poor drug metabolisers[76, 77].

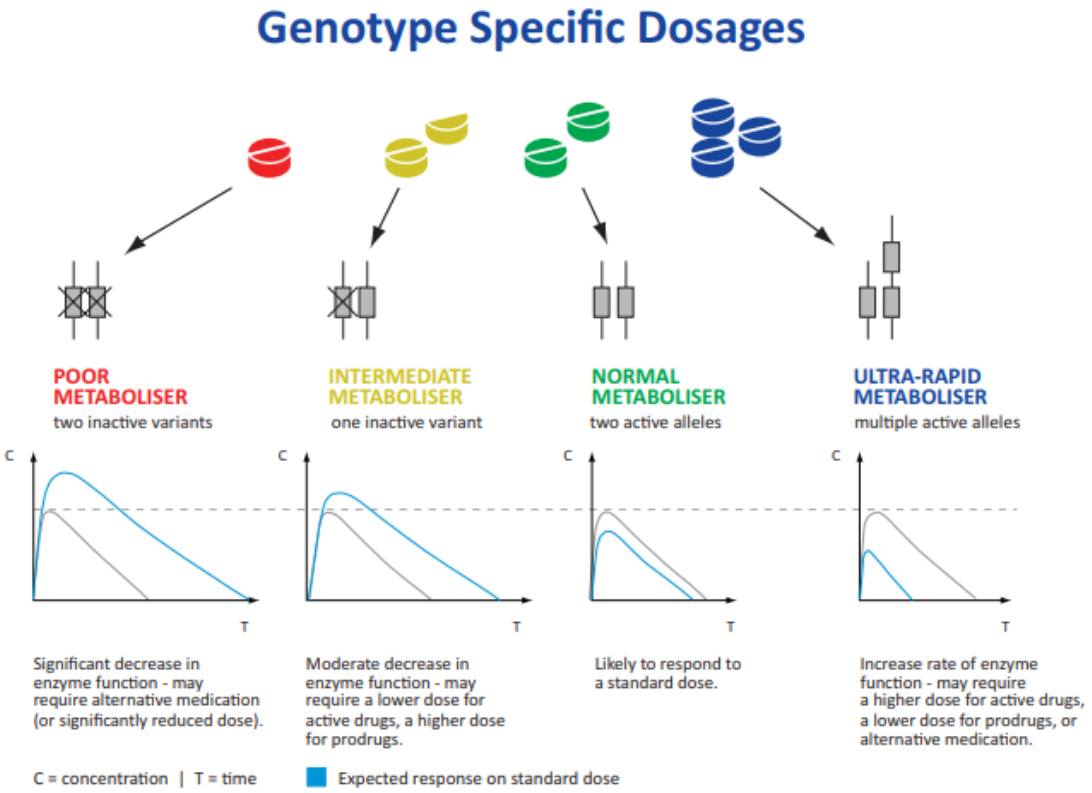


Figure 10: Cytochrome P450 genetic variations and corresponding metaboliser classification, GeneWay Laboratories.

In addition, other drugs may inhibit CYP2D6 activity, thereby causing a normal metaboliser to perform similar to an intermediate metaboliser. Furthermore, most SSRIs are substrates of CYP2D6, therefore SSRIs that are both inhibited and metabolised by CYP2D6, can inhibit their own metabolism and produce higher plasma concentrations[77]. It is therefore important to know and understand a patient's oxidative metabolism ability when prescribing any drugs.

1.9 The association between depression and obesity

Statistically females are more affected by depression than males. Females are also more overweight, in terms of numbers, compared to their male counterparts, suggesting an association between these two epidemics[78]. Studies propose that a portion of obese patients' weight problems may be related to uncontrolled carbohydrate intake with symptoms of atypical depression. However, the cause of this carbohydrate intake is not fully elucidated. A meta-analysis confirmed that obesity and depression are closely linked, in that each increases the risk of the other. Obese individuals have a 55% increased risk of developing depression while those with depression have a 58% increased risk of becoming overweight[79]. SSRIs have been used in the treatment of obesity with varying success. The hypothesis is that the cause of fluctuating serotonin levels in the brain may involve certain genetic factors. Studies have been able to show the following:

- After the intake of carbohydrate-rich foods, the serotonin-producing neurons synthesize and release more neurotransmitters[62, 80, 81].
- Animal studies have shown that carbohydrate intake is suppressed when serotonin-like drugs are systemically administered[82].
- Clinical studies showed that serotonin antagonists may cause weight gain[83, 84].
- Drugs which increase intra-synaptic serotonin levels have antidepressant properties[85].
- Animal studies have shown that systemic administration of nicotine increases serotonin release. This was confirmed by studies on long-term smokers attempting to free themselves from nicotine addiction[86].

However, the inter-relationship and role of different gene combinations coupled with environmental factors responsible for obesity and depression, has not been comprehensively elucidated. Therefore, the literature contains a multitude of contradictory reports.

1.10 Single Nucleotide Variations

Single nucleotide variations or SNVs are the most common form of genetic variation. It represents a change in a single nucleotide in the genome that differs between members of a species. The Human Genome project has enabled the identification of millions of SNVs. They occur approximately once in every 300 base pairs (bp). These SNVs act as biological markers, helping to locate genes and genetic variations that are associated with common diseases such as heart disease, diabetes mellitus, inflammation and cancer[87].

SNVs can occur in the coding sequence of genes, non-coding regions of genes or intergenic regions. Not all SNVs within the coding regions affect the amino acid sequence of the protein. There are two types of SNVs that occur in the coding region namely; synonymous and nonsynonymous SNVs. Synonymous SNVs do not affect the protein sequence, while nonsynonymous SNVs change the amino acid sequence and the protein. SNVs that occur in the non-coding regions can still affect gene splicing, transcription factor binding, messenger RNA degradation and RNA sequences[87].

In order to keep track of all the variations, the National Centre for Biotechnology Information (NCBI) in collaboration with the National Human Genome Research Institute (NHGRI), has created a free public archive for genetic variations within and across different species, called the Single Nucleotide Polymorphism Database (dbSNP). The database contains the following molecular variations: single nucleotide variations, short deletion and insertion variations, microsatellite markers or short tandem repeats, multi-nucleotide variations, heterozygous sequences and named variants. Every submitted variation receives a submitted single nucleotide polymorphism ID number (“SS#”). However unique submitted single nucleotide polymorphism records receive a reference SNV ID number (“rs#”, “refSNP cluster”), which is used to identify which SNVs are present during investigations[88].

Genetic predisposition involves the interaction between multiple genes and/or gene variations with an individual's exposure to environmental factors/stressors. Environmental factors/stressors typically include; family, friends, education, socio-economic status, environmental contaminates exposure and the availability of psychoactive drugs[89].

Food intake is controlled by the neurotransmitter's dopamine, GABA, noradrenaline and serotonin[90]. Potential genetic variables that could have an effect on hypodopaminergic function include serotonergic genes (serotonergic receptors (*5HTR2C*) and serotonin transfer *5-HTTLPR (SCL6A4*)), endorphinergic gene (mu *OPRM1* gene), GABAergic gene (*GABRA6*) and dopaminergic genes (*DRD2 / ANKKI Taq A1, DRD4, COMT*). Any of these genetic and environmental impairments could result in reduced release of dopamine and/or dopaminergic receptors.

The study focused on eight well described variations that have been linked to neurotransmitters in the brain associated with the brain reward cascade[38].

1.10.1 Serotonergic Receptor and Transport Variation

Serotonin-transporter-linked polymorphic region *(5-HTTLPR* and rs25531) is a degenerate repeat polymorphism (43 bp insertion/deletion), in the 5' regulatory region of the Serotonin transporter gene (*SLC6A4*). This polymorphism results in two common allelic variations of different lengths. The long (L) allele contains sixteen repeats and a short (S) allele contains fourteen repeats and each is associated with functional differences in serotonin reuptake. The "S" allele is associated with reduced transcription activity and lower levels of serotonin uptake in the lymphoblastic cells, hence increasing the availability of serotonin in the synapses [91]. The minor allele "G" is always in phase with the long (L) allele of the *5- HTTLPR*, which may be associated with less sensitivity to pain and lower levels of serotonin. While the major allele "A" is in the short phase and is associated with higher levels of available serotonin due to the lower levels of Serotonin uptake. Individuals with the A allele tend to be slightly less happy and may benefit from more support in terms of dietary guidance or medication utilization[92-98].

The 1000 Genomes project is the largest public data catalogue of human genetic variations and genotypes. The aim of the project was to find genetic variations with a

frequency of at least 1% in the populations studied[99]. The populations have been divided into 5 super populations namely, African (AFR), Ad Mixed American (AMR), East Asian (EAS), European (EUR) and South Asian (SAS). According to the statistics published by the 1000 Genomes project the allele frequency of the major allele (T) of the *SCL6A4* (rs25531) is 0.9105 and the minor allele (C) is 0.0895 for the European (EUR) population, and 77.9% and 22.01% respectively for the African (AFR) population, as shown in Figure 11 below. Mixed American (AMR) showed an allele frequency of 0.95 and 0.0504 for the major and minor allele respectively[100].

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	T=0.94959998	C=0.05040000
AFR	1322	T=0.77990001	C=0.22010000
EAS	1008	T=0.86809999	C=0.13190000
SAS	978	T=0.85689998	C=0.14309999
EUR	1006	T=0.91049999	C=0.08950000

Figure 11: 1000 Genomes Project allele frequency for *SCL6A4*, rs25531[100].

The serotonin (5-hydroxytryptamine (5-HT)) neurotransmitter has been linked to a variety of complex neurological, psychiatric and behavioural disorders. Behavioural disorders include depression, anxiety, schizophrenia, obsessive-compulsive disorder and appetite/satiety[101]. The serotonin 5-HT-2C receptor (*HTR2C*) gene is located on chromosome Xq24 and has been shown to be involved in the control of appetite and feeding behaviour and the development of obesity in knockout mice studies[102-106]. Rs3813929, also known as -759C/T, is the most frequent variation in the 5-*HTR2C* gene. The T allele showed borderline significant association with higher BMI and incidence of lifetime major depressive disorder among 4978 persons from the European Prospective Investigation into Cancer (EPIC)-Norfolk study. However, only the association with BMI remained borderline significant within the full EPIC-Norfolk cohort (20,981 persons)[107]. The statistics published by the 1000 Genomes Project reflect an allele frequency of 0.8807 for the major allele (C) and 0.1193 for the minor allele (T) for the European (EUR) population, and 0.9909 and 0.0091 respectively for

the African (AFR) population, as shown in Figure 12. The mixed American (AMR) population had a similar allele frequency distribution, 0.9107 for the major allele © and 0.0893 for the minor allele (T)[108].

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	C=0.91070002	T=0.08930000
AFR	1322	C=0.99089998	T=0.00910000
EAS	1008	C=0.88889998	T=0.11110000
SAS	978	C=0.79240000	T=0.20760000
EUR	1006	C=0.88069999	T=0.11930000

Figure 12: 1000 Genomes Project allele frequency for *HTRC2*, rs3813929[108].

1.10.2 Endorphinergic Gene Variation

Another potential pathway involved in food cravings and overeating is the opioid receptor system. The mu-opioid receptor gene (*OPRM1*) can be found on the q arm of chromosome 6 (6q24) and is one of four genes that protein product attach to opioids[11]. The G allele in exon 1 (rs1799971 / A118G) of the mu opioid receptor *OPRM1* gene causes the normal amino acid at residue 40, asparagine (Asn), to be replaced by aspartic acid (Asp). Carriers of at least one G allele appear to have a stronger craving for alcohol and are at a higher risk for alcoholism[109-112]. The allele frequency as calculated by the by the 1000 Genomes Project is 0.838 for the major allele (A) and 0.162 for the minor allele (G) for the European (EUR) population, 0.7997 and 0.2003 for the Mixed American (AMR) population and 0.9909 and 0.0091 respectively for the African (AFR) population, as shown in Figure 13[113].

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	A=0.79970002	G=0.20029999
AFR	1322	A=0.99089998	G=0.00910000
EAS	1008	A=0.60710001	G=0.39289999
SAS	978	A=0.58179998	G=0.41819999
EUR	1006	A=0.83800000	G=0.16200000

Figure 13: 1000 Genomes Project allele frequency for *OPRM1*, rs1799971[113].

1.10.3 GABAergic Gene Variation (*GABRA6*)

GABA is the major inhibitory neurotransmitter in the brain where it acts on GABA-A receptors. GABA is located on the q arm of chromosome 5 (5q34). *GABRA6* (rs3219151) is a new discovered variation situated in the 3' untranslated region of the GABA α-6 receptor subunit gene. There is a substitution of a T to a C at nucleotide 1521 which results in the loss of an AlwNI restriction site. The T allele showed high salivary cortisol levels, and carriers of the T allele tend to deal with stress better[114-117]. Figure 14 reflects the statistics published by the 1000 Genomes Project. The allele frequency of the major allele (T) is 0.5765 and the minor allele (C) is 0.4235 for the European (EUR) population. While for the African (AFR) population the major allele © had a frequency of 0.5666 and the minor allele (T) had a frequency of 0.4334. A similar allele frequency was observed for the mixed American (AMR) population, 0.611 and 0.389 for the major and minor alleles[118].

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	C=0.61100000	T=0.38900000
AFR	1322	C=0.56660002	T=0.43340001
EAS	1008	C=0.68150002	T=0.31850001
SAS	978	C=0.58080000	T=0.41920000
EUR	1006	T=0.57650000	C=0.42350000

Figure 14: 1000 Genomes Project allele frequency for *GABRA*6, rs3219151[118].

1.10.4 Dopaminergic Gene Variation

1.10.4.1 Dopamine D2 Receptor (*DRD2*)

Drug addiction and abuse have been strongly linked to the brain reward circuit. One of the most studied variations is the *Taq A1* allele of the dopamine D2 receptor (*DRD2*) gene which has been associated with cocaine, alcohol and opioid abuse. The brain reward system that promotes drug abuse, is involved in pleasure seeking behaviour, which in turn is linked to food cravings[11]. *Taq A1* variation (rs1800497) of the DRD2 gene is located 10 kbp downstream of the gene. This *Taq A1* (T) allele is associated with a reduced number of dopamine binding sites in the brain[119]. Abnormal craving-associated behaviours could be linked to defects in the DRD2 and the dopaminergic gene, Catechol-O-methyltransferase (*COMT*). Carriers of the A1 *DRD2* gene variant may exhibit unhealthy cravings, overeating, drug abuse tendencies and susceptible to chronic stress. Affected individuals often seek substances that activate the dopaminergic pathway in order to offset their low expression of D2 receptors. Studies have shown a link between dopamine D2 receptor quantity and leptin receptor genes in obese individuals, where these individuals display a decreased dopamine receptor availability[11, 120-123].

The 1000 Genomes Project revealed the following allele frequency distribution for the European population: 0.8121 for the major allele (G) and 0.1879 for the minor allele (A). In the African population the allele frequency for the major allele (G) was 0.615

and for the minor allele (A) 0.385[124], as shown in Figure 15. The mixed American (AMR) population showed an allele frequency of 0.6888 for the major allele and 0.3112 for the minor allele.

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	G=0.68879998	A=0.31119999
AFR	1322	G=0.61500001	A=0.38499999
EAS	1008	G=0.59420002	A=0.40580001
SAS	978	G=0.68510002	A=0.31490001
EUR	1006	G=0.81209999	A=0.18790001

Figure 15: 1000 Genomes Project allele frequency for *DRD2 / ANKK1*, rs1800497[124].

1.10.4.2 Catechol-O-Methyltransferase (*COMT*)

The *COMT* gene polymorphism rs4680 (Val158Met), is a well-studied single nucleotide variation. The *COMT* gene codes for the COMT enzyme, which breaks down dopamine in the brain's prefrontal cortex. The wild-type allele is a G, coding for a valine amino acid (Val) and the (A) substitution changes the amino acid to a methionine (Met). This alters the structure of the resultant enzyme such that its activity is only 25% of the wild type. As a result, A allele carriers have increased dopamine levels and lower COMT enzymatic activity in their prefrontal cortex, which may be responsible for the neuropsychological associations. Carriers of the A allele may have a lower pain threshold and tend to be vulnerable to stress. Carriers of the A allele are commonly termed “worriers”. Carriers of the G allele are termed “warriors” and have higher COMT enzymatic activity and lower dopamine levels. These carriers tend to have a higher pain threshold and may be more resilient to stress[125-130]. The AA genotype patients get more pleasure out of life but also more misery, they have more positive emotions, more creative, higher IQ, better working memory and cogitative function and better reading comprehension. But on the other hand, they also handle stress worse, suffer from anxiety, are more impulsive, more depressed and have a high risk of ADHD[125, 131].

In terms of the allele frequency the major (G) and minor (A) alleles are found in equal frequencies in the European populations. While in the African population, the major (G) allele is more common, 0.7193, as opposed to the minor (A) allele, 0.2806[132], similarly the mixed American (AMR) population 0.6225 and 0.3775 for the major and minor alleles respectively, refer to Figure 16 below.

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	G=0.62250000	A=0.37750000
AFR	1322	G=0.71939999	A=0.28060001
EAS	1008	G=0.72020000	A=0.27980000
SAS	978	G=0.55930001	A=0.44069999
EUR	1006		A=0.50000000 G=0.50000000

Figure 16: 1000 Genomes Project allele frequency for *COMT*, rs4680[132].

1.10.4.3 Dopamine D4 Receptor (*DRD4*)

Polymorphic dopamine D4 receptor (*DRD4*) has more than 200 identified SNVs and several variable number tandem repeats (VNTRs). This exonic variant contains a 48bp repeat sequence, which changes the length of the third intracellular loop of the receptor potentially affecting the signalling efficiency of the receptor. Two variants of the DRD4 gene have been well studied and have been reported to be associated with human approach related traits, such as novelty seeking and extraversion. The DRD4 seven repeat allele (long) which contains seven repeats of the 48bp sequence appears to have decreased sensitivity to dopamine compared to the four repeat alleles (short). Carries of the long allele (C) have been shown to have higher novelty seeking and risk-taking traits, compared to noncarriers[133-135]. In genetics, rs1800955, also known as C-521T and -521C/T, is the best studied variant in the promoter region upstream of the *DRD4* gene on chromosome 11[134, 136, 137].

The major or short allele (T) has frequency of 0.5885, while the minor or long allele (C) has frequency of 0.4115 in the European population. A similar distribution is seen in

the African population, T = 0.5923 and C = 0.4077, according to the statistic published by the 1000 Genomes Project[138], refer to Figure 17. The mixed American (AMR) population had a slightly higher allele frequency for the major allele (T) 0.6268 and 0.7372 for the minor allele (C).

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	T=0.62680000	C=0.37320000
AFR	1322	T=0.59230000	C=0.40770000
EAS	1008	T=0.59420002	C=0.40580001
SAS	978	T=0.57980001	C=0.42019999
EUR	1006	T=0.58850002	C=0.41150001

Figure 17: 1000 Genomes Project allele frequency for *DRD4*, rs1800955[138].

1.10.5 Leptin Receptor (*LEPR*) Variation

Leptin regulates satiety and energy expenditure. Its levels are dependent on the amount of fat present in the adipose cells. The leptin receptor is a single transmembrane protein belonging to a superfamily of cytokine receptors. The Leptin Receptor gene (*LEPR*) is located on chromosome 1 and variation rs1137101, also known as 668A>G, involves a substitution of an A for a G. The A allele encodes the Gln allele, while the G allele encodes the Arg allele. The G allele has been associated with an increased risk of obesity and type 2 diabetes[63, 65, 139-146].

The 1000 Genomes Project, revealed the following allele frequency distribution for the European population, 0.5308 for the major allele (A) and 0.4692 for the minor allele (G). In the African (AFR) population the allele frequency for the major allele (G) was 0.5923 and for the minor allele (A), 0.4077[147], Figure 18. For the mixed American (AMR) population the major allele was A and the minor allele G, with the respective allele frequencies of 0.5634 and 0.4366.

Population ID -Class	Sample (2n)	Major Allele Freq.	Minor Allele Freq.
AMR	694	A=0.56339997	G=0.43660000
AFR	1322	G=0.59230000	A=0.40770000
EAS	1008	G=0.86900002	A=0.13100000
SAS	978	G=0.50309998	A=0.49689999
EUR	1006	A=0.53079998	G=0.46919999

Figure 18: 1000 Genomes Project allele frequency for *LEPR*, rs1137101[147].

1.11 Prevention and Management of Obesity and the role of genomics

Although the urgency to mitigate the increasing rate of obesity has been acknowledged, research into the effectiveness of preventative strategies is still lacking. There is a need to explore the genetic determinants responsible for the biology behind weight regulation and how environmental changes affect gene expression, in order to establish effective long-lasting interventions. Suggested preventative strategies are described below.

1.11.1 Obesity preventative programs

As previously discussed in Sections 1.2 and 1.3, obesity is to a great extent hereditary. Studies into the effectiveness of obesity screening and preventative strategies based on family history, could be beneficial. This is especially important in young children who are at a high risk and preventing them from gaining excess weight at a young age.

1.11.2 Nutrigenomics

Nutrigenomics is the study of the interaction of nutrition and genes with regards to the prevention and treatment of disease, including weight management[122]. The rationale behind nutrigenomics is, that a person's optimal diet is determined by their genetic composition. Personalized intervention to assist in the prevention and management of obesity, is gaining support, however more research into the efficacy of this approach is needed.

1.11.3 Pharmacogenetics

Pharmacogenetics is the study of inherited genetic differences in terms of drug metabolic pathways, which can affect an individual's response to drugs, in terms of effectiveness and adverse effects. SNVs in important drug metabolising enzymes affect the transportation, metabolism and receptor binging ability of drugs, influencing the pharmacokinetic (what the body does to the drug) and pharmacodynamic (what the drug does to the body) properties of medications. Obesity genetics extend beyond understanding the effects of weight loss drugs, but also include how other drugs, such as SSRI's, may adversely affect individuals in terms of weight control. Furthermore, by understanding gene expression patterns, it could benefit researchers in identifying novel targets for the development of new anti-obesity drugs.

1.11.4 Environmental approaches

Lowe (2003) found that people who were at a high risk of developing obesity, showed difficulty in self-regulating their food intake to maintain an optimal weight, even though they had the knowledge of how to do so. The study reported that programs that focus on cognitive-behavioural interventions (such as lifestyle changes) were not as effective. The study did not take into consideration the threat of biological stimuli to eating, especially when there is an abundance of food and individuals are at a genetic predisposition. The study concluded that weight reduction programs advocating portion controlled, nutrient-dense meals may be more successful for long-term weight loss and management, as it eliminates the possibility for food stimuli (overabundance of foods, large portion sizes and high calories)[148].

The "one-size fits all" approach to reducing or preventing obesity is proving to be ineffective. Individual gene variations that affect behaviour, satiety and taste, which forms part of a genomic approach may appear to be useful. Considering the effects of relevant genes in the genome and their interaction with one another and the environment, could provide insight into novel and effective obesity prevention and intervention strategies. Appreciating obesity as being a complex disorder, the identification of specific genes associated with this disease may form the foundation to understanding the effect of environmental and lifestyle factors on the development of obesity.

1.12 Aims and Objectives

1.12.1 Study Aim

The aim of the study was to determine the prevalence of eight single nucleotide variations (SNVs) associated with the control and regulation of neurotransmitters in the brain reward cascade, which have been linked to addictive behaviour and food cravings in overweight and obese individuals.

1.12.2 Study Objectives

The objectives of the study were:

1. To validate the TaqMan® OpenArray™ Genotyping assay for SNV genotyping.
2. To determine the presence of eight SNVs, namely *SCL6A4, HTR2C, OPRM1, GABRA6, DRD2, DRD4, COMT* and *LEPR* known to be associated with food craving tendencies, in overweight and obese individuals.
3. To determine the allele and genotype frequency of the eight SNVs in the study population.
4. To evaluate how the allele frequencies observed in the study population compared to the general population reported by the 1000 Genomes Project and HapMap Project, in terms of demographic groups.
5. To determine the allele frequency of the eight SNVs in overweight/obese individuals compared to the control group of normal weight individuals.
6. To evaluate the difference in allele frequencies among ethnic groups.

CHAPTER 2: METHODOLOGY

2 MATERIALS, STUDY DESIGNS AND METHODS

2.1 Materials

2.1.1 Study population, sampling and sample size

This was a cross-sectional analytical study using DNA collected from a subject and a control group. Participants were selected according to the following two strata:

1. Normal weight (BMI score between 18.5 - 24.9 kg/m^2) – these participants formed the control group.
2. Overweight and obese (BMI score between 25 and greater than 40.0 kg/m^2) – these participants formed part of the subject group. Classification:
    - Overweight individuals with a BMI between 25.0 kg/m^2 and 29.9 kg/m^2 were collected.
    - Class 1 (low risk) obese individuals with a BMI between 30.0 kg/m^2 and 34.9 kg/m^2 were collected.
    - Class 2 (moderate risk) obese individuals with a BMI between 35.0 kg/m^2 and 39.9 kg/m^2 were collected.
    - Class 3 (high risk) obese individuals with a BMI greater than 40.0 kg/m^2 were collected.

A partially stratified sampling method was used. Participants were recruited from two independent locations: An obesity clinic and general family practice clinic. Both control and subject groups mimicked each other in terms of race and gender.

The primary analysis consisted of associating overweight and obesity participants with single nucleotide variations *SCL6A4, HTR2C, OPRM1, GABRA6, DRD2, DRD4, COMT* and *LEPR*. The expected proportion of overweight and obese participants (BMI 25.0 to greater than 40 kg/m^2) is 0.35. The association between overweight and obesity and the SNVs were derived from a logistic regression sample size based on events per variable that exceeded five[149]. From the expected prevalence, sample sizes are

130 for obesity. Thus, a sample size of 200 participants was recommend to be studied. The latter sample of participants assumed a cross-sectional study, Figure 19.

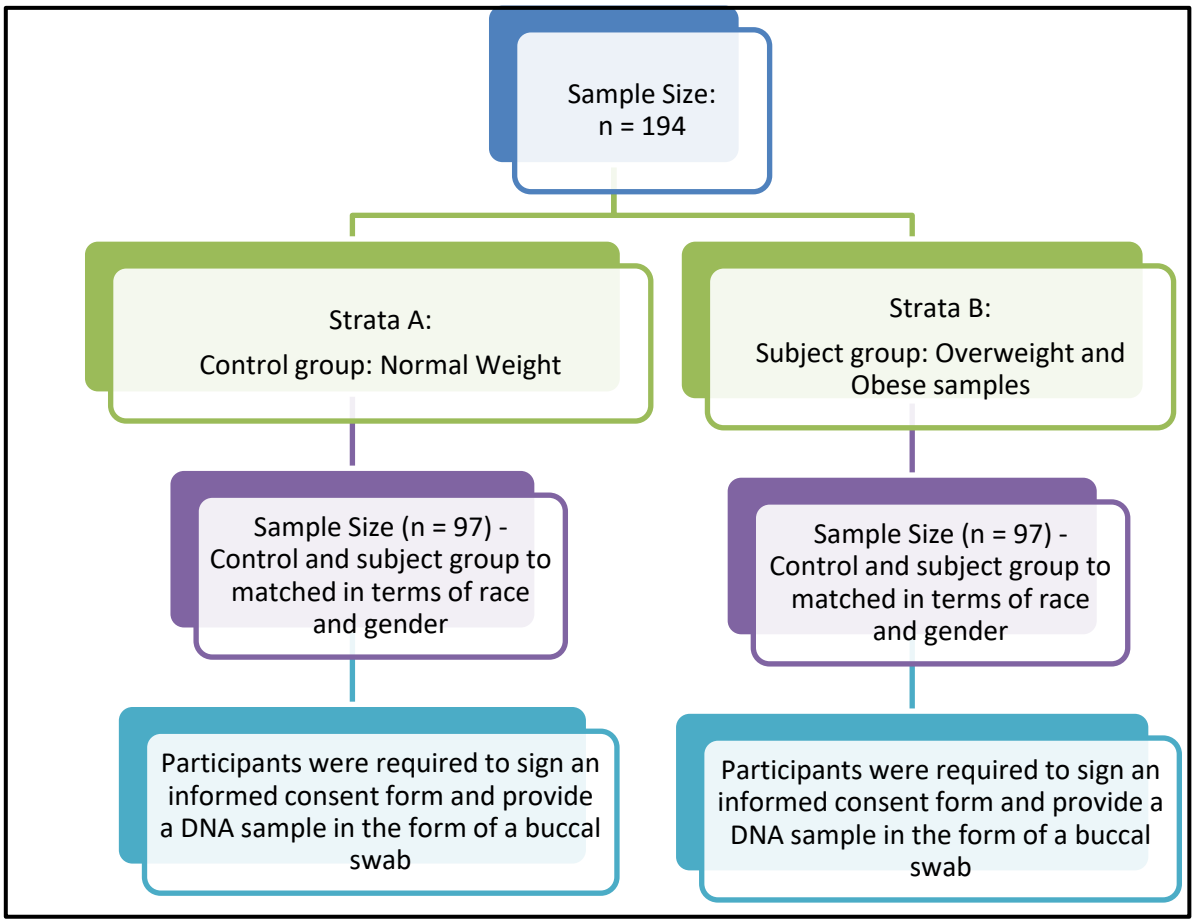


Figure 19: Flow diagram of sample collection strategy. Strata A will refer to the normal weight controls while Strata B are the group of overweight and obese samples.

2.1.1.1 Inclusion criteria

Willing participants over the age of 18 years, of any ethnic group with a BMI score greater than 18.5 kg/m^2.

2.1.1.2 Exclusions criteria

Participants were excluded from the study if their BMI score was below 18.5 kg/m^2, indicating severe thinness. Samples where the DNA samples collected were of poor quality were also excluded due to the potential for lack of amplification. Analysis was conducted on samples with complete genotypic results for all eight SNVs.

2.1.1.3 Participation requirements and assurances

Healthcare Practitioners from an Obesity Clinic and Family Medical Clinic were asked to assist with sample collection. They were made familiar with the protocol's aims and objectives and invited patients to participate in the study. The Healthcare Practitioners identified potential candidates for the study and introduced the study to the potential participants. Practice staff were trained in terms of the project aim and objectives in order to explain the project adequately to potential participants. The practice staff assisted with the explanation of the study and taking of the DNA sample in the form of a buccal swab. All participants completed and signed the informed consent form. Access to patient medical records by the investigator was not required. The Healthcare Practitioner and staff aided in the completion of the consent form. Information required included age, gender, race, anthropometrics and current medication use. Practitioners gave their written consent to conduct sampling within their practices. Refer to Appendix A4 for all permission letters.

All participants were given the necessary assurance that all information supplied was completely anonymous. No identifiable information was or will be in future revealed. Participants were assigned laboratory reference numbers as a means of associating DNA samples with consent forms. Participants were not penalized or negatively affected by their participation, nor were they remunerated for their participation.

2.1.2 Selection of Single Nucleotide Variations

After a bibliographic search for genetic variations previously studied relating to mood disorders and food cravings, which were also related to the brain reward cascade, eight genetic alterations (Table 5) present in the general population, were selected for the design of the customised OpenArray™ plate containing assays for all the indentified SNV's on a single plate.

The sequences (approximately 300–500 bp) flanking the single nucleotide variation changes were obtained from the database of the National Centre for Biotechnology Information (NCBI - http://www.ncbi.nlm.nih.gov/) and Ensembl (http://www.ensembl.org/index.html) websites. The changes were flagged, marked and customised using the FileBuilder software (Applied BioSystems).

The OpenArray™ format chosen was the 32 x 96 well plate. Each custom plate contained 28 assays and has the potential to accommodate 96 patient samples (94 samples with one positive and one negative control). Eight of the 28 SNVs selected and designed were specific for this project. Table 5 below lists the information of the 8 SNVs relevant to this study.

Table 5: Selected single nucleotide variations associated with mood disorders and food cravings.

No.	NCBI Reference Number	Symbol	Name	Sequence
1	rs1800497	*DRD2 / ANKK1*	Ankyrin repeat and kinase domain containing 1	CACAGCCATCCTCAAAGTG CTGGTC[A/G]AGGCAGGCG CCCAGCTGGACGTCCA
2	rs25531	*SLC6A4*	Solute carrier family 6 member 4	CTCGCGGCATCCCCCCTGC ACCCCC[A/G]GCATCCCCCC TGCAGCCCCCCCAGC
3	rs1799971	*OPRM1 A118G*	Opioid receptor mu 1 A118G	GGTCAACTTGTCCCACTTAG ATGGC[A/G]ACCTGTCCGAC CCATGCGGTCCGAA
4	rs4680	*COMT*	Catechol-O-methyltransferase	CCAGCGGATGGTGGATTTC GCTGGC[A/G]TGAAGGACAA GGTGTGCATGCCTGA
5	rs1800955	*DRD4*	Dopamine receptor D4	GGGCAGGGGGAGCGGGCG TGGAGGG[C/T]GCGCACGA GGTCGAGGCGAGTCC
6	rs3813929	*5HTR2C*	5-hydroxytryptamine receptor 2C	CTGCTCTTGGCTCCTCCCC TCATCC[C/T]GCTTTTGGCC CAAGAGCGTGGTGCA
7	rs1137101	*LEPR*	Leptin Receptor	ATCACATCTGGTGGAGTAAT TTTCC[A/G]GTCACCTCTAAT GTCAGTTCAGCCC
8	rs3219151	*GABRA6*	Gamma-aminobutyric acid type A receptor alpha6 subunit	AATTGGAAATCTGTAACGCA GCTTC[C/T]GTAAGCATGTG TGGGCAAAAAAGCA

2.1.3 Work Flow Schematic

Single nucleotide variations selection and OpenArray plate design

1. NCBI reference numbers and gene sequence acquired and tabulated into an Excel spreadsheet.

2. Excel spreadsheet sent through to manufacture for design and processing.

3. Positive Control sequenced by external company (Inqaba Biotechnology)

Sample (DNA) Collection - Buccal Swab.

Sample collection conducted at two locations:

1. Obesity Clinic &

2.Family Clinic

DNA Extraction: MagMax DNA Multi-sample ultra kit using magnetic beads in conjunction with MagMax 96 Express automated extraction equipment.

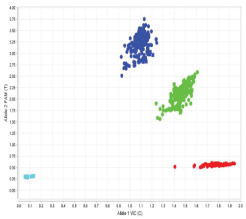
Quantification and Normalization of DNA: DNA quantified using Qubit reader, 20-50 ng of DNA is required per sample.

Pre-Amplification - TaqMan Pre-Amp master mix and Pre-Amp Pool (primers and probes) specific for the single nucleotide variations on OpenArray Plate.

DNA transferred to OpenArray Plate. PreAmplifed product transferred from the 96 well plate to a 384 well plate containing TaqMan OpenArray Genotyping Master Mix. The Accufill system transfers the product from the 384 well plate to the OpenArray plate.

Applied Biosystems QuantStudio (QS) Real-Time PCR. The OpenArray plate is loaded into the QS K12 flex, and the Genotyping protocol is initiated using the QuantStudio Software.

Genotype Analysis: The experiement is analysed using the Taqman Genotyper Software, which distinquishes between homozygotes and heterosygoes based on the fluorescent signal. Data is then exported in a csv file for statistical analysis.



Statistical Analysis

To ensure accuracy and reproducibility of results a portion of the samples was repeated.

2.1.4 Statistical Consideration

Of fundamental importance to this study was the association of SNVs *SLC6A4, DRD2, DRD4, COMT, OPRM1, 5HTR2C, GABRA6* and *LEPR* (Table 5) in normal weight of individuals compared to overweight and obesity individuals. The prevalence of these SNVs in terms of their allele and genotype frequencies was evaluated. The allele frequency differences in terms of ethnicity and gender was also investigated.

2.1.4.1 Data analysis

The data summary reported proportion and 95% confidence intervals overall, by overweight and obesity groups individually. Data analysis primarily made use of weighted logistic regression analysis, with odds ratio (OR) along with 95% confidence intervals being the statistics of interest. Allele and genotype frequencies in the sampled population (overweight/obese and normal weight participants) was determined. The allele frequencies were evaluated in terms of ethnic and gender group, and compared between subject group (overweight and obese) and control group (normal weight). The Hardy-Weinberg equilibrium was used to determine if the observed genotype frequencies in the population differed from the frequencies predicted by the equation.

2.1.4.2 Limitations and bias

There were two major limitations to the study:

1. The epigenetic interaction of the participant's environments was not investigated. Therefore, the environmental implications and influences were missed, of importance to the study have would been factors such as stress, diet, habits and medication. Meaning that despite the inherited genetics being monitored the genetic activity was not monitored.

2. The study was conducted using a relatively small number of DNA samples of participants suffering from overweight and obesity. One of the assumptions of the Hardy-Weinberg equilibrium is that the population size is infinitely large which could be a limitation to the statistical analysis. With the exception of patients attending the general family practice, sampling was not done

completely randomly. Sites have been selected where the prevalence of obesity is high. Therefore, a weighted logistic regression analysis was conducted to account for any bias introduced by the selected sampling sites. Any estimates derived from the study may not reflect representative patterns occurring in the general population. The Hardy-Weinberg equation does not include an indication of evolutionary processes acting on the population and requires further investigation in future studies.

2.1.5 Ethical Consideration

Approval from the MSc Scientific Committee and Research and Ethics Committee, Faculty of Health Sciences, University of Pretoria was obtained prior to the commencement of the project, protocol number 515/2017, Appendix A1. The study was conducted in accordance with the 2013 revised version of the declaration of Helsinki, 1975. Permission to conduct the study and recruitment of patients from the Healthcare Practitioners was obtained prior to the sampling and collection of DNA (Appendix A2). No identifiable information collected from participants was shared. Access to patient medical history files by investigator was not required.

Permission to obtain DNA samples from the identified participants was dependent on their written, signed informed consent (Appendix A3). The aims and objectives of the study was explained to each participant by either the Healthcare Practitioner at the participating site or the Healthcare Practitioner's practice staff who were trained on the protocol. Participants were issued with a patient information leaflet containing the researcher's contact information (Appendix A4). Participation was voluntary and participants were assured of their anonymity and confidentiality. Partakers were not remunerated for their efforts and were allowed to withdraw their consent at any time during the study.

2.1.6 Measures

Objective 1: To validate the TaqMan® OpenArray™ Genotyping assay for SNV genotyping.

TaqMan® OpenArray™ Genotyping results were compared to a sample that was sequenced through an external company in addition approximately 17% of the samples were repeated on the OpenArray™ analysis to ensure consistency of results.

Objective 2: To determine the presence of eight SNVs, namely *SCL6A4, HTR2C, OPRM1, GABRA6, DRD2, DRD4, COMT* and *LEPR* associated with food craving tendencies in overweight and obese individuals.

Objective two was achieved through quantitative analysis of collected DNA samples, using the Applied Biosystems QuantStudio 12K Flex Real-Time PCR and OpenArray™ custom designed plates. Participants were genotyped for all eight SNVs.

Objective 3: To determine the allele and genotype frequency of the eight SNVs in the study population.

STATA IC15.1 was used to statistically determine the allele and genotype frequencies of the eight SNVs. STATA makes use of the Hardy-Weinberg assumptions. A SNV is a bi-allelic locus, thus allele 1 has a frequency of p, and allele 2 has a frequency of $1-p$. Therefore the expected genotype frequencies follow $Binom(2,p)$. The TaqMan Genotyper Software, aids in distinguishing the homozygotes and heterozygotes by utilizing the fluorescent signal to divide the signals into four quadrands in order to determine the genotypes. An AIF file, which accompanies the OpenArray™ plates from the manufacturers is uploaded into the Genotyper program, and assigns the alleles accordingly to the VIC and FAM dyes. The output file from the TaqMan Genotyper Program, is in a tab delimited format. All unnecessary data is removed from the file, so that only the sample IDs and SNP genotypes are present in the file. This .csv file is then uploaded into STATA IC15.1. The statistical software reads the imported file to determine both the allele and genotype frequencies based on the above assumptions. The statistician, made the necessary assignments in order for the STATA program to read the imported file.

Objective 4: To evaluate how the allele frequencies observed in the study population compared to the general population reported by the 1000 Genomes Project and HapMap, in terms of demographic group.

Data from the 1000 Genomes Project was extracted from their web site and HapMap data was obtained from the NCBI website. Two populations were assessed, namely

European and African. A comparative table was drawn up to compare the observed values with the expected values obtained from the 1000 Genomes Project and HapMap to establish the affirmation or deviation in the study population.

Objective 5: To determine the allele frequency of the eight SNVs in overweight/obese individuals compared to the control group of normal weight individuals.

Allele frequency is the relative frequency of an allele (gene variant) in a particular gene in a population. It is used to describe the extent of variation in the population. From the quantitative data generated above, the allele frequencies for each of the eight SNVs were determined using the statistical software STATA IC15.1, and the Fisher exact test.

Objective 6: To evaluate the difference in allele frequencies among ethnic groups.

A multinomial logical regression test was performed using STATA IC15.1 software. Unfortunately, our study population was small, with the white population accounting for the majority of the samples collected. The White and African ethnic groups were compared. The Coloured and Indian ethnic groups were excluded from this analysis due to the small sample numbers, 1 and 3 respectively.

2.2 Methods

2.2.1 DNA collection

A DNA sample was collected from participants using a flocked buccal swab (FLOQSwabs™ from Copan Diagnostics Incorporated, Murrieta, USA). This was a simple non-invasive means of collecting DNA from participants. Participants were required not to eat or drink anything 30 minutes prior to taking the DNA swab. The polyester buccal swab was rubbed against the inside of the cheek for 30 seconds on either side. Cheek cells are rich in DNA and contain the same genetic material as blood cells.

2.2.2 DNA genomic extraction

Genomic DNA was extracted from the buccal swab using the MagMAX™ DNA Multi-Sample Ultra kit (Applied Biosystems, California, USA), according to the manufacturer's instruction in conjunction with the automated extraction instrument,

MagMax™ Express 96 (Applied Biosystems, California, USA). The MagMAX™ DNA Multi-Sample Ultra kit contains, Proteinase K, lysis buffer, magnetic beads, wash buffers, elution buffers and nuclease free water. Biological grade absolute Ethanol and Isopropanol were added to the wash buffers according to manufacturer's specifications, prior to use. Proteinase K was used for the digestion of any contaminating proteins and nucleases in the solution. The addition of the lysis buffer breaks down the cell membrane in order to release the DNA. Precipitation of the DNA was achieved through the addition of isopropanol into the solution as DNA is insoluble is Isoproponal. Magnetic beads act as a binding agent to, which the DNA attaches, according to the salt concentration of the solution. The MagMax™ Express 96 (Applied Biosystems, California, USA) is an automated system pulls the magnetic beads with the attached DNA out of the solution through the use of a that magnetic field. The beads are washed several times before the DNA is eluted and the magnetic beads removed[150].

2.2.3 Sample normalization and pre-amplification

The DNA concentration of the samples were quantified using the Qubit® 2.0 fluorometer (Invitrogen, California, USA). Samples were normalised to 15 ng/µL, dilutions were calculated using the formula, (C1) (V1) = (C2) (V2) and nuclease-free water was used as the dilutant. Samples with a DNA concentration between 1–19 ng/µL, were not normalized and used at the DNA concentration in the pre-amplification reaction. Pre-amplification was preformed using TaqMan® custom pool and TaqMan® PreAmp Master Mix following manufacturer's instructions. Each reaction requires 2.5 µL TaqMan® PreAmp Master Mix, 1.25 µL TaqMan® custom pool and 1.25 uL patient DNA. PCR conditions were as follows; initial denaturation at 95˚C for 10 minutes, followed by 12 cycles of denaturation at 95˚C for 15 seconds and annealing/extension at 60˚C for 4 minutes, a final inactivation step at 99.9˚C for 10 minutes, before cooling down to 4˚C[151]. The Pre-Amp products were diluted 1:20 with 10 mM Tris-HCl containing 1 mM EDTA, pH 8.0 (1 x TE Buffer).

2.2.4 96 and 384 Well Plate Set-up

Exactly 2.5 µL of diluted pre-amplification product of each sample was transferred from the 96 well plate to the 384-well plate, containing 2.5 µL of the TaqMan® OpenArray™

Genotyping Master Mix, using the distribution pattern determine by the Sample Tracker Software. An adjustable pipette was utilized for the transfer of the samples from the 96 well plate to the 384 well plate. The 384 well plate was sealed with a foil seal, gently shaken to ensure a uniform mixture of TaqMan® OpenArray™ Genotyping Master Mix and DNA sample, spun down for 10 seconds to remove any bubbles.

2.2.5 The TaqMan® OpenArray™ Genotyping Platform

The TaqMan® OpenArray™ Genotyping platform (Applied Biosystems, California, USA) allows for high performance and high-throughput real-time Polymerase Chain Reaction (PCR) technology, enabling the simultaneous analysis of several independent SNVs across several samples. The default plate layout that was employed was 28 SNVs by 96 samples[152]. Each run on one OpenArray™ plate, contained 94 samples, one positive control (sequenced sample) and one negative control (water sample).

The hydrolysis of the TaqMan® probes used for the TaqMan® OpenArray™ Genotyping technique, requires a pair of primers, which have been manufactured and fluorescently labelled by Thermo Fisher Scientific®, UK, when manufacturing the OpenArray™ plates. One primer corresponded to the wild-type sequence and the other corresponded to the mutant sequence. Two different Minor Groove Binder (MGB) probes for each assay were also required. The one probe was labelled with the VIC® fluorophore and the other probe was labelled with the FAM® fluorophore. Each probe was associated with either the wild-type or mutant sequence, Figure 20. The plates' layout for the OpenArray™ assays was composed of 48 sub-arrays (4.5 mm × 4.5 mm), each with 64 Nano-wells. The surface of the plate has hydrophobic properties, while the interior was hydrophilic in nature. The probes and primers required for the reactions were inserted in the interior of the wells by the manufacturer. These physical properties of the wells enabled small quantities (33 nL) of DNA to be loaded into the wells.

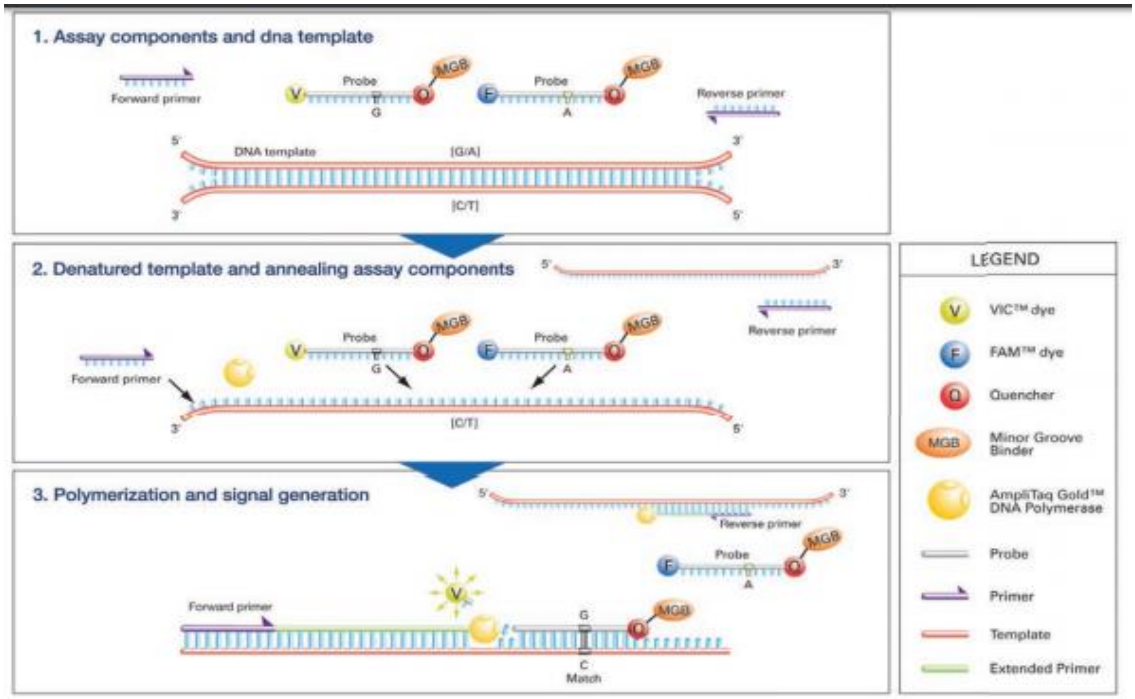


Figure 20: The TaqMan® SNV genotyping assay, Thermo Scientific TaqMan® OpenArray™ Genotyper User Manual[152].

2.2.6 Assay Design for OpenArray™ Plate design

Primer design is the corner stone to any successful PCR reaction. Thermo Fisher Scientific, the suppliers and manufacturers of the OpenArray™ plates, has 4.5 million predesigned genotyping assays, which have been functionally tested and validated. It is more practical to order OpenArray™ plates, with Thermo Fisher's predesigned assays as there is lower risk of assay failure, as they have already been both tested and optimised to work with TaqMan® OpenArray™ Master Mix. The assay makes use of the TaqMan® assay-based (5' nuclease) chemistry. Which amplifies and detects specific polymorphisms in genomic DNA. GC content is of utmost important when designing an assay, as it can affect both the annealing temperature and folding properties.

Table 6, shows the information extracted from Thermo Fishers Scientific's website of the eight assays selected. Seven of the eight assays, used predesigned assays and ordered accordingly. SLC6A4 required a customised assay design. The assays were designed according to the content sequence in the below table (Table 6). To ensure

optimal assay performance a quality check of the target sequence was performed. Appendix B1 shows the results of the BLAST® search results of the target sequence using the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The FASTA sequence underwent a BLAST analysis to ensure that the correct gene was being assessed. The primer labelled with the VIC® and FAM® dyes detecting Allele 1 and 2, have been highlighted below on the FASTA sequence in green and blue respectively.

Table 6: Thermo Fisher Scientific Predesign Information and Custom Assay Design

	Assay ID	Availability	Catalog No.	Assay Type	NCBI No.	Gene Symbol	Gene Name	Context Sequence [VIC/FAM]	NCBI Assembly Location	NCB Assembly Build	SNP Type
1	C___7486676_10	Made to Order	4351379	Functionally Tested	rs1800497	*ANKK1*	Ankyrin repeat and kinase domain containing 1	CACAGCCATCCTCAAAGTGCTGGTC[A/G]AGGCAGGCGCCCAGCTGGACGTCCA	Chr.11: 113400106 on GRCh38	38	Mis-sense Mutation,Transition Substitution,Intragenic
2	Requires Probe design				rs25531	*SLC6A4*	Solute carrier family 6 member 4	CTCGCGGCATCCCCCCTGCACCCCC[A/G]GCATCCCCCCTGCAGCCCCCCCAGC	Chr. 17: 30237328 on CRCh37.917		
3	C___8950074_1_	Made to Order	4351379	Validated	rs1799971	*OPRM1 A118G*	Opioid receptor mu 1 A118G	GGTCAACTTGTCCCACTTAGATGGC[A/G]ACCTGTCCGACCCATGCGGTCCGAA	Chr.6: 154039662 on GRCh38	38	Intron,Mis-sense Mutation,Transition Substitution,Intragenic
4	C__25746809_50	Inventoried	4362691	DME	rs4680	*COMT*	Catechol-O-methyltransferase	CCAGCGGATGGTGGATTTCGCTGGC[A/G]TGAAGGACAAGGTGTGCATGCCTGA	Chr.22: 19963748 on GRCh38	38	Transition Substitution;Mis-sense Mutation;Intragenic
5	C___7470700_30	Made to Order	4351379	Functionally Tested	rs1800955	*DRD4*	Dopamine receptor D4	GGGCAGGGGGAGCGGGCGTGGAGGG[C/T]GCGCACGAGGTCGAGGCGAGTCC	Chr.11: 636784 on GRCh38	38	Intron,Transition Substitution,Intragenic
6	C__27488117_10	Made to Order	4351379	Functionally Tested	rs3813929	*5HTR2C*	5-hydroxytryptamine receptor 2C	CTGCTCTTGGCTCCTCCCCTCATCC[C/T]GCTTTTGGCCCAAGAGCGTGGTGCA	Chr.X: 114584047 on GRCh38	38	Intron,Transition Substitution,Intragenic
7	C___8722581_10	Made to Order	4351379	Validated	rs1137101	*LEPR*	Leptin Receptor	ATCACATCTGGTGGAGTAATTTTCC[A/G]GTCACCTCTAATGTCAGTTCAGCCC	Chr.1: 65592830 on GRCh38	38	Mis-sense Mutation,Transition Substitution,Intragenic
8	C__11263956_10	Made to Order	4351379	Functionally Tested	rs3219151	*GABRA6*	Gamma-aminobutyric acid type A receptor alpha 6 subunit	AATTGGAAATCTGTAACGCAGCTTC[C/T]GTAAGCATGTGTGGGCAAAAAAGCA	Chr.5: 161701908 on GRCh38	38	Transition Substitution, UTR 3, Intragenic

Alignment of primers on FASTA sequence:

1. Target Sequence:

CACAGCCATCCTCAAAGTGCTGGTC[A/G]AGGCAGGCGCCCAGCTGGACGTCCA

rs1800497 *DRD2*

(Kindly note the below FASTA sequence is in the reverse compliment orientation, due the manner in which the discovery sequence has been published in the NCBI database)

```
TCCAGGCGAG AGGCCCCAAG TAGTCTAAAT TTCTTTCTTT CTTTCTTTTT TATATGGAGT
 CTCGCTCTGT TGCCCAGGCT GGAGTGCAGT GGTGCGATCT CGGCTCACTG CAACCTCTGC
 CTCCTGGGTT CAAGGAATTC TCCTGCCTCA GCCTCCCTGG TAGTTGGGAT TACAGGCACG
 TGCCACCATA CCCAGCTAAA TTTTGTATTT TTAGCAGAGA CAGGGTTTTG CCATGTTGGC
 CAGGCTGGCC TCAAACTCTT GATATCAGGT GATCTGCCTG CCTCAGCCTC CCAAAGTGCT
 GGGATTACAG ACGTGAGCCA CCACGGCTGG CCAAGTTGTC TAAATTTCCA TCTCGGCTCC
 TGGCTTAGAA CCACCCAGAG TGGCCACTGA CGGCTCCTTG CCCTCTAGGA AGGACATGAT
 GCCCTGCTTT CGGCTGCGGA GGGCCAGTTG CAGGGGTGTG CAGCTCACTC CATCCTGGAC
 GTCCAGCTGG GCGCCTGCCT
 [C/T]
 GACCAGCACT TTGAGGATGG CTGTGTTGCC CTTGAGGGCG GCCAGGTGGG CGGGTGTCCA
 GCCCACCTTG TTGCGGGCGT GGACATTTGC GTGATGTTCT AGGAGGTTGA TGACACTCAG
 GAAGGTGCTC CTCTGGACCG CCAGGTGGAG GGGTGTCCAG CCTGACTGCT CTGCAGCATT
 GGGGTCAGCC CCACACTGCA GCAGTGCTGA CACCACCGCC TCCTCCCCGT GGCGTGCAGC
 TAGGTGCAGG GGAGTCCAGT TCACAGCTCC AAGAGCACCC ATGTTTGCGT GGCTCTCTGC
 CAGCAGATGG ATGATCTCCA GGTGGCCCTT GTAGGCTGCT AGATGCAGGG GTGTCCAGCC
 CTGGTGGGTG GGCAGCTCAA GGCTGGCTCC GTACCTGAGC AGCATCTTGC AGATCAGGTA
 TTTGCCCCTG GCAGCTGCAG TGTGCAGTGG GCCGTAGCCG CTCTGGTCAA GGGCATCAGG
 GACCGCTCCA CTCTTCAGCA
```

2. Target Sequence:

CTCGCGGCATCCCCCCTGCACCCCC[A/G]GCATCCCCCCTGCAGCCCCCCCAG
C

rs25531 *SLC6A4*

```
TCTCCCGCCT GGCGTTGCCG CTCTGAATGC CAGCACCTAA CCCCTAATGT CCCTACTGCA
 GCCCTCCCAG CATCCCCCCT GCAACCTCCC AGCAACTCCC TGTACCCCTC CTAGGATCGC
 TCCTGCATCC CCCATTATCC CCCCCTTCAC CCCTCGCGGC ATCCCCCCTG CACCCCC
 [A/G]
 GCATCCCCCC TGCAGCCCCC CCAGCATCTC CCCTGCACCC CCAGCATCCC CCCTGCAGCC
 CTTCCAGCAT CCCCCTGCAC CTCTCCCAGG ATCTCCCCTG CAACCCCCAT TATCCCCCCT
 GCACCCCTCG CAGTATCCCC CCTGCACCCC CCAGCATCCC CCCATGCACC CCCGGCATCC
 CCCCTGCACC CCTCCAGCAT TCTCCTTGCA CCCTACCAGT ATTCCCCCGC ATCCCGGCCT
 CCAAGCCTCC CGCCCACCTT GCGGTCCCCG CCCTGGCGTC TAGGTGGCAC CAGAATCCCG
 CGCGGACTCC ACCCGCTGGG AGCTGCCCTC GCTTGCCCGT GGTTGTCCAG CTCAGTCCCT
 CTAGACGCTC AGCCCAACCG GCCGCACAGT TTTCAGGGGT CAGTTCCTCC AAGTACAAGG
 GGCGGTGGCT TCTCTGGAGC TGCAAACTTG TCACTGCTAT TTCCTTTCGG TCTTCTACTT
 CCTATCGTTC CTGGCCTCCT CTTGGGGAGA GGTAGAGCCC TCTCCTTTCC GCCTCAGGGA
 CAACCCAAAG CAAGTACTGC ATGTGCCCTT TTTAAAGTTT TAAATAATTT TAGCAAAAAG
 GATATTAACA TTAAATCAAT TTTTAAACTT TTTGAAAAAA TTATCAAAAC TACATGCACA
 TGGTTCAAAA CAATAGGCTC CTGCTGGGCC CTTTCAGATA ATTCAAATTG
```

3. Target Sequence:

GGTCAACTTGTCCCACTTAGATGGC[A/G]ACCTGTCCGACCCATGCGGTCCGAA

rs1799971 *OPRM1*

```
TGTGTTTGCA CAGAAGAGTG CCCAGTGAAG AGACCTACTC CTTGGATCGC TTTGCGCAAA
 ATCCACCCCT TTTCCCTCCT CCCTCCCTTC CAGCCTCCGA ATCCCGCATG GCCCACGCTC
 CCCTCCTGCA GCGGTGCGGG GCAGGTGATG AGCCTCTGTG AACTACTAAG GTGGGAGGGG
 GCTATACGCA GAGGAGAATG TCAGATGCTC AGCTCGGTCC CCTCCGCCTG ACGCTCCTCT
 CTGTCTCAGC CAGGACTGGT TTCTGTAAGA AACAGCAGGA GCTGTGGCAG CGGCGAAAGG
 AAGCGGCTGA GGCGCTTGGA ACCCGAAAAG TCTCGGTGCT CCTGGCTACC TCGCACAGCG
 GTGCCCGCCC GGCCGTCAGT ACCATGGACA GCAGCGCTGC CCCCACGAAC GCCAGCAATT
 GCACTGATGC CTTGGCGTAC TCAAGTTGCT CCCCAGCACC CAGCCCCGGT TCCTGGGTCA
 ACTTGTCCCA CTTAGATGGC
 [A/G]
 ACCTGTCCGA CCCATGCGGT CCGAACCGCA CCGACCTGGG CGGGAGAGAC AGCCTGTGCC
 CTCCGACCGG CAGTCCCTCC ATGATCACGG CCATCACGAT CATGGCCCTC TACTCCATCG
 TGTGCGTGGT GGGGCTCTTC GGAAACTTCC TGGTCATGTA TGTGATTGTC AGGTAAGGAA
 AGCGCCAGGG CTCCGAGCGG AGGGTTCAGC GGCTTAAGGG GGTACAAAGA GACACCTAAC
 TCCCAAGGCT CAATGTTGGG CGGGAGGATG AAAGAGGGGA GGTAAACTGG GGGGACTCTG
 GAGGAGACCA CGGACAGTGA TTGTTATTTC TATGAGAAAA CCTACTTTTC TGTTTTTTCT
 TCAACTGATA AAGAAAGAAT TCAAAATTTC AGGAGCAGAG AAGTTGCTTT GGTAAAAGCT
 ACAAATGTCT AGGGGTGGGG GGCGGAGGGA AGCTATAGCA TAGACTTGGA GCGCTTCCTT
 ATACTGAGCA AAGAGGGCTC
```

4. Target Sequence:

CCAGCGGATGGTGGATTTCGCTGGC[A/G]TGAAGGACAAGGTGTGCATGCCTGA

rs4680 *COMT*

```
AGAGGGCAGC TCTGTGTTAG GACACACTGG GGCCAGCCAG GAAGGGTGGA AAAGATAGGG
 ACCAGCGTGA GCATAGAGGC TAAGGGACCA TGGGAGCTCC AAGCGCGCTC ACAGTGGGGA
 CCAGGTCCTG GGGGCTGGGG ACACCAGGGA GGTGAAATAC CCCTCCAGCG GGTAGGGAGG
 GTGGGCAGAG GAGGGCCAGC GGCCAGGCAT TTGGGAGGGG CTCCTGCTCT TTGGGAGAGG
 TGGGGGGCCG TGCCTGGGGA TCCAAGTTCC CCTCTCTCCA CCTGTGCTCA CCTCTCCTCC
 GTCCCCAACC CTGCACAGGC AAGATCGTGG ACGCCGTGAT TCAGGAGCAC CAGCCCTCCG
 TGCTGCTGGA GCTGGGGGCC TACTGTGGCT ACTCAGCTGT GCGCATGGCC CGCCTGCTGT
 CACCAGGGGC GAGGCTCATC ACCATCGAGA TCAACCCCGA CTGTGCCGCC ATCACCCAGC
 GGATGGTGGA TTTCGCTGGC
 [A/G]
 TGAAGGACAA GGTGTGCATG CCTGACCCGT TGTCAGACCT GGAAAAAGGG CCGGCTGTGG
 GCAGGGAGGG CATGCGCACT TTGTCCTCCC CACCAGGTGT TCACACCACG TTCACTGAAA
 ACCCACTATC ACCAGGCCCC TCAGTGCTTC CCAGCCTGGG GCTGAGGAAA GACCCCCCCA
 GCAGCTCAGT GAGGGTCTCA CAGCTCTGGG TAAACTGCCA AGGTGGCACC AGGAGGGGCA
 GGGACAGAGT GGGGCCTTGT CATCCCAGAA CCCTAAAGAA AACTGATGAA TGCTTGTATG
 GGTGTGTAAA GATGGCCTCC TGTCTGTGTG GGCGTGGGCA CTGACAGGCG CTGTTGTATA
 GGTGTGTAGG GATGGCCTCC TGTCTGTGAG GACGTGGGCA CTGACAGGCG CTGTTCCAGG
 TCACCCTTGT GGTTGGAGCG TCCCAGGACA TCATCCCCCA GCTGAAGAAG AAGTATGATG
 TGGACACACT GGACATGGTC
```

5. Target Sequence:

GGGCAGGGGGAGCGGGCGTGGAGGG[C/T]GCGCACGAGGTCGAGGCGAGTCC

rs1800955 *DRD4*

```
TGGGGTCCCA CAGAGTGGTG CCCCCTTTTA GTGTCTTCTA GGCCCCTTAG TGACAGACTA
 CAGAAAATAC CTCTCAGGTC ACAGGTCACC CCTCTTTGGT GAAGAGTCCA TAGAATTCTC
 TGCTGCGCTT TGCAAGCACT TTCTCTTCTG CACGTTTGGA ACCTACCCCG GCCTGTCGTG
 TCTTTCTCCT GGCCTCCTCG CGAGCCGAAC CTACTGTCCG GTCCCGGGAC CCCCTGCCCA
 GGGTCAGAGG GGCGCCTACC TAGCTCACGG TCTTGGGCCG GAGGGAATGG AGGAGGGAGC
 GGGGTCGACC GCTCAGCTGT CCGCCCAGTT TCGGAGGCGG CCACGCGAGG ATCAACTGTG
 CAACGGGTGG GGCCGCGGCT GACCGTGGTG GTCGCGGGGG CTGAGGGCCA GAGGCTGCGG
 GGGGGGGGCG GCGGGATGAG CTAGGCGTCG GCGGTTGAGT CGGGCGCGGA GTCGGGGGCA
 GGGGGAGCGG GCGTGGAGGG
 [C/T]
 GCGCACGAGG TCGAGGCGAG TCCGCGGGGG AGGCGGGCAG AGCCTGAGCT CAGGTCTTTC
 TGCGTCTGGC GGAACGGGCC TGGGAGGGAG GTTTTGCCAG ATACCAGGTG GACTAGGGTG
 AGCGCCCGAG GGCCGGGACG CACGCACGGG CCGGGTAGGA TGGCGCTGGC GTCGATGCCC
 GCGCGCTTCA GGGCCTGGTC TGGCCGCCCC TCCATCCTTG TCGGTTTCTC GGGTCGCGGA
 CCCCGCGCGG CGCCGGGCGA TGCTGGCCTG CCCGTGGCCA CCACCTCGCT TCATTCCCGT
 CTCTTTGGGC CGCCGCATTC GTCCACGTGC CCGTCTCTCC CTGCGCAAAA TTCCAAGATG
 AGCAAATACT GGGCTCACGG TGGAGCGCCG CGGGGGCCCC CCTGAGCCGG GGCGGGTCGG
 GGGCGGGACC AGGGTCCGGC CGGGGCGTGC CCGAGGGGAG GGACTCCCCG GCTTGCGACC
 CGGCGTTGTC CGCGGTGCTC
```

6. Target Sequence:

CTGCTCTTGGCTCCTCCCCTCATCC[C/T]GCTTTTGGCCCAAGAGCGTGGTGCA

rs3813929 *5HTR2C*

```
GACAAGGATG GGGAAGTGGG CCTTATAACA GGATTGTGGC CTTTGCGCAC TCACCAAATG
 TTTGACCCTG TGAGTGCCTC AGTTGCTACT GTTGGAAGAA TGGGCAAGAG TCGGAACAGA
 GACCCTTGAA GGGAGTTTCA AAGCTTGATG AAATTTGCAA GACTTGAGAA TGCTGTTTGT
 TGAAATGAAA TGTACAGGGG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TTTGGGGAGG
 GGTATGCTAT GAATCTTTGA GGGTACATTC TTGAGAAAGC CTTCCCTCTC TCTCTATCCG
 GTGCCATGGC TGATCCTGGT TCCCCCTACT CTCTAGGCCT TGTGAATCAG ATTAATCATC
 ACCCCCACCC CCATCTCCAC CATGGGGTCT CGCGCCCCCT GCCAGCAGGC TCCAGATGCA
 CTAAGAGACC GGTCCAAACA GGCCCGGGGG CCACGTAATG CTGAGTGCTG ATTGGCTGCT
 CTTGGCTCCT CCCCTCATCC
 [C/T]
 GCTTTTGGCC CAAGAGCGTG GTGCAGATTC ACCCGCGCGA GGTAGGCGCT CTGGTGCTTG
 CGGAGGACGC TTCCTTCCTC AGATGCACCG ATCTTCCCGA TACTGCCTTT GGAGCGGCTA
 GATTGCTAGC CTTGGCTGCT CCATTGGCCT GCCTTGCCCC TTACCTGCCG ATTGCATATG
 AACTCTTCTT CTGTCTGTAC ATCGTTGTCG TCGGAGTCGT CGCGATCGTC GTGGCGCTCG
 TGTGATGGCC TTCGTCCGTT TAGAGTAGTG TAGTTAGTTA GGGGCCAACG AAGAAGAAAG
 AAGACGCGAT TAGTGCAGAG ATGCTGGAGG TGGTCAGTTA CTAAGCTAGA GTAAGATAGC
 GGAGCGAAAA GAGCCAAACC TAGCCGGGGG GCGCACGGTC ACCCAAAGGA GGTCGACTCG
 CCGGCGCTTC CTATCGCGCC GAGCTCCCTC CATTCCTCTC CCTCCGCCGA GGCGCGAGGT
 TGCGGCGCGC AGCGCAGCGC
```

7. Target Sequence:

ATCACATCTGGTGGAGTAATTTTCC[A/G]GTCACCTCTAATGTCAGTTCAGCCC

rs1137101 *LEPR*

```
CTTTGGTATG TCTGAAAAAA AAAGCCTTTA TTTCATCATT ATTTTGAAAG CTGTTTTCGC
 TGGGTATAGG ATTTTAGAAT TGCAGTTTTT CTTTTATTTT AGTACTTCAC TTTTACGTCA
 TTATCTTTTT GCTTATGTTA TTCCTGATGA TTAACCTGCT GTAATCTTTA TCTTTGTTTT
 TCTAATGTAG GGTTTTTTTT TTTCAGATAC CCTTTAAGCT GGGTGTCCCA AATAGTTTAC
 TTCAATTAGT ATTTAGTATC CTGCTTTAAA AGCCTATCCA GTATTTTCAT ATCTGTTTTA
 ATATTTAGCT CTTATTTTTC AATATAGGCC TGAAGTGTTA GAAGATTCAC CTCTGGTTCC
 CCAAAAAGGC AGTTTTCAGA TGGTTCACTG CAATTGCAGT GTTCATGAAT GTTGTGAATG
 TCTTGTGCCT GTGCCAACAG CCAAACTCAA CGACACTCTC CTTATGTGTT TGAAAATCAC
 ATCTGGTGGA GTAATTTTCC
 [A/G]
 GTCACCTCTA ATGTCAGTTC AGCCCATAAA TATGGGTAAG TTATGCACTA AAATGATGAT
 AATAGGTCTA AACATCAGTC ATATATAAAG GTTAAAAATT GCTTACAAAA ATATTTGCTA
 GCTTATCTCA CTTTGCTTAA CACTGTAATG ATGGTAGATG TAGTACTGGG GGTATTAAGA
 GTGGCTTCTA GAATGATTTA ACAATGGTAT GTATATCTCT GCCATTGTCA CTTAAATTCT
 GTTTTGAAAA CTGTTTTCTT TCAATCCTGG ATCTATGTAA TGGATGTATA TTGATTGGAT
 ATCACTTTTT CACATCTCAG ATAACTATTT TTGAAAATAG TAGCATGTTT CTTGCCTGAA
 TTTATTCCTT CAATAAATAT TTCTTAGAGG CTCATGTTTG TCAGAGACTG CTCCAGGAGC
 TGGAAAAAGA GTGGGACATT AGACATAGTT CCCACCTCAG AGAGCAGGGA CAAACAATAG
 TAGGCAGAGA CAATGATAAA
```

8. Target Sequence:

AATTGGAAATCTGTAACGCAGCTTC[C/T]GTAAGCATGTGTGGGCAAAAAAGCA

rs3219151 *GABRA6*

```
ATATTTGTCA ATGGTGAAAG AGTGAATAAA TAAGCAATTA AGCAATATCT ATTCTTTCAT
 TTGGGCTTAA TATTTGTCTT TTTTCCACAG CATCCTGACT CCAAATATCA TCTGAAGAAA
 AGGATCACTT CTCTGTCTTT GCCAATAGTT TCATCTTCCG AGGCCAATAA AGTGCTCACG
 AGAGCGCCCA TCTTACAATC AACACCTGTC ACACCCCCAC CACTCTCGCC AGCCTTTGGA
 GGCACCAGTA AAATAGACCA GTATTCTCGA ATTCTCTTCC CAGTTGCATT TGCAGGATTC
 AACCTTGTGT ACTGGGTAGT TTATCTTTCC AAAGATACAA TGGAAGTCAG TAGCAGTGTT
 GAATAGCTTG CGGCCAGGAC AACCTGAATT CTATAAGTTC TTGTTTTCTG TTTCCTATGT
 TTTCTTAAAA AATAGCATTG AGACTTGTGT AGATGCTTCT CAGAACATGA AATCAATTG
 GAAATCTGTA ACGCAGCTTC
 [C/T]
 GTAAGCATGT GTGGGCAAAA AAGCAATAAT CCTACTCCTC AAAATAGAAA GTTGAAGATT
 GCTGAAAAAT ATGACTTTTC TGTATGTTAG AGAAAAACTT TATGAGGATG AAATGGGTTC
 AAGATGAATT TGTCAACTTT TGTCTTCCAT TGTTCAGTAT TTTTAATTGT CACTGTAAAT
 AACATTTACC ACAAGGCAGA TAAAATAAGA AATGCTGACA CTTCCAAAGG TTGCCTTAAA
 ATATGTTTAT TTTGGCTTAG TTCCCGAGAG GGCAAAATAT AAATACAGTC TAAATATTTA
 TCAGTAGGTT AATACCAGCA TGTTGGAGGC CTTTATGCTA GTAAAATGGC TTTCAGTGGC
 ATTGTAAAGC CTACATTGAG CTTAGCCATT TGTTTTTAAC CTCGCTGTGC TCTTTTACCT
 CAATAAAATG TGGTGTTTGT ATACATATAA ATTATACATA GCTCATAAAT TATGTATGCA
 TATGTACATA GCTGTAGTTG
```

2.2.7 Genotyping using the TaqMan® OpenArray™ Genotyping Platform and validation

An automated pipetting system (OpenArray™ AccuFill® System) (Applied Biosystems, California, USA), connected to a computer controls the loading of the DNA from the 384-well plate to the OpenArray™ by means of the manufacturer's proprietary software.

The AccuFill® system transfers the DNA from the 384 well plate to the OpenArray™ plate. A volume of 33 nL is dispensed into each Nano-well. The OpenArray™ plate is sealed with a glass lid using the OpenArray™ Plate Press and stabilized with immersion fluid. The immersion fluid was slowly injected into the sealed OpenArray™ Plate at a constant uniformed rate, as not to disturb or dislodge the DNA in the wells. The AccuFill® system creates a loaded file by merging the imported 384 well plate design file and the OpenArray™ set-up file. The OpenArray™ plate was then loaded into the QuantStudio™ 12K Flex Real-Time PCR System, using the QuantStudio™ software. The QuantStudio™ software requires the importation of the loaded file. Standard thermocycling conditions were used: initial 10 minutes hold at 95˚C, followed by 40 cycles of a 15 second denaturation at 92˚C and one-minute annealing at 60˚C.

2.2.8 Control Samples

To ensure the accuracy of the OpenArray™ plate results, it is good laboratory practice to include one positive control and one negative control with any nucleic acid amplification technique[153]. One sample was selected to be sequenced for the eight SNV's which the study focuses on, which was used as the positive control throughout the study. The genotype for all eight SNVs of the positive control was confirmed through sequencing, conducted by an external company (Inqaba Biotechnology, Pretoria, RSA) to avoid any internal bias. No amplification of the positive control, would have resulted in the run being annulled. If the incorrect genotype for the positive control was analysed the run would also be considered nullified and need to be repeated. Incorrect analysis of genotype of the positive control could indicate the possibility of a sample switch. A negative control was also included to ensure the absence of any contamination during the processing of the samples. Amplification of the negative control, results in the run being null and void, in which event the run would need to be

repeated. If the negative control does not amplify the run is considered to be successful. The run is considered accurate and further analysis can then be conducted. Samples that showed no amplification for one or more of the SNVs, these samples were excluded from analysis.

2.2.9 Sequencing of the Positive Control

Sequencing is the gold standard for genetic genotyping. To ensure accuracy of the results and in order to validate the test, a positive control was included in each run. A positive control could have been created in one of two ways to validate the test:

1. A positive control could have been synthesized with a particular genotype.
2. Confirm the genotype of a sample through the sequencing (gold standard) and include the sample in each run to ensure consistency of results between runs.

The second option was opted for, due to the cost implication associated with synthesizing a positive control.

Sanger sequencing was conducted by Inqaba Biotechnology, Pretoria, RSA, utilising BrilliantDye™ Terminator V3.1 cycle sequencing kit (Nimagen, Netherlands) and the ABI 3500XL Genetics Analyser (Thermo Fisher Scientific®, UK) according to manufacturer's instructions. All sequencing products were purified prior to sequencing using the Zymo Kit D4053 from Zymo Research, California, USA, as per manufacturer's instructions. Primers were designed by Inqaba Biotechnology, Pretoria, RSA to amplify the target region of interest (Table 7). Results of the sequencing and report generate by Inqaba Biotechnology can be found in Appendix B3.

2.2.10 Primer Design for Sequencing

When designing Sanger sequence primers, there are a couple of important factors to consider. These include: Primer length (range between 18 to 22 bases), GC content (50% to 55%), GC lock at the 3' end, and melting temperature (approximately 50˚C). The melting temperature may require finding a sequence with a higher GC content or extending the length of the primer. Samples high in dinucleotide repeats (for example ACCCC or ATATATATAT) are additionally problematic for primer design and may require having to sequence a larger region of the gene[154]. Table 7 comprise a list of

primers designed by Inqaba Biotechnology, Pretoria, RSA used to sequence the SNV of interest.

2.2.11 Statistical Analysis

Statistical analysis of the data was done using STATA IC15.1 software, where Fisher's Exact Test p-values, OR and 95% confidence intervals for the OR were computed.

Table 7: Sanger Sequence Primer Design

Gene	Primer Orientation	Primer Sequence
DRD2 (rs1800497)	Forward	ACCTGGAGATCATCCATCTG
	Reverse	AATTTCCATCTCGGCTCCTG
SLC6A4 (rs25531)	Forward - V1	GTTGCAGGGGAGATCCTGGGAGAGG
	Reverse - V1	CCTCCTAGGATCGCTCCTGCATCC
	Forward - V2	GGCGTTGCCGCTCTGAATGC
	Reverse - V2	GAGGGACTGAGCTGGACAACCAC
	Forward - V3	GGTAGGGTGCAAGGAGAATGCTGGAG
	Reverse - V3	CTGCAACCTCCCAGCAACTCCCTGTAC
	Forward - V4	CTGAAGAGGAATCGGCTCTGGGC
	Reverse - V4	CGATGTTCACTCCAAATGATGTGC
OPRM1 (rs1799971)	Forward	GCTATACGCAGAGGAGAATG
	Reverse	ACATGACCAGGAAGTTTCCG
COMT (rs4680)	Forward	AAAAGATAGGGACCAGCGTG
	Reverse	TTTTCCAGGTCTGACAACGG
DRD4 (rs1800955)	Forward	CCCTTAGTGACAGACTACAGAAA
	Reverse	TAGTCCACCTGGTATCTGGCAAA
5HTR2C (rs3813929)	Forward	TCCAGATGCACTAAGAGACC
	Reverse	GCTAGGTTTGGCTCTTTTCG
LEPR (rs1137101)	Forward	CCTGCTTTAAAAGCCTATCCAGT
	Reverse	ACCCCCAGTACTACATCTACCAT
GABRA6 (rs3219151)	Forward	CAGTGTTGAATAGCTTGCGG
	Reverse	CTAGCATAAAGGCCTCCAAC

2.1.1 Reproducibility of results

To ensure the accuracy and reproducibility of the results, 39 (approximately 17.5%) of the patient samples were repeated, to the extent that the DNA sample was run and analysed for a second time, and results compared to ensure consistency.

2.1.2 Genotyping Analysis

Initial analysis was conducted using the QuantStudio™ software to evaluate the quality of the run. Quality control (QC) images were exported to assess accurate loading of the OpenArray™ Plate. The experiment was analysed and saved, where after it was imported into the TaqMan® Genotyper software for interpretation of the results according to the graphs of clusters generated. Results were validated by the inclusion of negative and positive controls (that were validated via sequencing).

CHAPTER 3: RESULTS

Obesity is a major public health problem and has increased significantly over the past few decades. The disease is notoriously difficult to treat. In general, society has historically perceived obesity as a behavioural disorder: with overweight and obese individuals lacking the willpower and self-control. However, this assumption is receiving more interest in the scientific community. The aim of the project was to assess the prevalence of eight SNVs associated with the brain's reward system in a population of normal weight and overweight/obese individuals. Furthermore, to establish whether there is any correlation between the eight SNVs and a higher BMI score.

3.1 Sequencing Results and Positive Control: Validation of TaqMan® OpenArray™ Genotyping Assay

To ensure accuracy and consistency of results across the three OpenArray™ plates that were processed and analysed, the sequenced positive control was included in each run. The positive control was sequenced by an external company, Inqaba Biotechnology, to ensure the results obtained were impartial. Table 8 shows the genotype results for the positive control of the eight SNVs across the three open OpenArray™ plates analysed, compared to the sequencing results obtained from Inqaba. The genotype results from the three OpenArray™ plates were in agreement with the sequencing results, with the exception of *SLC6A4*. There are no commercially designed probes and primers available for SLC6A4, the primers and probes were custom designed.

From the primer design information provided in Section 2.2.10 and Table 7, it can be seen that four different primer design versions were attempted to obtain sequencing results for *SLC6A4*, without success. The *SLC6A4* genotype remained undetermined through sequencing, kindly refer to Appendix B1 & B2, the FASTA and BLAST sequences. Although sequencing is considered the 'gold standard' for genotyping, this has highlighted a potential limitation to sequencing, where SNV genotype using the TaqMan® OpenArray™ assay may potentially be more beneficial. Result given for the failure of sequencing was the rich GC content (content greater that 65%) of the region. The BLAST sequence revealed a highly GC rich gene region making primer design

difficult due to none specific binding (Appendix B2). GC rich regions of DNA cause secondary structure formation, which can inhibit the denaturation, annealing and extension steps during PCR, which results in inefficient DNA sequencing[155]. Although the results for SCL6A4 could not be validated via sequencing, the consistency of the results across the three OpenArray™ plates analysed (both in terms of the control sample and the repeat samples), showed uniformity which provided reasonable confidence to include the results for further analysis. The sequencing certificate and full sequencing report from Inqaba are attached in Appendix B3. The full list of genotype results for all the samples analysed, are included in Appendix B4.

Table 8: Comparative display of genotype results for the positive control.

		GENOTYPE RESULTS			
Gene Name	**NCBI Ref.**	**Sequencing**	**AIE28***	**AIE29****	**AIE30*****
COMT	rs4680	A/G	A/G	A/G	A/G
DRD2	rs1800497	G/G	G/G	G/G	G/G
DRD4	rs1800955	C/C	C/C	C/C	C/C
GABRA6	rs3219151	C/T	C/T	C/T	C/T
HTR2C	rs3813929	C/C	C/C	C/C	C/C
LEPR	rs1137101	A/G	A/G	A/G	A/G
OPRM1	rs1799971	A/A	A/A	A/A	A/A
SLC6A4	rs25531	Undetermined	A/A	A/A	A/A

*AIE28 – Barcode of OpenArray® plate number 1
** AIE29 – Barcode of OpenArray® plate number 2
*** AIE30 – Barcode of OpenArray® plate number 3

A total of 39 (17.5%) out of the 223 genotyped samples were analysed in duplicate, to ensure consistency of the results. All samples were in concurrence with the genotype results obtained from the initial OpenArray™ plates analysed. These results can be found in Appendix B5.

3.2 TaqMan Genotyper Analysis and Interpretation

The initial analysis was performed using the QuantStudio™ 12K Flex Software v1.2.2. The QuantStudio™ software gives an indication of the quality of the run, but was likewise used as a double-checking mechanism to observe the sample amplification of the probes for each individual sample across the eight SNVs.

The QuantStudio™ 12K Flex has both qualitative and quantitative capabilities. Real-Time PCR enables a user to track the amount of fluorescence during the 50 cycles. Only the qualitative aspect of the TaqMan® genotyping assay, using the OpenArray™ system is reported to measure the genotype. Quantitative analysis is more relevant to the gene expression OpenArray™, where the level and amount of gene expression is of importance. The amount of fluorescence of the VIC® and FAM® dye is important to the study, in determining whether both dyes fluoresce equally (indicating heterozygosity) or whether either only the VIC® or only the FAM® dye fluoresced (indication homozygosity).

The first step was to export and analyse the QC images. The QC images give an indication of load quality as well as the presence of any potential fluorescent inhibitors, such as bubbles. The QC images in Figure 21A of Plate AIE30, shows a loading error in the centre of the plates, where approximately eight holes were miss-loaded. No amplification of the sample for these eight probes were observed. This was most likely due to insufficient master-mix, or a bubble that was taken up during the loading process. Figure 21B, shows post-amplification and the fluorescent for the VIC probe. There is no amplification in the holes that were missed. Samples that showed little to no amplification during analysis displayed no fluorescence in the QC images.

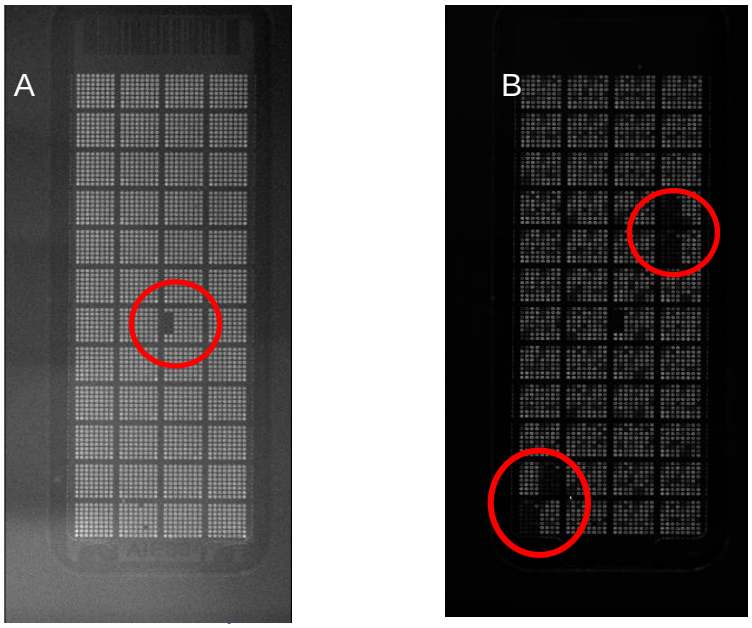


Figure 21: QC images exported from the QuantStudio software. A) QC image on load quality showing the miss loading of eight holes. B) Samples that do not amplify, shows no fluorescent in post-amplification QC images.

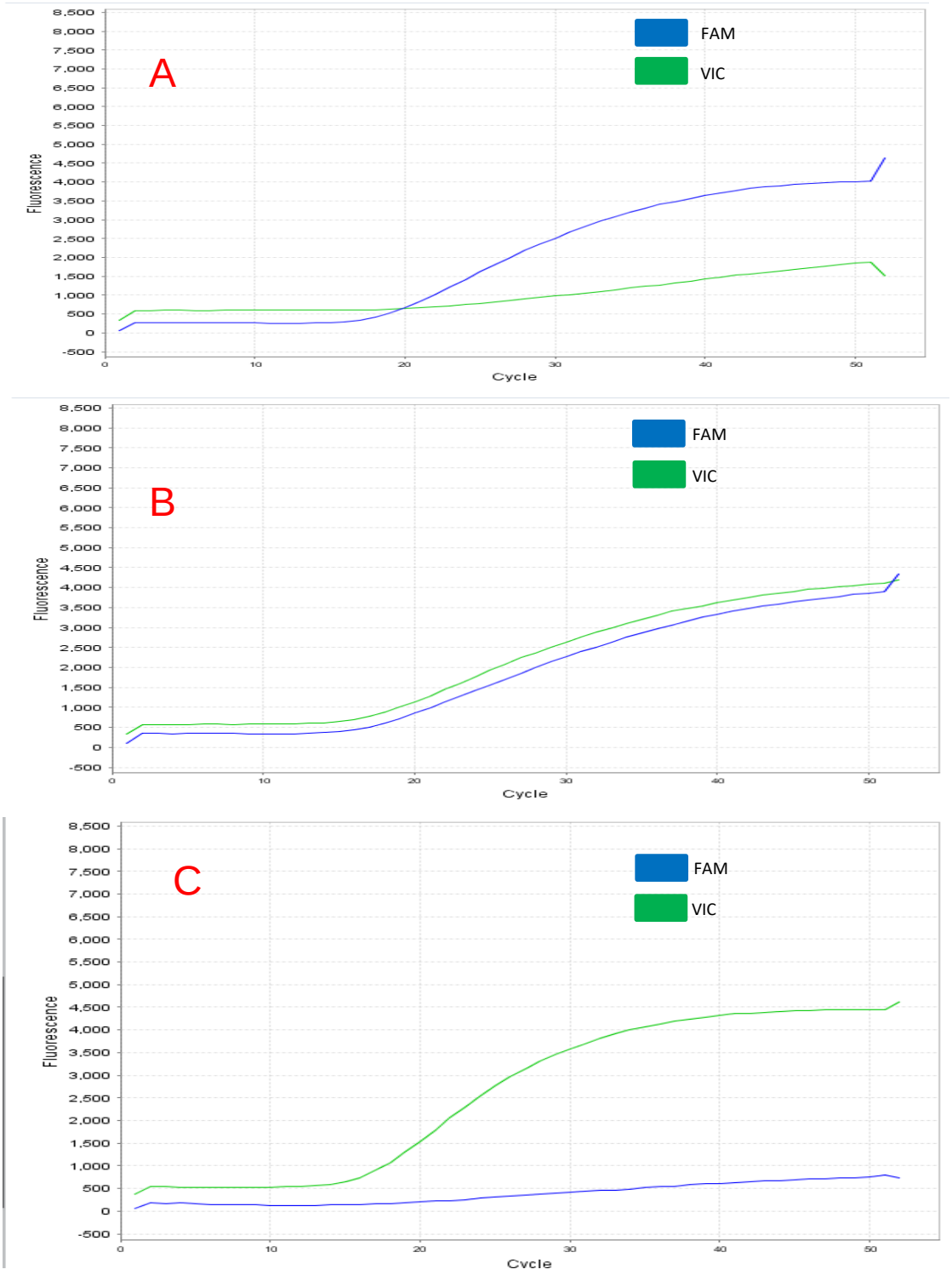


Figure 22: Multicomponent Plot from the QuantStudio™ 12K Flex Software. A) Homozygote for the FAM allele. B) Heterozygote for the FAM and VIC allele. C) Homozygote for the VIC allele.

When analysing the genotype results, it is important to check the multicomponent plot. The multicomponent plot shows which probe/s (VIC/FAM) fluoresce during the reaction. If only the FAM probe fluoresced, then only the blue line (Figure 22A), indicating a homozygote for the FAM allele. If only the VIC probe fluoresced, the green line (Figure 22C), indicating a homozygote for the VIC allele. If both the VIC and FAM probe fluoresced, increased intensity of both the green and blue lines (Figure 22B), indicating a heterozygote for the VIC/FAM alleles.

The second part of the analysis was done using the TaqMan® Genotyper Software version 1.3. The program groups the samples into 3 regions based on the fluorescence intensityof the probe, (Figure 23). The blue region are all the homozygotes for the FAM allele, the red region is the homozygote for the VIC allele and the green area is the heterozygotes (FAM and VIC alleles). The results showed good clustering of samples with minimal outliers. The blue dot at the bottom right is the negative control, showing no amplification.

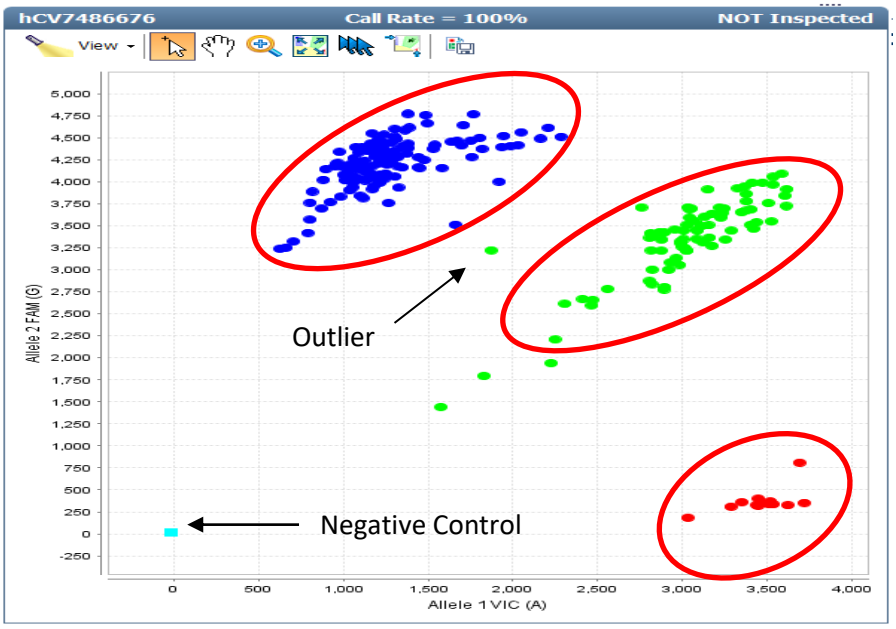


Figure 23: TaqMan® Genotyper results for DRD2 (rs1800497).

The cluster for SCL6A4 was not as clear due to the fact that there were only two genotypes present, AG and GG. No AA genotypes were observed in our study population, (Figure 24). Figures 22 to 30 are the graphical representations of each SNV and the clustering of the three genotypes across the eight SNVs. Any outliers

were double checked in the QuantStudio Program, by looking at the multiple component plot.

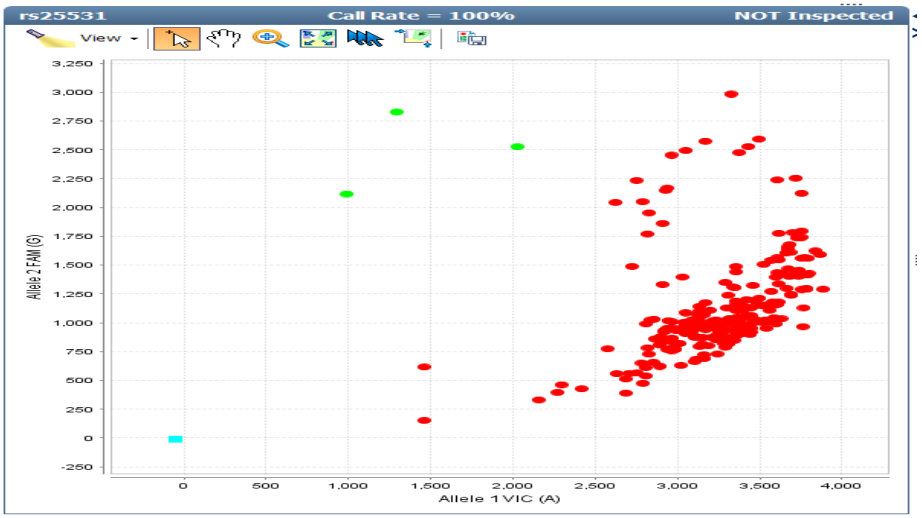


Figure 24: TaqMan® Genotyper results for SLC6A4 (rs25531).

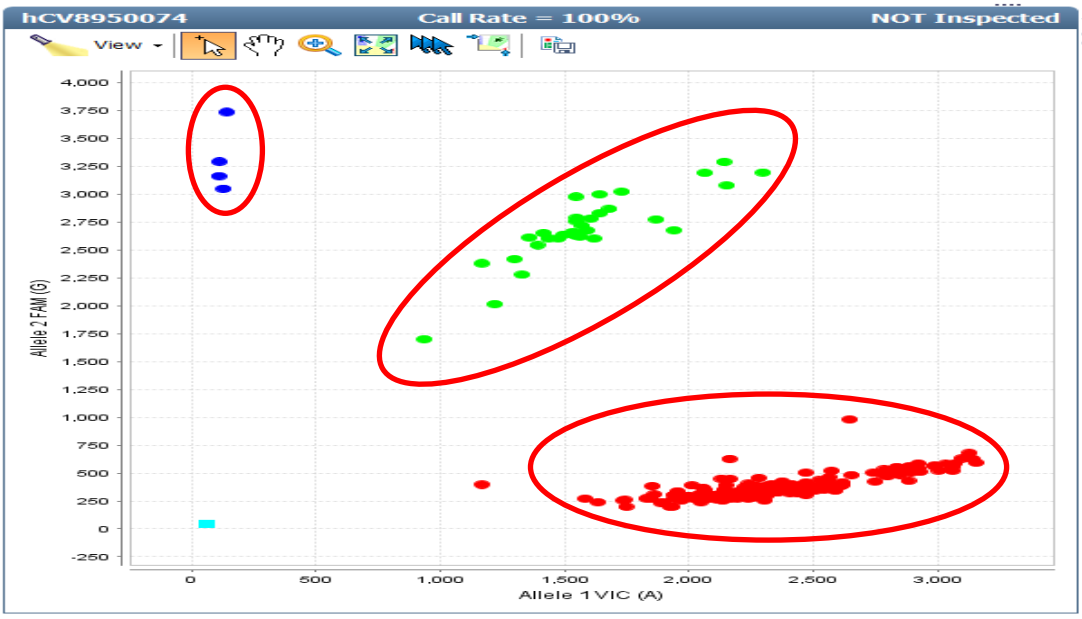


Figure 25: TaqMan® Genotyper results for OPRM1 (rs1799971).

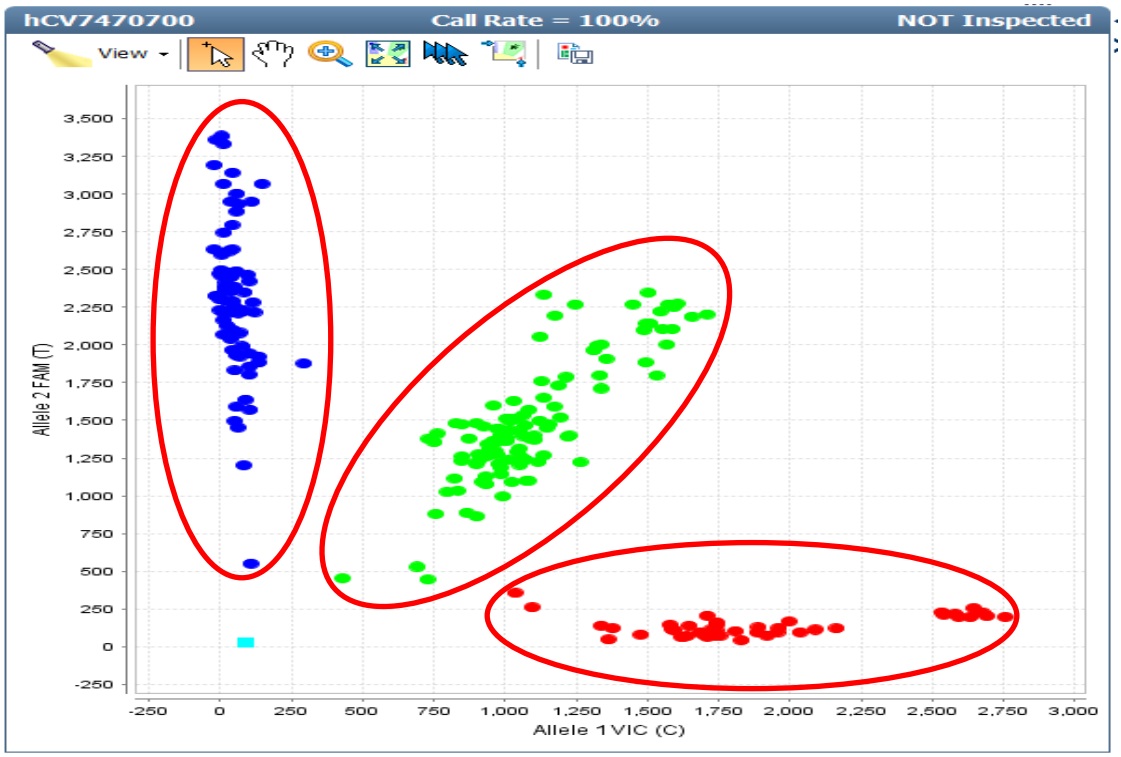


Figure 26: TaqMan® Genotyper results for DRD4 (rs1800955).

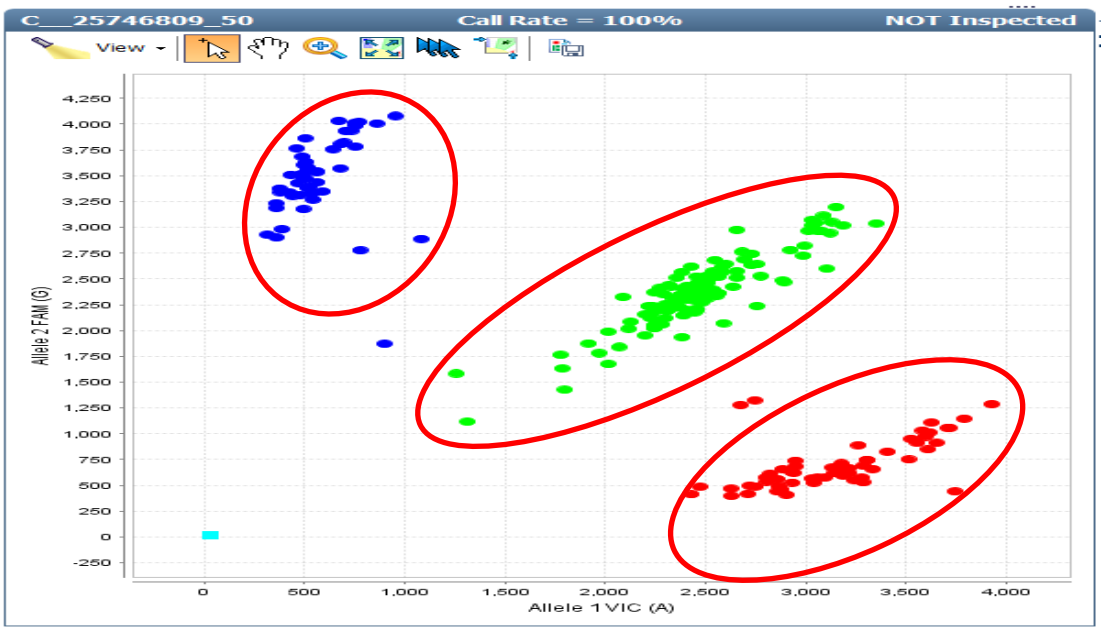


Figure 27: TaqMan® Genotyper results for COMT (rs4680).

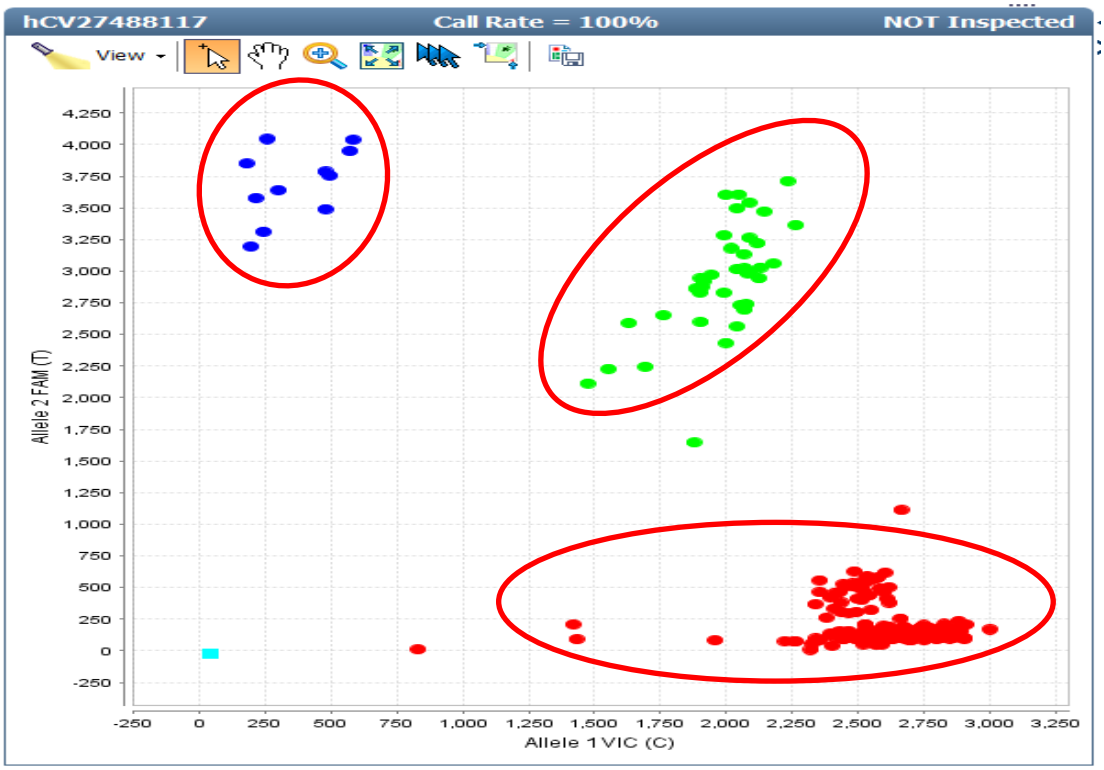


Figure 28: TaqMan® Genotyper results for HTR2C (rs3813929).

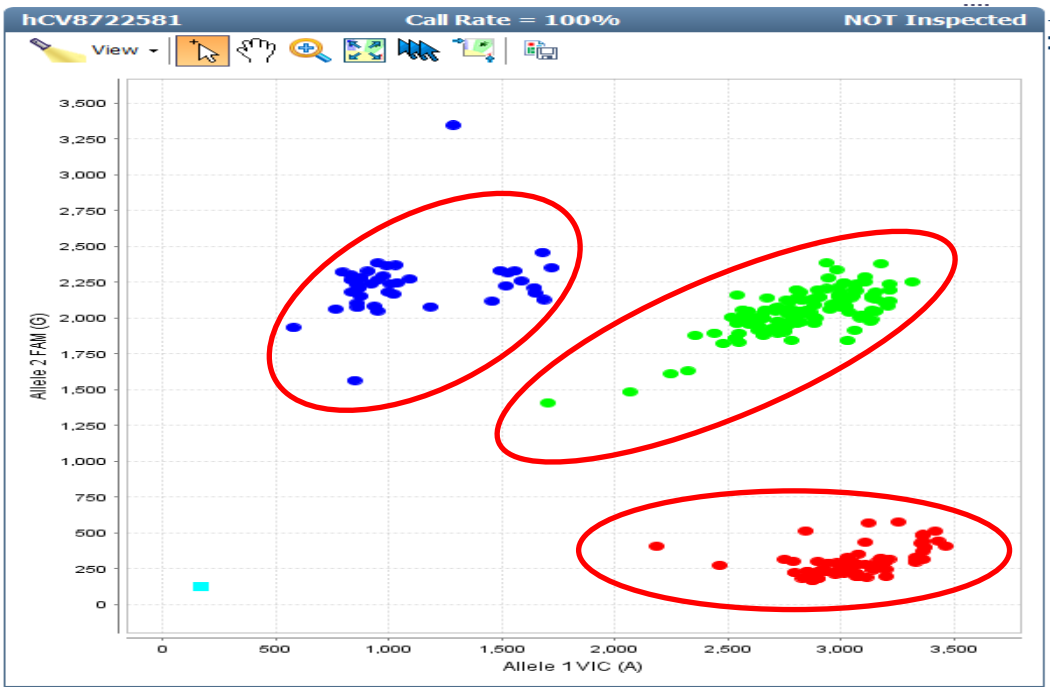


Figure 29: TaqMan® Genotyper results for LEPR (rs1137101).

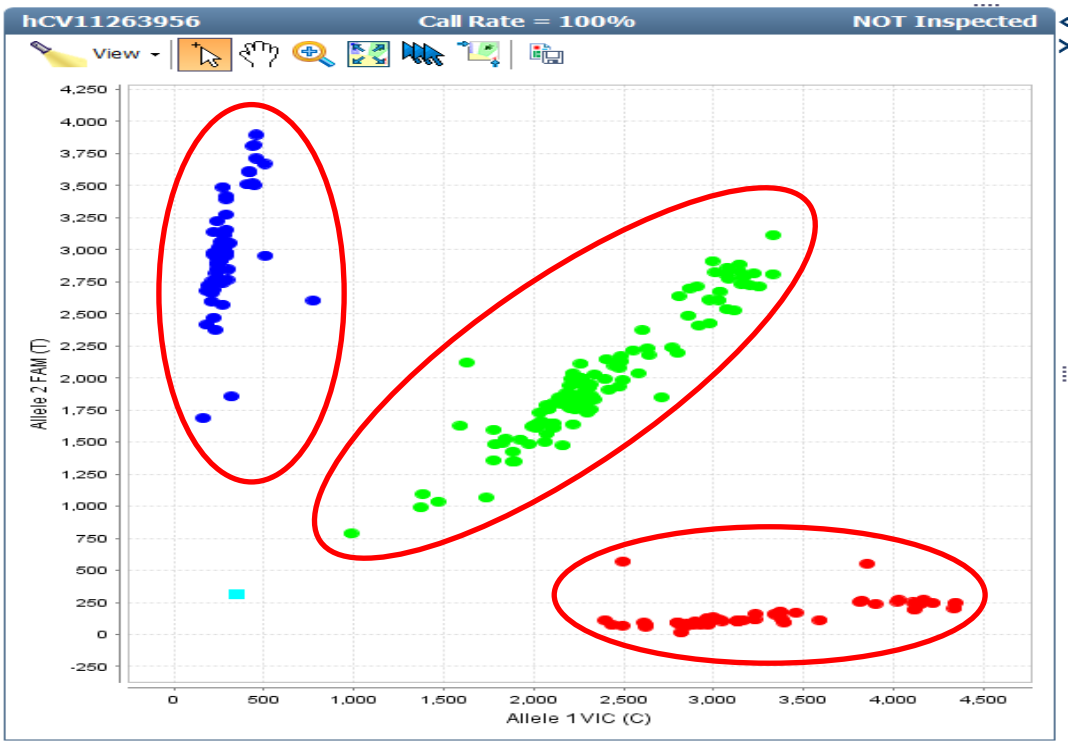


Figure 30: TaqMan® Genotyper results for GABRA6 (rs3219151).

3.3 Sample collection statistics

A total of 247 swabs were collected from the two sampling sites. Three samples were excluded from the study due to the participant's BMI score being below 18.5 kg/m^2. A further 21 samples were excluded from the study due to DNA extracted being of low concentration and poor quality, the sample failed to amplify, or results for all 8 SNVs were not obtained. Samples from Site A had an overall lower DNA concentration across the samples. For statistical purposes only 223 samples were included in the data for statistical analysis. One hundred and seven samples were collected from site A (family clinic) and 116 from site B (slimming clinic) (Table 9).

Table 9: Frequency distribution of samples collected from the two sampling sites

Site	Frequency	Percentage
A (Family Clinic)	107	47,98
B (Slimming Clinic)	116	52,02
Total	**223**	**100**

As per the initial study protocol a total of 194 samples were required, 97 normal weight individuals and 97 overweight and obese individuals. Of the 223 DNA samples collected and included in the statistical analysis, 107 were of normal weight and 116 were either overweight or obese, Table 10.

Table 10: Frequency distribution of the samples collected from the two strata

BMI Classification	Frequency	Percentage
Normal (Group 1)	107	47,98
Overweight & Class I – III (Group 2)	116	52,02
Total	**223**	**100**

The BMI distribution of the 223 samples into the individual classes according to the WHO's classification: 107 were of normal weight, 65 were overweight, 27 were classified as class I obesity, 13 as class II obesity and 11 as class III obesity. This data is summarised in Table 11 below.

Table 11: BMI Classification and distribution of samples collected

BMI Classification	Frequency	Percentage
Normal	107	47,98
Overweight	65	29,15
Class I	27	12,11
Class II	13	5,83
Class II	11	4,93
Total	**223**	**100**

Participants of the study were predominantly female. There were 157 female participant, and only 66 male participants, Table 12. Although sampling was conducted at two different sites, the presence of females was higher for both sites. Given the 2014 statistic from the WHO, where 62.2% of females and 41% males were overweight and/or obese[4]. The ratio of females to males who are overweight and/or obese are in agreement with global data. Tyrell *et al.* suggested the elevated incidence of obesity and depression in females, is likely linked to the lower socioeconomical standing of females[156]. Women handle the effects of stress in a different manner to men. Also it is important to note that females, are more self conscious of their health and weight and are more likely to seek help for health related concerns, which could have also contributed to the gender distibutor[157].

Table 12: Gender frequency distribution

Gender	Frequency	Percentage
Male	66	29,6
Female	157	70,4
Total	**223**	**100**

Ethic group distribution of the samples collected were as follows: of the 223 samples, 184 were Caucasians/whites, 31 African, 5 Coloured and 3 Indian. The distribution summary can be found in Table 13. Table 14 shows a summary of all the samples collected and their categorical distribution according to gender, race and BMI score. The observed distribution is likely due to the location of the two sampling sites, the family clinic is located in Doornpoort Pretoria and the Slimming Clinic in Brooklyn, Pretoria. Also, people tend to visit family doctors or doctors of same ethnic groups.

Table 13: Race distribution of samples

Race	Frequency	Percentage
White	184	82,51
African	31	13,9
Coloured	5	2,24
Indian	3	1,35
Total	**223**	**100**

Table 14 shows an overall summary of the sample distribution in terms of race, gender and BMI classification. The aim was to obtain samples that mimicked each other as much as possible (gender, race and BMI class) but to an extend this would have affected the randomization of the sampling of the study.

Table 14: Overall Summary of sample distribution of race, gender and BMI class.

Males	66	Females	157
White	49	White	135
African	15	African	16
Indian	1	Indian	4
Coloured	1	Coloured	2
White (Normal weight)	9	White (Normal weight)	76
White (Overweight and Obese)	40	White (Overweight and Obese)	59
African (Normal weight)	12	African (Normal weight)	7
African (Overweight and Obese)	3	African (Overweight and Obese)	9
Indian (Normal weight)	0	Indian (Normal weight)	2
Indian (Overweight and Obese)	1	Indian (Overweight and Obese)	0
Coloured (Normal weight)	0	Coloured (Normal weight)	1
Coloured (Overweight and Obese)	1	Coloured (Overweight and Obese)	3

3.4 Genotype and Allele frequency

In order to determine the prevalence of the eight SNVs, the frequency of the genotypes (allele 1/allele 1, allele1/allele2, allele2/allele2) and alleles (allele 1 and allele2) were determine for each of the eight SNVs.

The heterozygote genotype A/G for *COMT* (rs4860) was the most prevalent in the study population, with an equal allele distribution between the major and minor alleles, reflected in Table 15.

Table 15: Genotype and Allele frequency of *COMT*

COMT (rs4860)

Genotype	Frequency	Percentage
G/G	48	21,52
A/G	118	52,91
A/G	57	25,56
Total	**223**	**100**

Allele	Frequency	Percentage
G	107	47,98
A	116	52,02
Total	**223**	**100**

For the *DRD2* gene (also known as *ANKK1*) the homozygote genotype for the major allele G was most prevalent (57.4%). Table 16 show the genotype and allele frequency of *DRD2* (rs18000497).

Table 16: Genotype and Allele frequency for *DRD2*

***DRD2* / *ANKK1* (rs18000497)**

Genotype	Frequency	Percentage
G/G	128	57,4
A/G	82	36,77
A/A	13	5,83
Total	**223**	**100**

Allele	Frequency	Percentage
G	169	75,79
A	54	24,21
Total	**223**	**100**

The heterozygote genotype for *DRD4* was the most prevalent (48.43%), with T allele (major allele) being more prevalent than the minor allele, C, Table 17 and Figure 37.

Table 17: Genotype and Allele frequency for *DRD4*

***DRD4* (rs1800955)**

Genotype	Frequency	Percentage
T/T	71	31,84
C/T	108	48,43
C/C	44	19,73
Total	**223**	**100**

Allele	Frequency	Percentage
T	125	56,05
C	98	43,95
Total	**223**	**100**

GARBA6 showed a similar distribution to *DRD4*, with the heterozygote genotype (C/T) being the most prevalent and T allele having a slightly higher frequency, Table 18.

Table 18: Genotype and Allele frequency for GARBA6.

GARBA6 (rs3219151)

Genotype	Frequency	Percentage
C/C	50	22,42
C/T	107	47,98
T/T	66	29,6
Total	**223**	**100**

Allele	Frequency	Percentage
C	103,5	46,41
T	119,5	53,59
Total	**223**	**100**

HTR2C showed a slightly different distribution with the homozygote genotype for the major allele occurring in 78% of the population, 86.55% was the allele frequency for the C allele, Table 19. It is also important to note that HTR2C gene is located on the X-chromosome, so males would have either fallen into the T/T genotype of the C/C genotype.

Table 19: Genotype and Allele frequency for *HTR2C*

***HTR2C* (rs3813929)**

Genotype	Frequency	Percentage
T/T	11	4,93
C/T	38	17,04
C/C	174	78,03
Total	**223**	**100**

Allele	Frequency	Percentage
T	30	13,45
C	193	86,55
Total	**223**	**100**

The major allele, A, was most prevalent for the *OPRM1* gene (90.81%), with the homozygote genotype for the major allele (A/A) being the most frequently observed (83%), Table 23.

Table 20: Genotype and Allele frequency for *OPRM1*

***OPRM1* (rs1799971)**

Genotype	Frequency	Percentage
A/A	186	83,41
A/G	33	14,8
G/G	4	1,79
Total	**223**	**100**

Allele	Frequency	Percentage
A	202,5	90,81
G	20,5	9,19
Total	**223**	**100**

The minor allele A was slightly more prevalent in the study population, with the heterozygote genotype A/G occurring most frequently (53%), compared to the GG (18.83%) and AA (28.25%) genotype, Table 21.

Table 21: Genotype and Allele frequency for *LEPR*

***LEPR* (rs1137101)**

Genotype	Frequency	Percentage
G/G	42	18,83
A/G	118	52,91
A/A	63	28,25
Total	**223**	**100**

Allele	Frequency	Percentage
G	101	45,29
A	122	54,71
Total	**223**	**100**

The minor allele G, for *SLC6A4* was only seen in 0.67% of the study population, while A, the major allele was dominant in 99.33% of the study population. There was also no GG genotype observed in the study population, Table 22. Although the allele frequencies are similar to those observed in the 1000 Genome project, these results could be validated by the positive control, as the positive control failed sequencing. Therefore, the results should be interpreted with caution.

Table 22: Genotype and Allele frequency for *SLC6A4*

***SLC6A4* (rs25531)**

Genotype	Frequency	Percentage
A/G	3	1,35
A/A	220	98,65
Total	**223**	**100**

Allele	Frequency	Percentage
G	1,5	0,67
A	221,5	99,33
Total	**223**	**100**

Table 23: Summary of Genotype Results between Caucasians and Africans.

		Caucasians						Africans					
		Normal (n=85)	Overweight (n=59)	Class I (n=21)	Class II (n=9)	Class III (n=10)	Wt_NnN* (n=99)	Normal (n=19)	Overweight (n=5)	Class I (n=4)	Class II (n=2)	Class III (n=1)	Wt_NnN (n=12)
COMT	A/A	27 (0,32)	14 (0,24)	4 (0,19)	3 (0,33)	4 (0,4)	25 (0,25)	8 (0,42)	0 (0)	0 (0)	2	0 (0)	2 (0,17)
	A/G	43 (0,51)	34 (0,58)	13 (0,62)	6 (0,66)	4 (0,4)	57 (0,58)	8 (0,42)	2 (0,4)	1 (0,25)	0 (0)	1 (1)	4 (0,33)
	G/G	15 (0,18)	11 (0,19)	4 (0,19)	0 (0)	2 (0,2)	17 (0,17)	3 (0,16)	3 (0,6)	3 (0,75)	0 (0)	0 (0)	6 (0,5)
DRD2	A/A	2 (0,02)	4 (0,07)	0 (0)	1 (0,11)	0 (0)	5 (0,05)	6 (0,32)	1 (0,2)	1 (0,25)	0 (0)	0 (0)	2 (0,17)
	A/G	30 (0,35)	19 (0,32)	6 (0,29)	5 (0,55)	3 (0,3)	33 (0,33)	11 (0,58)	3 (0,6)	2 (0,5)	1 (0,5)	1 (1)	7 (0,58)
	G/G	53 (0,62)	36 (0,61)	15 (0,71)	3 (0,33)	7 (0,7)	61 (0,62)	2 (0,11)	1 (0,2)	1 (0,25)	1 (0,5)	0 (0)	3 (0,25)
DRD4	C/C	19 (0,22)	8 (0,14)	6 (0,29)	2 (0,22)	2 (0,2)	18 (0,18)	2 (0,11)	1 (0,2)	1 (0,25)	1 (0,5)	0 (0)	3 (0,25)
	C/T	46 (0,54)	32 (0,54)	9 (0,43)	3 (0,33)	3 (0,3)	47 (0,47)	7 (0,37)	2 (0,4)	1 (0,25)	1 (0,5)	1 (1)	5 (0,42)
	T/T	20 (0,24)	19 (0,32)	6 (0,29)	4 (0,44)	5 (0,5)	34 (0,34)	10 (0,53)	2 (0,4)	2 (0,5)	0 (0)	0 (0)	4 (0,33)
GABRA6	C/C	15 (0,18)	15 (0,25)	4 (0,19)	1 (0,11)	3 (0,3)	23 (0,23)	5 (0,26)	0 (0)	2 (0,5)	0 (0)	1 (1)	3 (0,25)
	C/T	45 (0,53)	16 (0,27)	9 (0,43)	7 (0,77)	5 (0,5)	37 (0,37)	8 (0,42)	2 (0,4)	1 (0,25)	2 (1)	0 (0)	5 (0,42)
	T/T	25 (0,29)	20 (0,34)	8 (0,38)	1 (0,11)	2 (0,2)	31 (0,31)	6 (0,32)	3 (0,6)	1 (0.25)	0 (0)	0 (0)	4 (0,33)
HTR2C	C/C	60 (0,71)	46 (0,78)	16 (0,76)	8 (0,88)	9 (0,9)	79 (0,80)	19 (1)	4 (0,8)	4 (1)	2 (1)	1 (1)	11 (0,92)
	C/T	20 (0,24)	10 (0,17)	4 (0,19)	1 (0,11)	0 (0)	15 (0,15)	0 (0)	1 (0,2)	0 (0)	0 (0)	0 (0)	1 (0,08)
	T/T	5 (0,06)	3 (0,05)	1 (0,05)	0 (0)	1 (0,1)	5 (0,05)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
LEPR	A/A	35 (0,41)	13 (0,22)	4 (0,19)	0 (0)	2 (0,2)	19 (0,19)	8 (0,42)	1 (0,2)	2 (0,5)	1 (0,5)	0 (0)	4 (0,33)
	A/G	33 (0,39)	37 (0,63)	14 (0,67)	9 (1)	7 (0,7)	67 (0,68)	10 (0,53)	3 (0,6)	2 (0,5)	1 (0,5)	1 (1)	7 (0,58)
	G/G	17 (0,2)	9 (0,15)	3 (0,14)	0 (0)	1 (0,1)	13 (0,13)	1 (0,05)	1 (0,2)	0 (0)	0 (0)	0 (0)	1 (0,08)
OPRM1	A/A	67 (0,79)	45 (0,76)	21 (1)	6 (0,66)	10 (1)	82 (0,83)	19 (1)	5 (1)	4 (1)	2 (1)	1 (1)	12 (1)
	A/G	15 (0,18)	14 (0,24)	0 (0)	3 (0,33)	0 (0)	17 (0,17)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	G/G	3 (0,04)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
SLC6A4	A/A	85 (1)	59 (1)	21 (1)	9 (1)	10 (1)	99 (1)	18 (0,95)	5 (1)	3 (0,75)	2 (1)	1 (1)	12 (1)
	A/G	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0,05)	0 (0)	1 (0,25)	0 (0)	0 (0)	1 (0.083)

Wt_NnN – Weight Not Normal (Overweight and Obese Class 1-3)
n – sample size

3.5 Overview of results

Due to the small sample numbers for Coloureds and Indians (8/223), no general (population) characteristics can be made for these two populations. Further analysis will focus mainly on the Caucasian and African race groups.

From Table 23, the following can be deduced:

1. *DRD4* – the TT genotype was found predominantly in the overweight and obese Caucasian population. While the TT genotype occurred in 53% of the normal weight African population.
2. *LEPR* – the AA genotype was dominate in the normal weight Caucasian population (41%), while the AG genotype was more prominent in the overweight and obese Caucasian population. The seemed to be an equal distribution of genotypes among the normal and not normal weight (Wt_NnN) African population.
3. *OPRM1* – the GG genotype was only observed in normal weight Caucasian population. The AG and GG genotypes were not observed in any of the African participants.
4. *COMT* – AA genotype was predominant in the normal weight African population, while the GG genotype was dominant in overweight and obese African population.
5. *HTR2C* – the TT genotype was not observed in the African population. Only one participant of the African population has the C/T genotype.

The other three SNVs namely, *DRD2, GABRA6* and *SLC6A4*, did not show any significant differences or any association between the genotypes and categorial divisions.

3.6 Allele and Genotype frequency comparison to previous studies

The 1000 Genomes Project was an international research effort to establish the most detailed public catalogue of human genetic variations. The aim was to sequence the genome of one thousand participants from various ethnic groups in three years. In the

final phase of the project, data from 2504 samples were combined and is to date the largest collection of genotypes of variant sites the project revealed. The study compared the allele frequency observed in the study population with that documented in the 1000 Genomes Project in terms of the African and European populations recorded (Table 24). The European population (EUR) consisted of Toscani (Italy), Finnish (Finland), British (England and Scotland) and Iberian (Spain). The African (AFR) population consisted of Yoruba (Nigeria), Wuhya (Kenya), Gambian (Gambia), Mende (Sierre Leone) and Esan (Nigeria). The Americans of African Ancestor (USA) and African Caribbean's (Barbados)[158]. Figure 31, shows the geographical display of the population samples.

The allele frequencies observed in the study population of Caucasians were very similar to the allele frequencies reported in the European population from the 1000 Genomes Project. The results from the African population in the study group was not in complete agreement with the African population from the 1000 Genomes Project. SNV *DRD4* (rs1800955), *HTRC2* (rs3813929) and *OPRM1* (rs1799971) show similar allele frequencies between the African Populations.

Figure 31: Geographical display of the populations sampled for the 1000 Genomes Project[158].

Table 24: Observed Allele Frequency

		In house Caucasian Participants	1000 Genomes Project (EUR)	In house African Participants*	1000 Genomes Project (AFR)
COMT	**rs4680**				
Major Allele	G	0.554	0.5000	0.484	0.7194
Minor Allele	A	0.446	0.5000	0.516	0.2806
DRD2	**rs1800497**				
Major Allele	G	0.791	0.8121	0.452	0.615
Minor Allele	A	0.209	0.1879	0.548	0.385
DRD4	**rs1800955**				
Major Allele	T	0.546	0.5885	0.645	0.592
Minor Allele	C	0.454	0.4115	0.355	0.4077
GABRA6	**rs3219151**				
Major Allele	C	0.473	0.4235	0.468	0.5666
Minor Allele	T	0.527	0.5765	0.532	0.4334
HTRC2	**rs3813929**				
Major Allele	C	0.851	0.8807	0.984	0.99089
Minor Allele	T	0.149	0.1193	0.016	0.0091
LEPR	**rs1137101**				
Major Allele	G	0.435	0.4692	0.339	0.5923
Minor Allele	A	0.565	0.5308	0.661	0.4077
OPRM1	**rs1799971**				
Major Allele	A	0.897	0.838	1	0.99089
Minor Allele	G	0.103	0.162	0	0.0091
SLC6A4	**rs25531**				
Major Allele	A	1	0.9105	0.968	0.7799
Minor Allele	G	0	0.0895	0.032	0.2201

AFR = African Population
EUR = European Population
* Important to note that the population size of the African ethnic group in the study was limited

As a second comparison the International HapMap Project was used, which aimed to develop a haplotype map (HapMap) of the human genome to describe the common patterns of human genetic variations. The International HapMap project was a collaboration among researchers from academic research centres, non-profit research groups and private companies from Canada, China, Japan, Nigeria, the United Kingdom and the United States. It started in October 2002 and ran until 2009, when the third phase of data was released. Populations studied included Utah residents with northern and western European ancestry, Han Chinese from Beijing, Japanese from

Tokyo, Yoruba residents from Nigeria, African ancestry from Southwest USA, Chinese from United States, Gujarati Indians from United States, Luhya residents from Kenya, Massai from Kenya, Mexican ancestry in the United States and Toscani residents from Italy[159].

The allele frequency observed in the Caucasian population in this study was similar to the European population from the HapMap project, while African study population showed different allele frequencies to the HapMap project, with the exception of *OPRM1* and *HTRC2*, as indicated in Table 25.

Table 25: International HapMap Project comparison with observed allele frequency

		In house Caucasian Participants	HapMap (CEU)	In house African Participants	HapMap (YRI)
COMT	**rs4680**				
Major Allele	G	0.554	0.5221	0.484	0,6858
Minor Allele	A	0.446	0.4779	0.516	0,3142
DRD2	**rs1800497**				
Major Allele	G	0.791	0.8053	0.452	0,5893
Minor Allele	A	0.209	0.1947	0.548	0,4107
DRD4	**rs1800955**				
Major Allele	T	0.546	N/A	0.645	N/A
Minor Allele	C	0.454	N/A	0.355	N/A
GABRA6	**rs3219151**				
Major Allele	C	0.473	0.3628	0.468	0,5398
Minor Allele	T	0.527	0.6372	0.532	0,4602
HTRC2	**rs3813929**				
Major Allele	C	0.851	0.8097	0.984	1
Minor Allele	T	0.149	0.1903	0.016	0
LEPR	**rs1137101**				
Major Allele	G	0.435	0.4732	0.339	0,5972
Minor Allele	A	0.565	0.5267	0.661	0,4028
OPRM1	**rs1799971**				
Major Allele	A	0.897	0.8451	1	1
Minor Allele	G	0.103	0.1549	0	0
SLC6A4	**rs25531**				
Major Allele	A	1	N/A	0.968	N/A
Minor Allele	G	0	N/A	0.032	N/A

YRI = Sub-Saharan African ancestry
CEU = European ancestry
N/A = Not applicable – no data available

The HapMap Project and the 1000 Genomes Project shared a small portion of samples.

3.7 Pooling of statistical data

The data was analysed individually using the various subgroups determined by the BMI score namely; normal weight, overweight, class I, class II and class III obesity. The data was also analysed using the simple pooling method, meaning data from the various categories or BMI class was combined without being weighted.

Pool 1: BMI Pool (Cl I-III) – observations from BMI Class I, Class II and Class III were combined.

Pool 2: Wt_NnN – observations from overweight, BMI Class I, Class II and Class III were combined.

The aim was to create a high precision estimate and increase statistical power of the small sample size and number of individuals sampled in the individual categories.

No variation across the categorical variables in the study population is designated the null hypothesis.

3.8 Fischer's Exact Test

Fisher's Exact test is a test of significance used in the place of chi-square test in the case of categorical data and small sample size. It tests the probability of getting a table that is strongly due to the chance of sampling. The Fisher Exact test uses a formula to obtain the probability (p-value) of the combination of the frequencies that are actually obtained. It also involves the finding of the probability of every possible combination which indicates more evidence of association. The Fischer Exact test is based on a couple of assumptions, namely:

- The samples have been drawn from the population randomly.
- Directional hypothesis – type of hypothesis which predicts either a positive association or a negative association.
- The value of the first person does not affect the value of the second person.

- Mutual exclusivity - probability theory, two events are mutually exclusive or disjoint if they cannot both occur.

The p-value is commonly defined as the probability of obtaining a result that is at least as extreme as the one observed, assuming the null hypothesis is true, thus p-values that are ≤0.05 are considered to be significant, where p-values that are >0.05 are not significant, using a 95% confidence interval [160].

The Fisher Exact test was performed on the natural data set as observed and two pooled data sets. The first pooled data set combined BMI Class I, II and III. The second pooled data set, combined all samples that were not of normal weight; overweight, class I, class II and class II.

Below are the tabulated results with the p-values of the Fisher Exact test for each SNV and the various pooled data sets. Results of the Fisher Exact tests, only *LEPR* showed significant variation (≥0.05) across all three Fisher Exact tests (Tables 36 and 37).

Table 26: Fisher Exact Test for significance with unpooled data, *COMT* (rs4680)

***COMT* (rs4680)**

	Normal	Overweight	Class I	Class II	Class III	Total
G/G	23	14	8	1	2	48
	47,92	29,17	16,67	2,08	4,17	100
	21,5	21,54	29,63	7,69	18,18	21,52
A/G	54	37	15	7	5	118
	45,76	31,36	12,71	5,93	4,24	100
	50,47	56,92	55,56	53,85	45,45	52,91
A/A	30	14	4	5	4	57
	52,63	24,56	7,02	8,77	7,02	100
	28,04	21,54	14,81	38,46	36,36	25,56
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,680 – Not significant

Table 27: Fisher Exact Test for significance with pooled data, *COMT* (rs4680).

***COMT* (rs4860)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
G/G	23	14	11	48
	47,92	29,17	22,92	100
	21,5	21,54	21,57	21,52
A/G	54	37	27	118
	45,76	31,36	22,88	100
	50,47	56,92	52,94	52,91
A/A	30	14	13	57
	52,63	24,56	22,81	100
	28,04	21,54	25,49	25,56
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,915 – Not significant

	wt_NnN		
Genotype	Normal	Not_Normal	Total
G/G	23	25	48
	47,92	52,08	100
	21,5	21,55	21,52
A/G	54	64	118
	45,76	54,24	100
	50,47	55,17	52,91
A/A	30	27	57
	52,63	47,37	100
	28,04	23,28	25,56
Total	107	116	223
	47,98	29,15	100
	100	100	100

Fisher's Exact = 0,694 – Not significant

Table 28: Fisher Exact Test for significance with unpooled data, *DRD2* (rs18000497)

***DRD2* (rs18000497)**

	Normal	Overweight	Class I	Class II	Class III	Total
G/G	60	38	17	6	7	128
	46,88	29,69	13,28	4,69	5,47	100
	56,07	58,46	62,96	46,15	63,64	57,4
A/G	41	22	9	6	4	82
	50	26,83	10,98	7,32	4,88	100
	38,32	33,85	33,33	46,15	36,36	36,77
A/A	6	5	1	1	0	13
	46,15	38,46	7,69	7,69	0	100
	5,61	7,69	3,7	7,69	0	25,56
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,975 – Not significant

Table 29: Fisher Exact Test for significance with pooled data, *DRD2* (rs18000497)

***DRD2* (rs18000497)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
G/G	60	38	30	128
	46,88	29,69	23,44	100
	56,07	58,46	58,82	57,4
A/G	41	22	19	82
	50	26,83	23,17	100
	38,32	33,85	37,25	36,77
A/A	6	5	2	13
	46,15	38,46	15,38	100
	5,61	7,69	3,92	5,83
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,923 – Not significant

	wt_NnN		
Genotype	Normal	Not_Normal	Total
G/G	60	68	128
	46,88	53,13	100
	56,07	58,62	57,4
A/G	41	41	82
	50	50	100
	38,32	35,34	36,77
A/A	6	7	13
	46,15	53,85	100
	5,61	6,03	5,83
Total	107	116	223
	47,98	52,02	100
	100	100	100

Fisher's Exact = 0,904 – Not significant

Table 30: Fisher Exact Test for significance with unpooled data, *DRD4* (rs1800955)

***DRD4* (rs1800955)**

	Normal	Overweight	Class I	Class II	Class III	Total
T/T	32	21	9	4	5	71
	45,07	29,58	12,68	5,63	7,04	100
	29,91	32,31	33,33	30,77	45,45	31,84
C/T	54	34	10	6	4	108
	50	31,48	9,26	5,56	3,7	100
	50,47	52,31	37,04	46,15	36,36	48,43
C/C	21	10	8	3	2	44
	47,73	22,73	18,18	6,82	4,55	100
	19,63	15,38	29,63	23,08	18,18	19,73
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,822 – Not significant

Table 31: Fisher Exact Test for significance with pooled data, *DRD4* (rs1800955)

***DRD4* (rs1800955)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
T/T	32	21	18	71
	45,07	29,58	25,35	100
	29,91	32,31	35,29	31,84
C/T	54	34	20	108
	50	31,48	18,52	100
	50,47	52,31	39,22	48,43
C/C	21	10	13	44
	47,73	22,73	29,55	100
	19,63	15,38	25,46	19,73
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,554 – Not significant

	wt_NnN		
Genotype	Normal	Not_Normal	Total
T/T	32	39	71
	45,07	54,93	100
	29,91	33,62	31,84
C/T	54	54	108
	50	50	100
	50,47	46,55	48,43
C/C	21	23	44
	47,73	52,27	100
	19,63	19,83	19,73
Total	107	116	223
	47,98	52,02	100
	100	100	100

Fisher's Exact = 0,810 – Not significant

Table 32: Fisher Exact Test for significance with unpooled data, *GARBA6* (rs3219151)

***GARBA6* (rs3219151)**

	Normal	Overweight	Class I	Class II	Class III	Total
C/C	21	15	8	2	4	50
	42	30	16	4	8	100
	19,63	23,08	29,63	15,38	36,36	22,42
C/T	55	27	10	10	5	107
	51,4	25,23	9,35	9,35	4,67	100
	51,4	41,54	37,04	76,92	45,45	47,98
T/T	31	23	9	1	2	66
	46,97	34,85	13,64	1,52	3,03	100
	28,97	35,38	33,33	7,69	18,18	29,6
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,309 – Not significant

Table 33: Fisher Exact Test for significance with pooled data, *GARBA6* (rs3219151)

***GARBA6* (rs3219151)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
C/C	21	15	14	50
	42	30	28	100
	19,63	23,08	27,45	22,42
C/T	55	27	25	107
	51,4	25,23	23,36	100
	51,4	41,54	49,02	47,98
T/T	31	23	12	66
	46,97	34,85	18,18	29,6
	28,97	35,38	23,53	5,83
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,528 – Not significant

	wt_NnN		
Genotype	Normal	Not_Normal	Total
C/C	21	29	50
	42	58	100
	19,63	25	22,42
C/T	55	52	107
	51,4	48,6	100
	51,4	44,83	47,98
T/T	31	35	66
	46,97	53,03	29,6
	28,97	30,17	5,83
Total	107	116	223
	47,98	52,02	100
	100	100	100

Fisher's Exact = 0,538 – Not significant

Table 34: Fisher Exact Test for significance with unpooled data, *HTR2C* (rs3813929)

***HTR2C* (rs3813929)**

	Normal	Overweight	Class I	Class II	Class III	Total
T/T	6	3	1	0	1	11
	54,55	27,27	9,09	0	9,09	100
	5,61	4,62	3,7	0	9,09	4,93
C/T	21	11	5	1	0	38
	55,26	28,95	13,16	2,63	0	100
	19,63	16,92	18,52	7,69	0	17,04
C/C	80	51	21	12	10	174
	45,98	29,31	12,07	6,9	5,75	100
	74,77	78,46	77,78	92,31	90,91	78,03
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,804 – Not significant

Table 35: Fisher Exact Test for significance with pooled data, *HTR2C* (rs3813929)

***HTR2C* (rs3813929)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
T/T	6	3	2	11
	54,55	27,27	18,18	100
	5,61	4,62	3,92	4,93
C/T	21	11	6	38
	55,26	28,95	15,79	100
	19,63	16,92	11,76	17,04
C/C	80	51	43	174
	45,98	29,31	24,71	100
	74,77	78,46	84,31	78,03
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,787 – Not significant

	wt_NnN		
Genotype	Normal	Not_Normal	Total
T/T	6	5	11
	54,55	45,45	100
	5,61	4,31	4,93
C/T	21	17	38
	55,26	44,74	100
	19,63	14,66	17,04
C/C	80	94	174
	45,98	54,02	100
	74,77	81,03	78,03
Total	107	116	223
	47,98	52,02	100
	100	100	100

Fisher's Exact = 0,538 – Not significant

Table 36: Fisher Exact Test for significance with unpooled data, *LEPR* (rs1137101)

***LEPR* (rs1137101)**

	Normal	Overweight	Class I	Class II	Class III	Total
G/G	27	10	3	1	1	42
	64,29	23,81	7,14	2,38	2,38	100
	25,23	15,38	11,11	7,69	9,09	18,3
A/G	42	41	16	11	8	118
	35,59	34,75	13,56	9,32	6,78	100
	39,25	63,08	59,26	84,62	72,73	52,91
A/A	38	14	8	1	2	63
	60,32	22,22	12,7	1,59	3,17	100
	35,51	21,54	29,63	7,69	18,18	28,25
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,018 – SIGNIFICANT

The unpooled Fisher Exact Test for the *LEPR* gene, showed a p-value = 0.018, which is less than 0.05, thus the null hypothesis is rejected and the variation observed is not solely due to chance (Table 35). The GG and AA genotypes appeared to be more prominent in the normal weight BMI category than any of the other classes.

Table 37: Fisher Exact Test for significance with pooled data, *LEPR* (rs1137101)

***LEPR* (rs1137101)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
G/G	27	10	5	42
	64,29	23,81	11,9	100
	25,23	15,38	9,8	18,83
A/G	42	41	35	118
	35,59	34,75	29,66	100
	39,25	63,08	68,63	52,91
A/A	38	14	11	63
	60,32	22,22	17,46	100
	35,51	21,54	21,57	28,25
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,03
SIGNIFICANT

	wt_NnN		
Genotype	Normal	Not_Normal	Total
G/G	27	15	42
	64,29	35,71	100
	25,23	12,93	18,83
A/G	42	76	118
	35,59	64,41	100
	39,25	65,52	52,91
A/A	38	25	63
	60,32	39,68	100
	35,51	21,55	28,25
Total	107	116	223
	47,98	52,02	100
	100	100	100

Fisher's Exact = 0,00
SIGNIFICANT

The Fisher Exact test for *LEPR*, using the pooled data showed, produced similar results to the unpooled data set. In addition, it also found a significant association with the AG genotype, which was more prominent in overweight and obese individuals (Table 37). The p-value for pooled data set 1 was 0.03, while the p-value for the pooled data set 2 was 0. Thus, both values are less than 0.05 (95% confidence interval), once again rejecting the null hypothesis.

Table 38: Fisher Exact Test for significance with unpooled data, *OPRM1* (rs1799971)

***OPRM1* (rs1799971)**

	Normal	Overweight	Class I	Class II	Class III	Total
A/A	88	51	26	10	11	186
	47,31	27,42	13,98	5,38	5,91	100
	82,24	78,46	96,3	76,92	100	83,41
A/G	15	14	1	3	0	33
	45,45	42,42	3,03	9,09	0	100
	14,02	21,54	3,07	23,08	0	14,8
G/G	4	0	0	0	0	4
	100	0	0	0	0	100
	3,74	0	0	0	0	1,76
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,172 – Not significant

Table 39: Fisher Exact Test for significance with pooled data, *OPRM1* (rs1799971)

***OPRM1* (rs1799971)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
A/A	88	51	47	186
	47,31	27,42	25,27	100
	82,24	78,46	92,16	83,41
A/G	15	14	4	33
	45,45	42,42	12,12	100
	14,02	21,54	7,84	14,8
G/G	4	0	0	4
	100	0	0	100
	3,74	0	0	1,79
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,090 – Not significant

	wt_NnN		
Genotype	Normal	Not_Normal	Total
A/A	88	98	186
	47,31	52,69	100
	82,24	84,48	83,41
A/G	15	18	33
	45,45	54,55	100
	14,02	15,52	14,8
G/G	4	0	4
	100	0	100
	3,74	0	1,79
Total	107	116	223
	47,98	52,02	100
	100	100	100

Fisher's Exact = 0,134 – Not significant

Although the p-value for pooled data set 1 is 0.09, which is greater than 0.05, it should be noted that the GG genotype only appeared in the normal weight group of individuals.

Table 40: Fisher Exact Test for significance with unpooled data, *SLC6A4* (rs25531)

***SLC6A4* (rs25531)**

	Normal	Overweight	Class I	Class II	Class III	Total
A/G	2	0	1	0	0	3
	66,67	0	33,33	0	0	100
	1,87	0	3,7	0	0	1,35
A/A	105	65	26	13	11	220
	47,73	29,55	11,82	5,91	5	100
	98,13	100	96,3	100	100	98,65
Total	107	65	27	13	11	223
	47,98	29,15	12,11	5,83	4,93	100
	100	100	100	100	100	100

Fisher's Exact = 0,464 – Not significant

Table 41: Fisher Exact Test for significance with pooled data, *SLC6A4* (rs25531)

***SLC6A4* (rs25531)**

	BMI Pool			
Genotype	Normal	Overweight	Class I - III	Total
A/G	2	0	1	3
	66,67	0	33,33	100
	1,87	0	1,96	1,35
A/A	105	65	50	220
	47,73	29,55	22,73	100
	98,13	100	98,04	98,65
Total	107	65	51	223
	47,98	29,15	22,87	100
	100	100	100	100

Fisher's Exact = 0,603 – Not significant

	wt_NnN		
Genotype	Normal	Not_Normal	Total
A/G	2	1	3
	66,67	33,33	100
	1,87	0,86	1,35
A/A	105	115	220
	47,73	52,27	100
	98,13	99,14	98,65
Total	107	116	223
	47,98	5202	100
	100	100	100

Fisher's Exact = 0,470 – Not significant

None of the other seven SNVs showed a p-value less than 0.05. Hence the null hypothesis was rejected across the other seven variations.

3.9 Multinomial Logistic Regression

A multinomial logistic regression is a statistical predictive analysis test performed to compare numerous variables. When the dependent variable is nominal with more than two levels, to explain the relationship between one nominal dependent variable and one or more independent variables. STATA IC15.1 software was used to generate this information.

The test is based on the following assumptions:

1. The dependent variable should be measured at a nominal level, e.g. ethnicity.
2. There are one or more independent variables that are continuous, ordinal or nominal, e.g. weight in this study using the BMI score.
3. There should be independence of observations and the dependent variable should have mutually exclusive and exhaustive categories.
4. No multicollinearity should occur (two or more independent variables highly correlated to one another).
5. A linear relationship needs to exist between any continuous independent variables and the logit transformation of the dependent variable.
6. There should be no outliers.

Two multinomial logistic regression tests (mlogit) were performed on the two pooled data sets mentioned above (Cl I-III and Wt_NnN). Firstly, looking at the relationship to the genotypes and then looking at the relationship to ethnic groups.

Only the *LEPR* and *OPRM1* SNVs showed significance, the *COMT, DRD2, DRD4, HTR2C, SLC6A* and *GABRA6* SNVs p-values did not show any significance.

LEPR

Pooled odds ratios (ORs) and corresponding 95% confidence intervals (CI) were used to estimate the association strength between *LEPR* polymorphisms and obesity. When setting up the mlogit model, a number of parameters needed to be set. Firstly, the base comparison group needs to be established and, in this case, a “Normal” BMI was set as the comparison group. Depending on the pooled data set, the groups varied. In pooled data set 1, overweight was given the value “2” and pooled dated “Cl I – II” was assigned the value “3”. In pooled data set 2, “normal” weight remained the comparison group, while wt_NnN (pooled date from overweight and obesity class I-III) was designated the value 2. The indicators or categorical variables were each of the eight SNVs. The race groups were assigned as follows, Caucasians were the comparison group (purely because of the majority sample size) and Africans was assigned group number 2. The risk allele homozygotes genotypes or major alleles were given the priority group number 1, followed by heterozygotes group 2 and minor allele homozygotes group 3. The number of observations was 215 as opposed to 223, this is because the statistical program excluded the ethnic groups Coloured and Indians due to the small sample numbers (223 – 8 = 215).

The likelihood ratio Chi-square (LR chi (4)) along with the p-value < 0.05 (prob > chi2) gives an indication of how the model fits significantly better than the empty model (the model with no predications). When comparing the α level, which is regarded as the willingness to accept a type 1 error, a smaller p-value shows that at least one of the regression coefficients in the model are equal to zero. The α value is the determinant of the level of significance, the value of 0.05 was used for this study.

When reflecting on the statistical analysis regarding the *LEPR* gene, the LR chi is 18.08, with a p-value<α (0.0012). The GG genotype (double mutant allele or risk allele) is the base group. One unit increase in the variable of over-weight is associated with a 0.37 decrease in the relative log odds of being in the overweight group versus the

normal group (OR = 0.37) [95% CI = 0.179; 0.765]. Similarly, for one unit increase in the variable BMI Pool Cl I-III is associated with a 0.148 decrease in the relative log odds of being in the BMI pooled group versus the normal groups (OR = 0.148) [95% CI = 0.0518; 0.423]. The relative log odds of being in the overweight group versus the normal group will increase 1.05 if moving from the GG genotype to the AA genotype and 1.68 for the BMI pool versus the normal group. The odds ratio of being overweight with the AG genotype is 2.63 [95% CI = 1.13; 6.13], and overweight with an AA genotype is 1.05 [95% CI = 0.403; 2.722]. Similarly, the odds ratio of being in the pooled BMI class with the AG genotype is 3.39 [95% CI = 1.782; 17.577], and AA genotype is 1.687 [95% CI = 0.469; 0.064]. The GG genotype is more prominent in the normal BMI group. The AG genotype showed significance for both overweight and pooled BMI classes, 0.025 and 0.003 respectively. Refer to Appendix B6 for Statistical Data

For the pooled data set 1 and ethnicity showed very similar results. The LR chi is 21.65, with a p-value < 0.005 (0.0014). The Caucasian ethnic group is the base group, because of the largest number of observations and GG genotype (double mutant allele or risk allele) being the base group. One unit increase in the variable of overweight is associated with a 0.438 decrease in the relative log odds of being in the overweight group versus the normal group (OR = 0.438) [95% CI = 0.214; 0.978]. Similarly, for one unit increase in the variable BMI pool is associated with a 0.135 decrease in the relative log odds of being in the pooled data set 1 versus the normal groups (OR = 0.135) [95% CI = 0.032; 0.459]. The relative log odds of being in the overweight group versus the normal group will increase 0.903 if moving from the GG genotype to the AA genotype and 1.63 for the BMI pool versus the normal group. There was no significance between the genotypes and ethnicity for *LEPR*.

Given that there were only two variables being compared namely, normal weight versus wt_NnN, a logical regression analysis was performed as opposed to a multinomial logical regression. The LR chi is 16.46, with a p-value < 0.005 (0.0003), for pooled data set 2 (wt_NnN), once again showing significance. The baseline odds, showed a one unit increase in the variable of wt-NnN and is associated with a 0.318 decrease in the relative log odds of being in the not normal weight group versus the normal group (OR = 0.318) [95% CI = 0.272; 0.989], with a significance value of 0.046.

The AG genotype still showed the greatest significant p> |z| 0.001, indicating that the AG genotype was more prominent in overweight and obese individuals. Ethnicity did not appear to have any significance on the *LEPR* genotypes observed between the normal group and the wt_NnN group.

OPRM1

The mlogit regression for the *OPRM1* gene showed significance with the LR chi value 9.90 and with a p-value of 0.0421 (<0.05). The AA genotype (double wild type allele) was the base variable. One unit increase in the variable of overweight is associated with a 0.38 decrease in the relative log odds of being in the overweight group versus the normal group (OR = 0.38) [95% CI = 0.41; 0.824]. Similarly, for one unit increase in the variable BMI pool is associated with a 0.31 decrease in the relative log odds of being in the BMI pooled group versus the normal groups (OR = 0.31) [95% CI = 0.355; 0.735]. The odds ratio of being overweight with the AG genotype is 1.6 [95% CI = 0.713; 3.6], and overweight with an AA genotype is 1.05 [95% CI = 0.403; 2.722]. Similarly, the odds ratio of being in the pooled BMI class with the AG genotype is 3.39 [95% CI = 1.782; 17.577]. There were no participants in the pooled BMI class with the GG genotype. Only three normal weight Caucasians had the GG genotype.

The AA genotype was the only genotype observed in the African population. The GG genotype was only observed in the Caucasian population. The standard error is questionable as there were only three observations for the GG genotype.

When conducting the two variable logical regression on the wt_NnN pooled data, no significance was found (p-value = 0.9249), which is close to 1. In addition, the number of observations decreased to 212 from 215. The program excluded the three GG genotypes observed for normal weight individuals, to correct for the questionability of the standard error. One unit increase in the variable of wt_NnN is associated with a 1.093 increase in the relative log odds of being in overweight or obese versus normal, giving a 1:1 ratio (OR = 1.093) [95% CI = 0.814; 1.464]. The AA genotype was more prominent than the AG or GG genotype, by more than double.

Furthermore, there was no significance observed between ethnic groups, except that only the AA genotype was observed in the African population. This makes the relative frequency of the genotype equal between normal weight and overweight/obese individuals.

COMT

Although there was no significance, to be noted was that the AA and AG genotype was observed more frequently in the normal weight African group than any other genotype. The GG genotype was observed equally between the normal, overweight and pooled BMI groups with three observations each. The binomial logical regression for COMT showed no significance between normal weight Caucasian group with a GG genotype.

The AA genotype appeared to be more common in normal weight African individuals, while the GG genotype is more prominent in the overweight and obese individuals.

Remaining SNVs

For the *DRD2 DRD4, GABRA6* and *HTR2C* no significance was observed across the four logistical regression models. For *SLC6A4*, only the AA genotype was observed in the Caucasian population. The AA and AT genotypes were found in equal frequencies in the African population. However, still no significance was observed across the four logical regression models.

3.10 Single nucleotide variation combinations associated with a high BMI

The final study objective was to investigate whether there was a combination of the eight SNVs that was more prominent in overweight and obese individuals and linked more significantly to obesity. Due to the lack of significance across the eight SNVs, this objective could not be achieved. A larger sample size and a less biased sampling strategy could have increased the statistical power of the project. Participants were selected based on their BMI score and then pre-classified into the respective BMI classes, which created a certain level of sample bias. Selecting samples randomly without prior knowledge of their BMI and investigating the lifestyle and environmental factors impacting on the individuals, could have provided a more accurate indication of the prevalence within the population. However, socio-environmental factors were

purposefully excluded during the planning phase of this study. Although normal weight individuals were selected based on their current BMI score, this is not to say that the person has always had a normal BMI, or maybe they have but they have better control of their eating behaviour to control previous habits of overeating. Weight maintenance could still be a continuous struggle.

CHAPTER 4: DISCUSSION

The TaqMan® OpenArray™ genotyping allowed for accurate high-through put selected SNP genotyping of the eight SNVs across the 223 samples analysed. The technique was validated through sequencing and repeat testing of a portion of the samples. Three OpenArray™ plates were processed and analysed, with 17.5% of the samples repeated for accuracy and reproducibility of results.

Of the 247 swabs collected only 223 samples were used for analysis purposes based on the inclusion criteria. There was a 50:50 split between normal weight and overweight samples. Approximately 82% of the samples collected were Caucasians. There was a 70:30 split between females and males. The observed split is in line, with the WHO's obesity statistics where 62% of females are overweight or obese and 41% of males.

Of the eight SNVs investigated that have been previously linked to the brain reward system, only one of these variants showed a significant association to a high BMI score, namely *LEPR* (rs1137101). Results were partially in agreement with previous international studies, where the AG genotype was associated with a higher BMI. A meta-analysis found the AG genotype to result in a 1.25 increased risk in obesity and type 2 diabetes. While the GG genotype results in a 1.5 increase[161, 162]. In contrast, this study did not show similarity with the GG genotype. The GG genotype was seen predominately in normal weight participants.

The *LEPR* gene encodes the protein called leptin receptor. The leptin receptor is involved in the regulation of body weight. Although, the protein if found on the surface of many cells and tissues in the body, its most important function is in the hypothalamus, where it controls hunger, thirst, sleep, mood, body temperature and hormone functions. Leptin, which is released by the body's fat cells, activate the leptin receptor. As adipose cells become larger, the more leptin is produced. The rise in leptin indicates that fat stores increased. Binding of leptin to the leptin receptor in the hypothalamus triggers a series of chemical signals that affect hunger and produces the feeling of satiety[163]. The G allele of the *LEPR* gene has been associated with an increased risk of obesity and type 2 diabetes. From Table 23 it was noted that 41% of normal weight Caucasian population had the AA genotype, while 19% of pooled overweight and obese Caucasian population had the AA genotype. A similar

percentage of normal weight (20%) and pooled overweight and obese (13%) Caucasian individuals had the GG genotype. In terms of the African population, the AA genotype was slightly more dominant in normal weight individuals (42%) than overweight or obese individuals (33%), where the AG (50% and 58% respectively) and GG (5% and 8% respectively) genotype showed similar frequencies between normal and overweight/obese individuals. Our current data is not in agreement with previous published work, however, this could be due to the small sample size, given the meta-analysis study.

By doing a simple search on the NCBI database of SNVs for the *LEPR* gene, there are 47962 reported Leptin SNVs. When deciding on which SNV to select, a comprehensive review of previous research was conducted, the most relevant and well researched SNVs were selected. If more than 1% of a population carries the same nucleotide at a specific position in the DNA sequence, then this variation can be classified as a SNV.

Carriers of at least one *OPRM1* (rs1799971), G allele appear to have stronger cravings for alcohol and carbohydrates than carriers of two A alleles[112]. The GG genotype was only found in normal weight females, three were of Caucasian ethnicity and one was of Indian ethnicity. The AA genotype was more prominent in the overweight and obese participants; however, it did not reach significance.

In the *COMT* (rs4680) gene, the A allele carriers tend to have more dopamine in their prefrontal cortex, due to lower *COMT* enzymatic activity. They also have a lower pain threshold making them more susceptible to stress[133-135]. Increased stress levels may lead to poor eating habits, either binge eating or absence of eating. The AA genotype in the study was found at a high relative frequency in normal weight African followed by the normal weight Caucasian group. The AA genotype did not have a high frequency distribution among the various BMI categories among Caucasians.

The *DRD4* (rs1800955) AA genotype was more prevalent in the overweight and obese population of Caucasian, while it was more prevalent in the normal weight African group. Carries of the long allele (C) have shown to have higher novelty seeking traits and risk-taking traits compared to non-carriers[133-135].

HTR2C (rs3813929) has a strong link with an increased BMI and unhealthy feeding behaviour. The CC genotype has decreased serotonin receptor binding, resulting in an increased risk for antipsychotic induced weight gain, high BMI and increased

overeating feeding behaviour. The T allele has increased serotonin receptor binding associated with protection against antipsychotic induced weight gain, increased glucose and acute insulin. The TT genotype was not found in the African population and the CT genotype was only observed in one of the African populations. The CC genotype was observed more often, however, there was no significant association between BMI categories.

It appears that SNVs targeted during this study that have been associated with the brain reward system have a limited effect on BMI. In the study we are only focused on the inherited genetics and not the active genetics, in other words what effect the environment has of the genes and their functioning.

CHAPTER 5: GENERAL INTRODUCTION TO STUDY BIAS

5.1 Study Bias & Limitations

Bias is defined as the “inclination or prejudice or against one person or group, in a manner that is considered to be unfair”, according to the Oxford English Dictionary definition (https://en.oxforddictionaries.com/definition/bias). In research this is a form of systematic error which affects scientific investigation and affects the measurement process. It is nearly impossible to almost completely eliminate bias from any study[164, 165]. It is essential to understand research and the implications of bias for several reasons:

1. Bias is present in all research and is difficult to eliminate 100%
2. Bias can also occur at every stage of the research process
3. Bias impacts the validity and reliability of the study findings, leading to incorrect interpretation of the data which could potentially have other consequences.

There are various forms of bias namely; design, sampling/selection, statistical, procedural, measurement and reporting bias. The effect of each type of bias and how it relates to the study is discussed below.

5.1.1 Design Bias

When inherent biases in experiments, are not considered, that is when design bias is introduced. The definition of risk and outcome should be clearly defined prior to study implementation, both the objectives and study validation methods[166].

Data collection of participants could have included more comprehensive information about the participants, such as an inclusion of a lifestyle questionnaire and/or medical history, to also understand the environmental impacts of the phenotypes observed in the participants and the study. Collection of samples should have been more blind and random. Participants were collected on the basis of their current BMI score. This did not take into account whether the higher BMI observed has persisted from childhood (has a higher BMI being problematic from childhood to adulthood), or if it was attributed to a new development due to environmental stressors. Although the protocol for data collection was standardised, it included external personnel in the collection process.

There were various collectors of the buccal swabs which could have affected the study design, even if personnel were trained in the participant criteria and sample collection of participants. Lab work and data capturing was performed by the author, which was standard and consistent throughout.

The study design was incomplete and did not consider external variables such as swab collection variation. This also affected the quality of the collected DNA samples. Samples from Site A had an overall lower DNA concentration across the samples.

5.1.2 Sampling/Selection Bias

Selection bias occurs during the identification of the study population. A study population should be clearly defined, accessible, reliable and be at a decreased likelihood of the outcome of interest. Selection bias also occurs when the recruitment and enrolment of participant is not uniform, resulting in collecting sampling data incorrectly[166].

Selection bias was introduced into the study, as participants were selected from two distinct groups, namely normal weight controls and overweight and obese participants. By creating two strata and selecting participant with a certain BMI of interest we selected participants with an increased likelihood of developing the outcome of obesity. Furthermore, it would have been more beneficial to have equal numbers of male and female participants, as well as equal numbers of White and African participants.

Prospective studies, such as this one, where the outcome is unknown at the time of enrolment are less prone to selection bias. However, the sample collection was not entirely randomized, the two groups were predefined, which in itself created both design and selection bias. Participants were selected with a certain BMI based exclusively on measured values, which could have been biased towards the expected result.

5.1.3 Statistical Bias

Statistical bias is introduced during statistical analysis, when the data is analysed is such a matter that it gives preference to the conclusion of the research hypothesis. There are various ways bias can be introduced, i.e. fabrication of data (reporting of

non-existing data), abusing or manipulating data (either eliminating data which does not support the hypothesis, use of inappropriate statistical tests, performing multiples testing fishing for a P values by pair-wise comparisons). Usually researchers try to torturer the data to find the association they are aiming to achieve[167]. There is an unpublished quote by Ronald Coase that says, “If you torture the data long enough, it will confess to anything”.

Essentially this is exactly what was attempted to achieve by conducting several Fisher Exact Tests using pooled data sets. It was only the LEPR gene that remained constant throughout, suggesting the significant difference is real and did not occur purely by chance. A larger sample size, which randomized sample collection could have increased the statistical power.

5.1.4 Procedural Bias

When subjects are required to complete questionnaires or documentation, but pressure is applied to complete the information and responses quickly this is when procedural data is introduced. This form of bias would not be relevant to our study, as the participants were not required to report any additional information and questionnaires. Procedural bias could also relate to the manner in which other personal data was collected. For the study we required the patient’s anthropometrics, age, gender and race. Some of the participants were only weighed, while some patients had a complete body assessment done. Other information was completed by the participants and checked by staff against the patient file to ensure accuracy. Or the information was solely completed by the Healthcare Practitioner’s staff. The diversity also created procedural bias.

5.1.5 Performance Bias

When establishing cause and effect relationships between procedure and outcomes, performance bias may complicate this. A key means of avoiding performance bias is to cluster participants. This will minimise performance variability within a group and decrease performance bias[166], however it is study relevant.

Data was collected using two strata namely; normal weight and overweight or obese. However, due to the fact that study’s focus was on the prevalence of eight SNVs within

a geographical location and the cause and effect relationship between the genotype and phenotype, it might have been better first randomly selecting participants for the study (in excess of 200 samples) and then divide them into clusters based on the phenotype and supporting information from the lifestyle questionnaire.

5.1.6 Measurement Bias

Measurement bias or rather measurement error refers to systemic or non-random error that occurs when a certain measurement process is not uniform resulting in either an underestimate or overestimate of the actual measure[166]. Human error is always present, and error in data recording can occur. Different scales were used between the two sites, and measurements were not done by one single person, which could create inconsistency. However, samples were collected across a wide BMI range which counted this potential bias.

Measurement bias was not a major contributing factor in the data bias seen in the study. Essentially since there were only two sites and one data capturer of the information. All samples were prepared and analysed by the author. Therefore, the variability in the processing and analysis would be very limited. Genotyping analysis was done using a computer software which was then exported into a .csv file for further analysis. There was very limited chance of human error to occur in this part of the study. The necessary controls were also in place to ensure consistency of results.

5.1.7 Reporting Bias

Reporting bias is closely linked to statistical bias and usually introduced when results are either selective revealed or certain information suppressed by the participant. Example of such information can include, past medical history, alcohol abusing, smoking habits and sexual experience[166].

Lack of supporting information was a major pitfall to this study. Reporting bias was absent, due to the lack of information required from participants. This once again, highlighted the important of a lifestyle questionnaire. Although it is unrealistic to assume that the information collected would be a 100% accurate way of measuring participant clustering, it may have created greater insight into the environmental impact and better classification of participants. Instead of classifying participants as either

normal weight or overweight/obese, participants could have been additionally classified as having no history of an elevated BMI and those that have struggled with their weight and BMI. This might have resulted in some normal weight individuals being classified into a different phenotypic category, and vice versa, if a particular individual is now classified as overweight as a result of environmental stress, but in the past always attained a normal weight. Given the sensitive and personal nature of obesity, people would rather hide the stigma of being overweight and obese than being honest about their weight, especially if they have a low self-esteem. This may contribute to reporting bias.

CHAPTER 6: CONCLUSION

The need to decrease the rate of obesity in South Africa has been acknowledged and Government six broad goals to lower the rate of obesity in the country namely;

1. Create an institutional framework to support intersectoral engagement;
2. Create and enabling environment that supports availability and accessibility to healthy food choices in various settings;
3. Increase the percentage of the population engaging in physical activity.
4. Support obesity prevention in early childhood;
5. Communication with, educate and mobilise communities; and
6. Establish a surveillance system and strengthen monitoring, evaluation and research[168].

The South African Government has also introducing a sugar tax on sweetened drinks, to motivate healthier soft drink and food choices[169]. The notion of “one size fits all” is no longer applicable to the treatment of obesity with new research into the effects of an individual’s genetic makeup and SNVs which give an indication of a person’s predisposition to various diseases, obesity being one of the diseases of focus.

The aim of the study was to determine the prevalence of eight SNVs associated with the control and regulation of the neurotransmitters in the brain reward cascade, which have been linked to addictive behaviour and food cravings in overweight and obese individuals. Investigating the genetics and hereditary basis of obesity, could help with pre-screening and implementation of early intervention of preventative strategies. Understanding how an individual respond’s to certain nutrients could assist in the prevention and treatment of obesity, including personalizing an individual’s intervention strategies and meal plans. Obesity is a multifactorial disorder and of importance are the effects of an individual’s environmental and lifestyle factors.

Limited genetic research with respect to obesity has been conducted locally and it was envisaged with this study to gain some insight into the allele frequency distribution within the South African population.

When increasing the sample size the standard error becomes smaller and the sample mean more accurate. The sample size decreased the statistical power and the lack of environmental and lifestyle information on the participants narrowed the interpretation

of the results. A more detailed history into the lifestyle of participants and daily challenges or stressors they face relating to weight management, might have shed light into the discrepancies observed in the data. Although, the allele frequencies were in agreement with data published in the 1000 Genomes Project and HapMap, the discrepancies in terms of the phenotype presentation could be due to a lack of supporting information environmental factors. For example, a normal weight individual was not asked, whether he/she experience weight management problems, have they always been of normal weight (from childhood to adulthood), or whether they employ any strategies to maintain a normal weight. This information might have changed the categorization of the participant, rather than a simple scale and height measurement.

Overweight and obese individuals were not asked about their occupation or emotional state. Whether they have always been overweight, or whether this has been a recent change in their physiology due to certain stressors or medication they have started taking. Also learning more about the participant's personality could have added additional benefits, and what their eating tendencies are. Do they binge eat, enjoy snacking, prefer sweet or salty foods, or do they experience loss of appetite.

Knowing that obesity is linked to several other diseases, it would have been of importance to gain insight into the medical history of the patients, i.e do they currently suffer from type II diabetes, insulin resistance, heart disease, cholesterol or blood pressure problems. This information might have linked other co-morbidities to the genetic variations investigated.

When incorporating the lifestyle and medical history information into the interpretation of the results, the classified of individuals into various risk groups might have been better suited (low, moderate and high), rather than classifying the participants according to their BMI.

CHAPTER 7: STUDY LIMITATIONS AND RECOMMENDATIONS

7.1 Study Limitations

Although the sample size was statistically calculated to be a representation of the general population, the sample size of 197 recommended and the number of samples collected (n=223) was still too small to make a generalisation of the population. A small sample size decreases statistical power and increases the margin of error (likelihood of Type II error). There were limited samples collected from other race groups other than Caucasian and Africans. South Africa is a diverse population and it would have been beneficial to have a wider distribution of sample demographics. Sample bias was potentially generated through the selection of overweight and obese participants, and through the collection of samples from a slimming clinic, where the proportion of overweight and obese individuals are naturally high.

7.2 Recommendations

There were a couple of lessons learnt from this study, that a better sampling strategy should have been employed. More sampling sites beneficial to broadening the collection of various ethnic groups would have been more ideal. Sampling from a Slimming clinic could have created some sampling bias, as the population of overweight and obese individual are naturally higher.

It will be beneficial to incorporate a lifestyle questionnaire in future studies. The following information might have been important to the study:

1. Have you struggled with your weight during your life time?
2. If yes, for how long? Childhood, Adulthood or both?
3. Do you currently follow a meal plan of sorts? Are you self-conscience of your eating habits?
4. What is your current occupation?
5. What sort of environmental or lifestyle stressors are you currently experiencing?
6. Are you on any medication or have you been diagnosed with any aliments, i.e blood-pressure, cholesterol, depression, etc.
7. Do you suffer from any addictive tendencies such as binge eating or stress eating?
8. What types of foods to you crave?

9. Do you prefer to cook a meal or is take out more convenient?

Obesity is a complex disorder, thus combining both genetic and environmental information is of utmost importance to understand the interplay of various influencing factors. Simply dividing the population in either normal weight or overweight, does not supply full and complete picture. It would have been better to divide the population into low risk, moderate risk and high risk, based on their lifestyle information as well as their BMI score. Although any lifestyle questionnaire utilised would need to be validated. Some individuals naturally do not experience any problems with managing their weight, in which event they would be classified into the low risk group. Other individuals, although of normal weight, maybe more proactive in their weight management, making the necessary lifestyle changes to reduce their risk. Or they have recently become overweight due to a change in their lifestyle, e.g. change of job, personal problems or prescription medication. Or they could potentially be in a profession where their image is of utmost importance and are thus self-conscience of their weight, i.e. dietician, tv personality or sportsman. Overweight and obese individuals who find weight loss a constant battle, where they lose weight but regain it again, would be considered the high-risk group.

Measuring other factors such as blood pressure, cholesterol, glucose levels, insulin cortisol, leptin, gherlin and thyroid hormones might also be of importance. Given that our body is a network, it might prove to be of importance when assessing potential co-morbidities of obesity.

REFERENCES

1. Sanderson S. SA declared the fattest nation in Sub-Saharan Africa. The South African. 2014 4 June.
2. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C. Global, regional,and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2014;384(9945):766-81.
3. World Health Organization. Obesity & Overweight Fact Sheet. Medica Centre. 2016.
4. World Health Organization. Global Status Report on Non-Communicable Diseases. WHO Library Cataloguing in Publication Sata; 2014.
5. Cowley MA, Brown W A CR. Obesity: The Problem and Its Management In: J Larry Jameson LJDG, David M de Kretser, Linda C Giudice, Ashley B Grossman, Shlomo Melmed, John T Potts Jr, Gordon C Weir, editor. Endocrinology: Adult & Pediatric (7th ed, pp 468-478). Philadelphia PA USA: Elsevier; 2016. p. 468-78.
6. Xia Q, Grant SF. The genetics of human obesity. Annals of the New York Academy of Sciences. 2013;1281:178-90.
7. Loos RJF, Bouchard C. Obesity - is it a genetic disorder? Journal of Internal Medicine. 2003;254:401-25.
8. MHA News [Internet]. How Will Obesity Rates in the Future Affect Your Healthcare Career? MHA Degree2014 [02 February 2018]. Available from: https://mhadegree.org/obesity-and-your-healthcare-career/.
9. Baleta A, Mitchell F. Country in Focus: diabetes and obesity in South Africa. Lancet Diabetes Endocrinology. 2014;2:687-8.
10. Jessica B, Gabriel E, Crickmore C. Cardiovascular Disease Statistics Reference Document. The Heart Foundation of South Africa; 2016.
11. Carpenter CL, Wong AM, Li Z, Noble EP, Heber D. Association of dopamine D2 receptor and leptin receptor genes with clinically severe obesity. Obesity (Silver Spring). 2013;21(9):E467-73.
12. Tomlinson M, Grimsrud AT, Stein DJ, Williams DR, L M. The Epidemiology of Mental Disorders. South African Medical Journal. 2009;99(5 Pt 2):367 - 73.
13. Oswald R. Body mass index and depression related to serotonin-transporter gene and COMT gene polymorphisms: University of Wein; 2013.

14. World Health Organization [Internet]. Media Centre: Depression Fact Sheet No 3692017 [10 December 2017]. Available from: http://www.who.int/mediacentre/factsheets/fs369/en/.
15. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). 5th ed: American Psychiatric Publisher; 2013. 1-1520 p.
16. American Psychiatric Association. Supplement to Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). 5th ed. Washington, DC: American Psychiatric Association Publishing; 2016. 1-56 p.
17. Reeves GM, Postolache TT, Snitker S. Childhood Obesity and Depression: Connection between these Growing Problems in Growing Children. International Journal of Child Health and Human Development. 2008;1(2):103-14.
18. Heber D, Carperter CL. Addictive genes and the relationship to obesity and inflammation. Molecular Neurobiology. 2011;44(2):160-5.
19. Nemiary D, Shim R, Mattox G, Holden K. The Relationship Between Obesity and Depression Among Adolescents. Psychiatric Annals. 2012;42(8):305-8.
20. Stunkard A.J, Foch TT, Hrubec Z. A twin study of human obesity. Journal of the American Medical Association. 1986;256(1):51-4.
21. Bell CG, Walley AJ, P F. The genetics of human obesity. Nature Reviews Genetics. 2005;6(3):221-34.
22. Berrettini W. The Genetics of Eating Disorders. Psycharity. 2014;November 2014:18-25.
23. Bogg RA. Genetic Component of Obesity. The Lancet. 1976;308(7996):1205.
24. Lehner B. Modelling genotype-phenotype relationships and human diseas with genetic interaction networks. Journal of Experimental Biology. 2007;210:1159-566.
25. Newell A, Zlot A, Silvey K, Ariail K. Addressing the Obesity Epidemic: A Genomics Perspective. Preventing Chronic Disease. 2017;4(2):1-6.
26. Bush WS, Moore JH. Chapter 11: Genome-Wide Association Studies. PLoS Compitational Biology. 2012;8(12).
27. Camp KM, Trujillo E. Position of the Academy of Nutrition and Dietetics: Nutritional Genomics. Journal of the Academy of Nutrition and Dietetics. 2014;114(2).

28. Stunkard AJ, Harris JR, Pedersen NL, McClearn GE. The body mass index of twins who have been reared apart. The New England Journal of Medicine. 1990;322(21):1483-7.
29. Tholin S, Rasmussen F, Tynelius P, Karlsson J. Genetic and environmental influences on eating behavior: the Swedish Young Male Twins Study. American Journal of Clinical Nutrition. 2005;81:564–9.
30. Whitaker RC, Wright JA, Pepe MA, Seidel KD, WH D. Predicting obesity in young adulthood form chilhood and parental obesity. The New England Journal of Medicine. 1997;337(13):869-73.
31. Maffeis C, Talamini G, L T. Influence of diet, physical activity and parents' obesity on children's adiposity: a four-year longitudinal study. International Journal of Obesity and Related Metabolic Disorders. 1998;22(8):758-64.
32. Wardle J , Guthrie C, Sanderson S, Birch L, R P. Food and activity preferences in children of lean and obese parents. International Journal of Obesity and Related Metabolic Disorders. 2001;25(7):971-7.
33. Kunej T, Skok DJ, Zorc M, Ogrinc A, Michal JJ, Kovac M, et al. Obesity Gene Atlas in Mammals. Journal of Genomics. 2013;1:45-55.
34. Herrera BM, Lindgren CM. The Genetics of Obesity. Current Diabetes Reports. 2010;10:498–505.
35. Shawky RM, Sadik DI. Genetics of Obesity. The Egyptian Journal of Medical Human Genetics. 2012;13:11-7.
36. Blum K, Sheridan PJ, Wood RC, Braverman ER, Chen TJ, Cull JG, et al. The D2 dopamine receptor gene as a determinant of reward deficiency syndrome. Journal of the Royal Society of Medicine. 1996;89(7):396-400.
37. Blum K, Badgaiyan RD, Agan G, Fratantonio J, Simpatico T, Febo M, et al. Molecular Genetic Testing in Reward Deficiency Syndrome (RDS): Facts and Fiction. Journal of Reward Deficiency Syndrome and Addiction Science. 2015;1(1):65-8.
38. Blum K, Fornari F, Downs BW, Waite RL, Giordano J, Smolen A, et al. Genetic Addiciton Risk Score (GARS): Testing for Polygenetic Presisposition and Risk to Reward Deficiency Syndrome (RDS). In: Kang PC, editor. Gene Therapy Applications: InTech; 2011.
39. The High Emotional Cost Of South Africa's Financial Crisis: The South African Depression and Anxiety group; [18 June 2018]. Available from:

http://www.sadag.org/index.php?option=com_content&view=article&id=314:the-high-emotional-cost-of-south-africa-s-financial-crisis&catid=64&Itemid=132.

40. Torres SJ, Nowson CA. Relationship between stress, eating behavior, and obesity. Nutrition. 2007;23(11-12):887-94.
41. Hammen C. Stress and depression. Annual Review of Clinical Psychology. 2005;1:293-319.
42. van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, et al. The industrial melanism mutation in British peppered moths is a transposable element. Nature. 2016;534:102-5.
43. Unger RH, Orci L. Lipotoxic diseases of nonadipose tissues in obesity. International Journal of Obesity and Related Metabolic Disorders. 2000;24(Suppl 4):S28-32.
44. Eaton SB, Konner M, M S. Stone Agers in the Fast Lane: Chronic Degenerative Diseases in Evolutionary Perspective. American Journal of Medicine. 1988;84:739-49.
45. Meyer R [Internet]. The biggest threat to healthcare World of DTC Marketing.com2015 [06 February 2018]. Available from: http://worldofdtcmarketing.com/the-biggest-threat-to-healthcare/cost-of-healthcare-in-the-u-s/.
46. Genetics Home Reference [Internet]. What is a gene? : U.S National Library of Medicine; 2018 [07 November 2018]. Available from: https://ghr.nlm.nih.gov/primer/basics/gene.
47. Mou Z, Hyde TM, Lipska BK, Martinowich K, Wei P, Ong G, et al. Human Obesity Associated with an Intronic SNP in the Brain-Derived Neurotrophic Factor Locus. Cell Report. 2015;13(6):1073-80.
48. What is Epigenetics? Epigenetics: Fundamentals [04 May 2017]. Available from: http://www.whatisepigenetics.com/fundamentals/.
49. Ray M. Epigenetics. Environmental Health Perspectives. 2006;114(3):A160-A7.
50. Ochoa MC , Marti A, Azcona C, Chueca M, Oyarzábal M, Pelach R, et al. Gene-gene interaction between PPAR gamma 2 and ADR beta 3 increases obesity risk in children and adolescents. International Journal of Obesity and Related Metabolic Disorders. 2004;28 Suppl(3):S37-41.
51. Argyropoulos G , Rankinen T, Neufeld DR, Rice T, Province MA, Leon AS, et al. A polymorphism in the human agouti-related protein is associated with late-onset

obesity. The Journal of Clinical Endocrinology and Metabolism. 2002;87(9):4198-202.

52. C B. The biological predisposition to obesity: beyond the thrifty genotype scenario. International Journal of Obesity. 2007;31:1337–9.

53. Martı´nez JA, Corbala´ n MS, Sa´ nchez-Villegas A, Forga L, Marti A, Martı´nez-Gonza´ lez MA. Obesity Risk is Associated with Carbohydrate Intake in Women Carrying the Gln27Glu 2-Adrenoceptor Polymorphism. Journal of Nutrition. 2003;133(8):2549-54.

54. Ravussin E , Valencia ME, Esparza J, Bennett PH, LO S. Effects of a traditional lifestyle on obesity in Pima Indians. Diabetes Care. 1994;17(9):1067-74.

55. Cassano PA, Rosner B, Vokonas PS, Weiss ST. Obesity and Body Fat Distribution in Relation to the Incidence of Non-Insulin-dependent Diabetes Mellitus. American Journal of Epidemiology. 1992;136(120):1474-86.

56. Olson JE, Atwood LD, Grabrick DM, Vachon CM, Sellers TA. Evidence for a major gene influence on abdominal fat distribution: The Minnesota breast cancer family study. Genetic Epidemiology. 2001;20(4):458–78.

57. Barry D, Clarke M, Petry NM. Obesity and Its Relationship to Addictions: Is overeating a Form of Addictive Behaviour. The American Journal of Addiction. 2009;18(6):439-51.

58. Holden C. 'Behavioral' addictions: do they exist? Science. 2001;294:980-2.

59. Kenny PJ [Internet]. Is Obesity an Addiction? : Scientific American; 2013 [07 July 2018]. Available from: https://www.scientificamerican.com/article/is-obesity-an-addiction/.

60. Blum K, Braverman ER, Holder JM, Lubar JF, Monastra VJ, Miller D, et al. The Reward Deficiency Syndrome: A Biogenetic Model for the Diagnosis and Treatment of Impulsive, Addictive and Compulsive Behaviors. Journal of Psychoactive Drugs. 2000;32(sup1):1-112.

61. Blum K, Oscar-Berman M, Demetrovics Z, Barh D, Gold MS. Genetic Addiction Risk Score (GARS): Molecular neurogenetic evidence for predisposition to Reward Deficiency Syndrome (RDS). Molecular Neurobiology. 2014;50(3):765-96.

62. Wurtman RJ, Wurtman JJ. Brain Serotonin, Carbothydrate-Craving, Obesity and Depression. Obesity Research. 1995;3(Supp 3).

63. Enns JE, Taylor CG, Zahradka P. Variations in Adipokine Genes AdipoQ, Lep, and LepR are Associated with Risk for Obesity-Related Metabolic Disease: The Modulatory Role of Gene-Nutrient Interactions. Journal of Obesity. 2011;2011:168659.
64. Phillips CM, Goumidi L, Bertrais S, Field MR, Ordovas JM, Cupples LA, et al. Leptin receptor polymorphisms interact with polyunsaturated fatty acids to augment risk of insulin resistance and metabolic syndrome in adults. Journal of Nutrition. 2010;140(2):238-44.
65. Rojano-Rodriguez ME, Beristain-Hernandez JL, Zavaleta-Villa B, Maravilla P, Romero-Valdovinos M, Olivo-Diaz A. Leptin receptor gene polymorphisms and morbid obesity in Mexican patients. Hereditas. 2016;153:2.
66. Brennan AM, Mantzoros CS. Drug Insight: the role of leptin in human physiology and pathophysiology—emerging clinical applications. Nature Reviews Endocrinology. 2006;2:318-27.
67. Ahima RS, Antwi DA. Brain regulation of appetite and satiety. Endocrinology and Metabolism Clinics of North America. 2008;37(4):811–23.
68. Marais A. The Pharmacological Management of Obesity. South African Family Practice 2016;58(4):16-21.
69. Australian Medicines Handbook: MIMS annuals; 2017 [13 November 2018]. Available from: ttp://www.mimsonline.com.au.
70. Santarsieri D, Schwartz TL. Antidepressant efficacy and side-effect burden: a quick guide for clinicians. Drug Context. 2015;4.
71. What are SSRIs? : WebMD; [24 May 2017]. Available from: www.webmd.com/depression/ssris-myths-and-facts-about-andtidepressants#1.
72. Osuch E, Marais A. The pharmacological management of depression - update 2017. South African Family Practice. 2017;59(1):6-16.
73. Ogbru O [Internet]. MOAIs (Monoamine Oxidase Inhibitors, MAOI) Medicine Net [25 May 2017]. Available from: www.medicinenet.com/mao_inhibitors-oral/article.htm.
74. What are Serotonin-norepinephrine reuptake inhibitors Drugs.com [25 May 2017]. Available from: www.drugs.com/drugs-class/ssnri-antidepressamts.html.
75. How different antidepressants work WebMD [25 May 2017]. Available from: www.webmd.com/depression/how-different-antidepressants-work#1.

76. John R. Horn, Hansten PD [Internet]. Get to Know an Enzyme: CYP2D6 Pharmacy Times 2008 [01 October 2017]. Available from: http://www.pharmacytimes.com/publications/issue/2008/2008-07/2008-07-8624.
77. Kagimoto M, Heim M, Kagimoto K, Zeugin T, Meyer UA. Multiple mutations of the human cytochrome P450IID6 gene (CYP2D6) in poor metabolizers of debrisoquine. Study of the functional significance of indivdual mutations by expression of chimeric genes. The Journal of Biological Chemistry. 1990;265(28):17209-14.
78. Albert PR. Why is depression more prevalent in women? Journal of Psychiatry Neuroscience. 2015;40(4):219-21.
79. Simon GE, Ludman EJ, Linde JA, Operskalski BH, Ichikawa L, Rohde P, et al. Association between obesity and depression in middle-aged women. General Hospital Psychiatry. 2008;30(1):32-9.
80. John D. Fernstrom, Wurtman RJ. Brain Serotonin Content: Increase Following Ingestion of Carbohydrate Diet. Science. 1971;174(4013):1023-5.
81. Wurtman RJ, Wurtman JJ. Carbohydrate craving, obesity and brain serotonin. Appetite. 1986;Volume 7:99-103.
82. Andrews P, Bharwani A, Lee K, Fox M, Thomsom Jr J. Is serotonin an upper or a downer? The evolution of the serotonergic system and it role in depression and the antidepression response. Neuroscience & Behavioural Reviews. 2015;51:164-88.
83. Nihalani N, Schwartz TL, Siddiqui UA, Megna JL. Weight Gain, Obesity and Psychotropic Prescribing. Journal of Obesity. 2010;2011:9.
84. Pijl H, Meinders AE. Bodyweight change as an adverse effect of drug treatment. Mechanisms and management. Drug Safety. 1996;14(15):329-42.
85. Sun-Edelstein C, Tepper S, Shapiro R. Drug-induced serotonin syndrome: a review. Expert Opinion on Drug Safety. 2008;7(5).
86. Ribeiro E, Bettiker R, Bogdanov M, Wurtman RJ. Effect of systemic nicotine on serotonin release in rat brain. Brain Research. 1993;1993:311-8.
87. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nature Genetics. 1999;22:231-8.

88. Sherry ST, Ward M, Sirotkin K. dbSNP—Database for Single Nucleotide Polymorphisms and Other Classes of Minor Genetic Variation. Genome Research. 1999;9(8):677-9.
89. Mazzeo SE, Bulik CM. Environmental and genetic risk factors for eating disorders: what the clinician needs to know. Child and Adolescent Psychiatric Clinics of North America. 2009;18(1):67-82.
90. Briguglio M, Dell'Osso B, Panzica G, Malgaroli A, Banfi G, Dina CZ, et al. Dietary Neurotransmitters: A Narrative Review on Current Knowledge. Nutrients. 2018;10(5):591.
91. Pascale E, Ferraguti G, Codazzo C, Passarelli F, Mancinelli R, Bonvicini C, et al. Alcohol Dependence and Serotonin Transporter Functional Polymorphisms 5-HTTLPR and rs25531 in an Italian Population. Alcohol and Alcoholism. 2015;50(3):259–65.
92. Humberto Nicolini. Serotonin transporter gene polymorphisms & obsessive-compulsive disorder. Indian Journal of Medicine. 2010;132(6):663-5.
93. Karg K, Burmeister M, Shedden K, Sen S. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. Archives Of General Psychiatry. 2011;68(5):444-54.
94. Ruhe HG, Ooteman W, Booij J, Michel MC, Moeton M, Baas F, et al. Serotonin transporter gene promoter polymorphisms modify the association between paroxetine serotonin transporter occupancy and clinical response in major depressive disorder. Pharmacogenet Genomics. 2009;19(1):67-76.
95. Shioe K, Ichimiya T, Suhara T, Takano A, Sudo Y, Yasuno F, et al. No association between genotype of the promoter region of serotonin transporter gene and serotonin transporter binding in human brain measured by PET. Synapse. 2003;48(4):184-8.
96. Heinz A, Jones DW, Mazzanti C, Goldman D, Ragan P, Hommer D, et al. A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity. Biological Psychiatry. 2000;47(7):643-9.
97. Willeit M, Stastny J, Pirker W, Praschak-Rieder N, Neumeister A, Asenbaum S, et al. No evidence for in vivo regulation of midbrain serotonin transporter availability by serotonin transporter promoter gene polymorphism. Biological Psychiatry. 2001;50(1):8-12.

98. van Dyck CH, Malison RT, Staley JK, Jacobsen LK, Seibyl JP, Laruelle M, et al. Central Serotonin Transporter Availability Measured With [123I]β-CIT SPECT in Relation to Serotonin Transporter Genotype. American Journal of Psychiatry. 2004;161:525-31.
99. IGSR: The International Genome Sample Resource [Internet]. Providing ongoing support for the 1000 Genomes Project data: EMBL-EBI; 2008 [20 June 2018]. Available from: www.internationalgenome.org.
100. 1000 Genomes Project [Internet]. ss1358325909 [ss1358325909]. NCBI dbSNP Short Genetic Variations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1358325909.
101. Lucki I. The spectrum of behaviors influenced by serotonin. Biological Psychiatry. 1998;44:151-62.
102. Drago A, Serretti A. Focus on HTR2C: A possible suggestion for genetic studies of complex disorders. American Journal of Medical Genetics. 2009;150B:601–37.
103. Rocha BA, Goulding EH, O'Dell LE, Mead AN, Coufal NG, Parsons LH, et al. Enhanced locomotor, reinforcing, and neurochemical effects of cocaine in serotonin 5-hydroxytryptamine 2C receptor mutant mice. The Journal of Neuroscience: the Official Journal of the Society of Neuroscience. 2002;22:10039–45.
104. Nonogaki K, Abdallah L, Goulding EH, Bonasera SJ, LH T. Hyper-activity and reduced energy cost of physical activity in serotonin 5-HT(2C) receptor mutant mice. Diabetes. 2003;52 315–20.
105. Tecott LH, Sun LM, Akana SF, Strack AM, Lowenstein DH, Dallman MF, et al. Eating disorder and epilepsy in mice lacking 5-HT2c serotonin receptors. Nature. 1995;374:542–6.
106. Vickers SP, Easton N, Webster LJ, Wyatt A, Bickerdike MJ, Dourish CT, et al. Oral administration of the 5-HT2C receptor agonist, mCPP, reduces body weight gain in rats over 28 days as a result of maintained hypophagia. Psychopharmacology. 2003;167(3):274–80.
107. Vimaleswaran K, Zhoa J, Wainwright N, Surtees P, Wareham N, Loos R. Assocaition between serotonin 5-HT-2C receptor gene (HTR2C) polymorphism and obesity- and mental health-related phenotypes in a large population-based cohort. International Journal of Obesity. 2010;34(6):1028-33.

108. 1000 Genomes Project [Internet]. ss1555817030 NCBI dbSNP Short Genetic Variations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1358325909.
109. Alcoholism: Clinical & Experimental Research. "G Allele Of Mu-opioid Receptor Gene Is Linked To Craving For Alcohol ScienceDaily: ScienceDaily; 2007 [01 October 2017]. Available from: www.sciencedaily.com/releases/2007/01/070103201526.htm.
110. Oslin DW, Berrettini W, Kranzler HR, Pettinati H, Gelernter J, Volpicelli JR, et al. A functional polymorphism of the mu-opioid receptor gene is associated with naltrexone response in alcohol-dependent patients. Neuropsychopharmacology. 2003;28(8):1546-52.
111. Ray LA, Barr CS, Blendy JA, Oslin D, Goldman D, Anton RF. The role of the Asn40Asp polymorphism of the mu opioid receptor gene (OPRM1) on alcoholism etiology and treatment: a critical review. Alcoholism: Clinical and Experimental Research. 2012;36(3):385-94.
112. Ray LA, Hutchison KE. A Polymorphism of the u-Opioid Receptor Gene (OPRM1) and Sensitivity to the Effects of Alcohol in Humans. Alcoholism: Clinical & Experimental Research. 2004;28(12):1789-95.
113. Project G [Internet]. ss1322962066 NCBI dbSNP Short Genetic Vartiations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1322962066.
114. Uhart M, McCaul ME, Oswald LM, Choi L, Wand GS. GABRA6 gene polymorphism and an attenuated stress response. Molecular Psychiatry. 2004;9:998-1006.
115. Pham X, Sun C, Chen X, van den Oord EJ, Neale MC, Kendler KS, et al. Association study between GABA receptor genes and anxiety spectrum disorders. Depression and Anxiety. 2009;26(11):998–1003.
116. Skilbeck KJ, Johnston GA, Hinton T. Stress and GABA receptors. Journal of Neurochemistry. 2010;112(5):1115-30.
117. Uhart M, McCaul ME, Oswald LM, Choi L, Wand GS. GABRA6 gene polymorphism and an attenuated stress response. Molecular Psychiatry. 2004;9(11):998-1006.

118. 1000 Genomes Project [Internet]. ss1318056825 NCBI dbSNP Short Genetic Variations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1318056825.
119. Pohjalainen T, Rinne JO, Någren K, Lehikoinen P, Anttila K, Syvälahti EK, et al. The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. Molecular Psychiatry. 1998;3(3):256-60.
120. Donofry SD. Multilocus Dopamine Gene Variation, Reward Sensitivity and Aberrant Eating: University of Pittsburgh; 2014.
121. He Q, Xue G, Chen C, Lu ZL, Chen C, Lei X, et al. COMT Val158Met polymorphism interacts with stressful life events and parental warmth to influence decision making. Scientific Reports. 2012;2:677.
122. Downs BW, Chen AL, Chen TJ, Waite RL, Braverman ER, Kerner M, et al. Nutrigenomic targeting of carbohydrate craving behavior: can we manage obesity and aberrant craving behaviors with neurochemical pathway manipulation by Immunological Compatible Substances (nutrients) using a Genetic Positioning System (GPS) Map? Medical Hypotheses. 2009;73(3):427-34.
123. Wallace DL, Aarts E, d'Oleire Uquillas F, Dang LC, Greer SM, Jagust WJ, et al. Genotype status of the dopamine-related catechol-O-methyltransferase (COMT) gene corresponds with desirability of "unhealthy" foods. Appetite. 2015;92:74-80.
124. 1000 Genomes Project [Internet]. ss1343024241 NCBI dbSNP Short Genetic Variations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1343024241.
125. Stein DJ, Newman TK, Savitz J, Ramesar R. Warriors versus worriers: the role of COMT gene variants. CNS Spectrums. 2006;11(10):745-8.
126. Andersen S, Skorpen F. Variation in the COMT gene: implications for pain perception and pain treatment. Pharmacogenomics. 2009;10(4):669-84.
127. Belfer I, Segall S. COMT genetic variants and pain. Drugs Today. 2011;47(6):457-67.
128. Foulkes T, Wood JN. Pain genes. PLoS Genetics. 2008;4(7):e1000086.
129. Li W, Chen Y, Yin B, Zhang L. Pain in Parkinson's disease associated with COMT gene polymorphisms. Behavioural Neurology. 2014;2014:304203.
130. Rakvag TT, Ross JR, Sato H, Skorpen F, Kaasa S, Klepstad P. Genetic variation in the catechol-O-methyltransferase (COMT) gene and morphine requirements in cancer patients with pain. Molecular Pain. 2008;4:64.

131. Cohen J [Internet]. Worrier or Warrior? Explaining The COMT V158M Gene (rs4680) Selfhacked 2018 [21 October 2018]. Available from: https://www.selfhacked.com/blog/worrier-warrior-explaining-rs4680comt-v158m-gene/.
132. 1000 Genomes Project [Internet]. ss1366683122 NCBI dbSNP Short Genetic Variations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1366683122.
133. Benjamin J, Li L , Patterson C, Greenberg BD, Murphy DL, Hamer DH. Population and familial association between the D4 dopamine receptor gene and measures of Novelty Seeking. Nature Genetics. 1996;12:81–4.
134. Jo Lusher, Chris Chandler. Dopamine D4 receptor gene (DRD4) is associated with Novelty Seeking (NS) and substance abuse: the saga continues. Molecular Psychiatry. 2001;6:497–9.
135. Laucht M, Becker K, Blomeyer D, Schmidt MH. Novelty seeking involved in mediating the association between the dopamine D4 receptor gene exon III polymorphism and heavy drinking in male adolescents: results from a high-risk community sample. Biological Psychiatry. 2007;61(1):87-92.
136. Thomson CJ, Hanna CW, Carlson SR, Rupert Jl. The -521C/T variant in the dopamine-4-receptor gene (DRD4) is associated with skiing and snowboarding behavior. Scandinavian Journal of Medicine & Science in Sports. 2013;23(2):e108-13.
137. Derringer J, Krueger RF, Dick DM, Saccone S, Grucza RA, Agrawal A, et al. Predicting sensation seeking from dopamine genes. A candidate-system approach. Psychological Science. 2010;21(9):1282-90.
138. 1000 Genomes Project [Internet]. ss1339780819 NCBI dbSNP Short Genetic Variations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1339780819.
139. Yang MM, Wang J, Fan JJ, Nq TK, Sun DJ, Guo X, et al. Variations in the Obesity Gene "LEPR" contribute to risk of type 2 diabetes mellitus: Evidence from a Meta-Analysis. Journal of Diabetes Research. 2016;5412084.
140. Boumaiza I, Omezzine A, Rejeb J, Rebhi L, Ouedrani A, Ben Rejeb N, et al. Relationship between leptin G2548A and leptin receptor Q223R gene polymorphisms and obesity and metabolic syndrome risk in Tunisian volunteers. Genetic Test Molecular Biomarkers. 2012;16(7):726-33.

141. Murugesan D, Arunachalam T, Ramamurthy V, Subramanian S. Association of polymorphisms in leptin receptor gene with obesity and type 2 diabetes in the local population of Coimbatore. Indian Journal of Human Genetics. 2010;16(2):72–7.
142. Farooqi IS, O'Rahilly S. New advances in the genetics of early onset obesity. International Journal of Obesity. 2005;29(10):1149-52.
143. Hastuti P, Zukhrufia I, Padwaswari MH, Nuraini A, Sadewa AH. Polymorphism in leptin receptor gene was associated with obesity in Yogyakarta, Indonesia. Egyptian Journal of Medical Human Genetics. 2016;17(3):271-6.
144. Sahin S, Rustemoglu A, Tekcan A, Tasliyurt T, Guven H, Yigit S. Investigation of associations between obesity and LEP G2548A and LEPR 668A/G polymorphisms in a Turkish population. Disease Markers. 2013;35(6):673-7.
145. Su S, Zhang C, Zhang F, Li H, Yang X, Tang X. The association between leptin receptor gene polymorphisms and type 2 diabetes mellitus: A systematic review and meta-analysis. Diabetes Research and Clinical Practice. 2016;121:49-58.
146. Yang M, Peng S, Li W, Wan Z, Fan L, Du Y. Relationships between plasma leptin levels, leptin G2548A, leptin receptor Gln223Arg polymorphisms and gestational diabetes mellitus in Chinese population. Scientific Reports. 2016;6:23948.
147. 1000 Genomes Project [Internet]. ss1291186628 NCBI dbSNP Short Genetic Variations: NCBI; 2014 [21 June 2018]. Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ss.cgi?ss=ss1291186628.
148. Lowe MR. Self-Regulation of Energy Intake in the Prevention and Treatment of Obesity: Is It Feasible? Obesity Research. 2003;11(Supplement):44S-59S.
149. Peduzzi P, Concato J, Kemper E, Holford T, Feinstein A. A simultaion study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology. 1996;49(12):1373-9.
150. Applied Biosystems by Life Technologies. MagMAX™ DNA Multi-Sample Ultra Kit User Guide. ThermoFischer; 2016. p. 1-5.
151. Technologies ABbL. TaqMan® PreAmp Master Mix Kit Protocol. 2010.
152. Corporation LT. TaqMan OpenArray Genotyping User Manual. In: Biosystems A, editor. Part Number 4377476 Rev E2010. p. 1-158.
153. Furrows S, Ridgway G. Good laboratory practice'in diagnostic laboratories using nucleic acid amplification methods. Clinical Microbiology and Infection. 2001;7(5):227-9.

154. Biosoft P [Internet]. PCR Primer Design Guidelines: PCR (Polymerase Chain Reaction) 2018 [25 October 2018]. Available from: http://www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html.
155. Choi J, Kim J, Joe C, Kim S, Ha K, Y P. Improved cycle sequencing of GC-rich DNA template. Experimental and Molecular Medicine. 1999;31(1):20-4.
156. Tyrrell J, Anwar Mulugeta, Andrew R Wood, Ang Zhou, Robin N Beaumont, Marcus ATuke, et al. Using genetics to understand the causal influence of higher BMI on depression. International Journal of Epidemiology, . 2018;dyy223(https://doi.org/10.1093/ije/dyy223).
157. Barber N [Internet]. Why Women Feel Bad About Their Appearance: Lookism does not explain body image discontent. Psychology Today: Sussex Publishers; 2013 [updated 2013. 13 November 2018]. Available from: https://www.psychologytoday.com/intl/blog/the-human-beast/201305/why-women-feel-bad-about-their-appearance.
158. Resource ITIGS [Internet]. Providing ongoing support for the 1000 Genomes Project data 2008 [14 September]. Available from: http://www.internationalgenome.org/home.
159. Institute TNHG [Internet]. About the International HapMap Project 2012 [14 September 2018]. Available from: https://www.genome.gov/11511175/.
160. Fay DS, Gerow K. A Biologist's Guide to Statistical Thinking and Analysis: WormBook: The online review of C. eleganc Biology; 2013.
161. Yang MM, Wang J, Fan JJ, Ng TK, Sun DJ, Guo X, et al. Variations in the Obesity Gene "LEPR" Contribute to Risk of Type 2 Diabetes Mellitus: Evidence from a Meta-Analysis. Journal of Diabetes Research. 2016;5412084.
162. Yang MM, Wang J, Fan JJ, Ng TK, Sun DJ, Guo X, et al. Variations in the Obesity Gene "LEPR" Contribute to Risk of Type 2 Diabetes Mellitus: Evidence from a Meta-Analysis. Journal of Diabetes Research. 2016;2016:5412084.
163. Genetics Home Reference [Internet]. LEPR gene - Leptin Receptor: U.S National Library of Medicine; 2018 [24 September].
164. Sica GT. Bias in research studies. Radiology. 2006;238(3):780-9.
165. Smith J, Noble H. Bias in Research. BMJ Journals - Evidence-Based Nursing. 2014;17(4):100-2.
166. Pannucci CJ, Wilkins EG. Identifying and Avoiding Bias in Research. Plastic and Reconstructive Surgery 2010;126(2):619-25.

167. Šimundić AM. Bias in Research. Biochemia Medica. 2013; 23(1):12-5.
168. Motsoaledi PA. Strategy for the Prevention and Control of Obesity in South Africa 2015-2020. In: Health Do, editor. 2015. p. 57.
169. Economics Tax Analysis Chief Directorate. Taxation of Sugars Sweetended Beverages - Policy Paper. In: Treasury N, editor. Goverment Gazette: South African Government; 2016. p. 1-30.

APPENDICES

Appendix A1 – Medical Ethics Approval

The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.
- FWA 00002567, Approved dd 22 May 2002 and Expires 03/20/2022.
- IRB 0000 2235 IORG0001762 Approved dd 22/04/2014 and Expires 03/14/2020.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Health Sciences Research Ethics Committee

16/01/2018

Approval Certificate
New Application

Ethics Reference No: 515/2017

Title: The prevalence of eight single nucleotide variations in overweight and obese participants

Dear Ms Bianca Sansom

The **New Application** as supported by documents specified in your cover letter dated 27/11/2017 for your research received on the 4/12/2017, was approved by the Faculty of Health Sciences Research Ethics Committee on its quorate meeting of 16/01/2018.

Please note the following about your ethics approval:
- Ethics Approval is valid for 2 years
- Please remember to use your protocol number (**515/2017**) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, or monitor the conduct of your research.

Ethics approval is subject to the following:
- The ethics approval is conditional on the receipt of **6 monthly written Progress Reports**, and
- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

We wish you the best with your research.

Yours sincerely

Dr R Sommers; MBChB; MMed (Int); MPharMed,PhD
Deputy Chairperson of the Faculty of Health Sciences Research Ethics Committee, University of Pretoria

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes, Second Edition 2015 (Department of Health).

☎ 012 356 3078 fhsethics@up.ac.za http://www.up.ac.za/healthethics
✉ Private Bag X323, Arcadia, 0007 - Tswelopele Building, Level 4, Room 61, Gezina, Pretoria

Appendix A2 - Permission letter

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Department of Pharmacology
Faculty of Health Sciences

PERMISSION TO ACCESS PATIENTS AT PRIVATE FAMILY MEDICAL CLINIC

To: Dr Andre Marais
Private Practice
Doornpark Sentrum
Airport Road
Doornport, Pretoria
0186

From: The Investigator
Ms Bianca Sansom
Department of Pharmacology
University of Pretoria
Bianca.sansom86@gmail.com
012 942 9602/ 074 671 4143

Re: Permission to do research at Private Family Medical Clinic

I am a researcher working at the Department of Pharmacology, University of Pretoria. I am requesting permission on behalf of myself, supervisor and co-supervisor of my project to conduct a study at your facility that involves access to patients who will form part of either my control group or study group. The control group will consist of patients with a normal body weight (BMI between 18.5 - 24.9 kg/m^2). The study group will consist of patients who suffer from overweight (BMI between 25 kg/m^2 to 29.9kg/m^2), obesity (BMI between 30 kg/m^2 to 39.9kg/m^2) and morbid obesity (BMI between greater than 40 kg/m^2). .

This is done concerning my MSc degree in Pharmacology. The title of the study is:

The prevalence of eight single nucleotide variations in overweight and obese participants

We hypothesize that individuals harbouring certain genes curbed towards addictive behaviours, such as food cravings and overeating, have a genetic predisposition to becoming obese and/or suffering from depression. It has long been known that there is a link between obesity and depression, recent studies have also shown there to be a genetic component. The interaction between gene combinations and the environment have a large part to play in the expression of the genes. By establishing a pattern, a better understanding into the cause of obesity can be gained, leading to novel treatment strategies.

Ninety-seven participants from your facility will constitute the control group (normal weight, BMI 18.5-25kg/m^2), which will be compared to the study groups of patients (overweight - BMI between 25 kg/m^2 to 29.9kg/m^2, obese - BMI between 30 kg/m^2 to 39.9kg/m^2 and morbidly obese - BMI between greater than 40 kg/m^2). DNA will be collected from all participants by using a cotton buccal swab. The study will be conducted in accordance with the Declaration of Helsinki, and we undertake not to proceed with the study until we have received approval from the Faculty of Health Sciences Research Ethics Committee, University of Pretoria. All personal and identifiable patient information will be kept confidential, and will only be used by the primary investigator to contact a

patient in the event where additional data might be required. Data emanating from this study will therefore be published anonymously.

Yours sincerely

Ms Bianca Sansom

Permission to do the research study at this facility and to access the admitted patients as requested, is hereby approved.

Dr Andre Marais
Private Practice
Doornpark Sentrum
Airport Road
Doornport, Pretoria
0186

9/11/2017.

Signature Date

Hospital Official Stamp

Dr Andre Marais
MB.ChB, B.Pharm, MSc.Pharm
AC.Clin Pharm (SA)
CLINICAL PHARMACOLOGIST
PR 0507547, MP 0680818

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Department of Pharmacology
Faculty of Health Sciences

PERMISSION TO ACCESS PATIENTS AT SLIMMING CLINIC

To: Dr Cecile Baard
Private Practice
447 May Street
Brooklyn
Pretoria
0186

From: The Investigator
Ms Bianca Sansom
Department of Pharmacology
University of Pretoria
Bianca.sansom86@gmail.com
012 942 9602/ 074 671 4143

Re: Permission to do research at Slimming Clinic

I am a researcher working at the Department of Pharmacology, University of Pretoria. I am requesting permission on behalf of myself, supervisor and co-supervisor of my project to conduct a study at your facility that involves access to patients who suffer from overweight (BMI between 25 kg/m^2 to 29.9kg/m^2), obesity (BMI between 30 kg/m^2 to 39.9kg/m^2) and morbid obesity (BMI between greater than 40 kg/m^2).

This is done concerning my MSc degree in Pharmacology. The title of the study is:

The prevalence of nine single nucleotide polymorphisms in overweight and obese participants.

We hypothesize that individuals harbouring certain genes curbed towards addictive behaviours, such as food cravings and overeating, have a genetic predisposition to becoming obese. The neurotransmitters associated with the brain reward cascade, have been shown to have an effect on one's mood, food craving and feeling of satiety. The interaction between gene combinations and the environment have a large part to play in the expression of the genes. By establishing a pattern, a better understanding into the cause of obesity can be gained, leading to novel treatment strategies.

Ninety-seven participants from your facility will constitute the study group (overweight - BMI between 25 kg/m^2 to 29.9kg/m^2, obese - BMI between 30 kg/m^2 to 39.9kg/m^2 and morbidly obese - BMI between greater than 40 kg/m^2), which will be compared to a control group of participants (97 normal weight participant with a BMI between 18.5 and 24.9 kg/m2). DNA will be collected from all participants by using a cotton buccal swab. The study will be conducted in accordance with the Declaration of Helsinki, and we undertake not to proceed with the study until we have received approval from the Faculty of Health Sciences

Research Ethics Committee, University of Pretoria. All personal and identifiable patient information will be kept confidential, and will only be used by the primary investigator to contact a patient in the event where additional data might be required. Data emanating from this study will therefore be published anonymously.

Yours sincerely

Ms Bianca Sansom

Permission to do the research study at this facility and to access the admitted patients as requested, is hereby approved.

Dr Cecile Baard
Private Practice
447 May Street
Brooklyn
Pretoria
0186

Baard 3/11/17

Signature Date

Official Practice Stamp/Other Official Proof

Appendix A3 - Patient information leaflet

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

PATIENT/PARTICIPANT INFORMATION LEAFLET & INFORMED CONSENT

Name of Principle Investigator: Bianca Sansom
Contact details (daytime and afterhours): 074 671 4143 or bianca@geneway.co.za
Name of Institute: Department of Pharmacology, University of Pretoria
Degree: MSc Pharmacology
Supervisor & Co-supervisor: Dr Andre Marais and Dr Christa North.

Date and Time of First Informed Consent Discussion:

Date: _____ Time: _____

Dear Participant,

Title: The prevalence of eight single nucleotide variations in overweight and obese participants

I, Bianca Sansom, am an MSc student in the Department of Pharmacology undertaking a research project in order to obtain my master's degree in Pharmacology. You are invited to partake in my research project as a voluntary participant to investigate the prevalence of several gene variations relating to the brain reward cascade and food cravings that may leading to weigh gain and obesity.

This letter provides the necessary information to equip participants with the knowledge in order to make an informed decision whether or not to partake in this study. It should be noted that before accepting to participate in the study, one must have read and understood the following explanation of the purpose of the study, the study procedures, benefits, risks and precautions as well as your right to withdraw from the study at any time.

If you have any questions with regards to any aspect of the study or the document, please do not hesitate to ask me.
If you are not satisfied about the procedure involved in the research, you are not obligated to participate.
If you are satisfied with the procedures and agree to participate in this study, you will be asked to sign the informed consent document to confirm that you understand the study and your role within the research. You will also be given a copy of this document to keep.

Purpose of the Study
Obesity is a serious epidemic in South Africa and is the leading cause of major health problems such as Type II diabetes, cardiovascular disease and cholesterol. Although

overeating is a major contributing factor to becoming over weight and obese, there also appears to be a genetic link involved. This project is aimed at studying the relationship between several gene variations that have been associated with addictive behaviours, including overeating and/or binge eating, and the brain's reward cascade. By investigating single nucleotide changes in the DNA, we will be able to establish whether a genetic link exists and establish better treatment strategies. Should you accept to participate in the study a cheek swab will be taken, the DNA extracted and genotyped.

What does participation in the study entail?
We will require to provide a DNA sample (1 x buccal/cheek swab). If you agree, you will be asked to complete and sign the informed consent form, and provide basic information namely; age, gender, race, anthropometrics and current medication usage. Although your information will be available to me, you will not be identified in any collection of data. The processing of the samples will be handled by myself and any data generated will therefore be anonymous and confidential.

You have the right to privacy. All samples and information will be dealt with in confidence. Information regarding your participation will be kept in locked files. Your sample will be labelled with a number to ensure your privacy. All samples will be stored in the laboratory for the duration of the study. The principle investigators involved in the study will not disclose any medical information and results to any external party. The information and data generated will only be used for statistical purposes, including publication of results.

You have the right to agree or refuse to participate in this research. If, however you decided to participate, you are free to withdraw from the study at any intervening time thereafter.
Refusal to participate will not result in any penalty or loss of benefits to which you are entitled. Your refusal to participate will not affect your legal rights or health care that you are receiving.

Possible benefits of this study
There will be no direct benefit to you, however your participation would enable us to gain a greater insight into the ever growing epidemic. The collection of the DNA sample and molecular testing done thereafter will be at no cost you or your medical aid. You will also not receive any financial benefit for partaking in the study.

Ethical Approval
The research study protocol has been granted approval by the Faculty of Health Science Research Ethics Committee, University of Pretoria.

They may be contacted on 012 356 3085, fax 086 609 5893 or email manda.smith@up.ac.za. You may contact myself on 074 671 4143, or email bianca@geneway.co.za, should you have any questions relating to this study.

Your participation is sincerely appreciated.
Kind regards,

Bianca Sansom

Appendix A4 - Informed consent

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

INFORMED CONSENT

Informed consent form for patients/participants partaking in the research project titled: The prevalence of eight single nucleotide variations in overweight and obese participants

Name of Principle Investigator: Bianca Sansom
Contact details: 074 671 4143 or bianca@geneway.co.za
Name of Institute: Department of Pharmacology, University of Pretoria
Degree: MSc Pharmacology
Supervisor & Co-supervisor: Dr Andre Marais and Dr Christa North.

I, the participant, have been invited to partake in this research project to investigate the prevalence of eight genetic variations associated with the brain reward cascade in overweight and obese individuals.

I have read the patient information leaflet or it has been read to me, and I have received a copy thereof. I agree to supply the following information:

Participant Reference No.:		**Medication:**	
Age:		**Height:**	
Gender:		**Weight:**	
Ethnicity:		**BMI Score:**	

Your signature below will certify the following:

1. That you have read the information provided in the information leaflet.
2. That you have understood the contents of the consent form
3. That you have received feedback and answers to all your questions
4. That you have freely decided to participate in this research
5. That you are aware that your legal rights will not be affected.

Declaration

I have received, read and understood the contents of the document and the information regarding the study. I have no further questions and I am prepared to participate in the study.

Name of Participant (Print): _____

Signature of Participant: _____ **Date:** _____

Contact details of Participant: _____

Witness name: _____

Witness Signature: _____**Date:** _____

VERBAL PATIENT INFORMED CONSENT

I, the undersigned, Dr _____ have read and have explained fully to the patient, name _____ and/or his/her relative, the patient informed leaflet, which has indicated the nature and purpose of the study in which I have asked the patient to participate. The explanation I have given has mentioned both the possible risks and benefits of the study and the alternative treatment available for his/her illness. The patient indicated that he/she understands that he/she will be free to withdraw from the study at any time for any reason and without jeopardizing his/her treatment.

I hereby certify that the patient has agreed to participate in this study.

Name of Participant (Print): _____

Signature of Participant: _____ **Date:** _____

Contact details of Participant: _____

Witness's Name _____

Witness's Signature _____ **Date:** _____

RESEARCHER'S STATEMENT

Declaration

I have accurately read out the information sheet to the potential participant, and to the best of my ability made sure that the participant understood the content of the information leaflet.

I confirm that the participant has been given the opportunity to ask questions about the study and that all questions asked were answered to the best of my knowledge and ability.

I confirm that the participant has not been coerced into giving consent and participating in this study. Consent by the participant was given freely and voluntarily.

A copy of the information leaflet and consent form was given to the participant.

Name of Researcher: Bianca Sansom

Signature: _____

Date: _____

Appendix B1: BLAST® search of FASTA sequence for Quality Control check

1. Target Sequence:

CACAGCCATCCTCAAAGTGCTGGTC**[A/G]**AGGCAGGCGCCCAGCTGGACGTCCA

rs1800497 *DRD2*

BLAST® search results:

Homo sapiens ankyrin repeat and kinase domain containing 1 (ANKK1), RefSeqGene on chromosome 11
Sequence ID: NG_012976.1Length: 19628Number of Matches: 2

Related Information
Range 1: 16816 to 17816GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1845 bits(999)	0.0	1000/1001(99%)	0/1001(0%)

```
Query  1      TCCAGGCGAGAGGCCCCAAGTAGTCTAAAtttctttctttctttctttttttATATGGAGT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17816  TCCAGGCGAGAGGCCCCAAGTAGTCTAAATTTCTTTCTTTCTTTCTTTTTTATATGGAGT  17757

Query  61     CTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCTGC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17756  CTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCTGC  17697

Query  121    CTCCTGGGTTCAAGGAATTCTCCTGCCTCAGCCTCCCTGGTAGTTGGGATTACAGGCACG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17696  CTCCTGGGTTCAAGGAATTCTCCTGCCTCAGCCTCCCTGGTAGTTGGGATTACAGGCACG  17637

Query  181    TGCCACCATACCCAGCTAAATTTTGTATTTTTAGCAGAGACAGGGTTTTGCCATGTTGGC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17636  TGCCACCATACCCAGCTAAATTTTGTATTTTTAGCAGAGACAGGGTTTTGCCATGTTGGC  17577

Query  241    CAGGCTGGCCTCAAACTCTTGATATCAGGTGATCTGCCTGCCTCAGCCTCCCAAAGTGCT  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17576  CAGGCTGGCCTCAAACTCTTGATATCAGGTGATCTGCCTGCCTCAGCCTCCCAAAGTGCT  17517

Query  301    GGGATTACAGACGTGAGCCACCACGGCTGGCCAAGTTGTCTAAATTTCCATCTCGGCTCC  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17516  GGGATTACAGACGTGAGCCACCACGGCTGGCCAAGTTGTCTAAATTTCCATCTCGGCTCC  17457

Query  361    TGGCTTAGAACCACCCAGAGTGGCCACTGACGGCTCCTTGCCCTCTAGGAAGGACATGAT  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  17456  TGGCTTAGAACCACCCAGAGTGGCCACTGACGGCTCCTTGCCCTCTAGGAAGGACATGAT  17397

Query  421    GCCCTGCTTTCGGCTGCGGAGGGCCAGTTGCAGGGGTGTGCAGCTCACTCCATCCTGGAC  480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17396  GCCCTGCTTTCGGCTGCGGAGGGCCAGTTGCAGGGGTGTGCAGCTCACTCCATCCTGGAC  17337

Query  481    GTCCAGCTGGGCGCCTGCCTYGACCAGCACTTTGAGGATGGCTGTGTTGCCCTTGAGGGC  540
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  17336  GTCCAGCTGGGCGCCTGCCTCGACCAGCACTTTGAGGATGGCTGTGTTGCCCTTGAGGGC  17277

Query  541    GGCCAGGTGGGCGGGTGTCCAGCCCACCTTGTTGCGGGCGTGGACATTTGCGTGATGTTC  600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17276  GGCCAGGTGGGCGGGTGTCCAGCCCACCTTGTTGCGGGCGTGGACATTTGCGTGATGTTC  17217

Query  601    TAGGAGGTTGATGACACTCAGGAAGGTGCTCCTCTGGACCGCCAGGTGGAGGGGTGTCCA  660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17216  TAGGAGGTTGATGACACTCAGGAAGGTGCTCCTCTGGACCGCCAGGTGGAGGGGTGTCCA  17157

Query  661    GCCTGACTGCTCTGCAGCATTGGGGTCAGCCCCACACTGCAGCAGTGCTGACACCACCGC  720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17156  GCCTGACTGCTCTGCAGCATTGGGGTCAGCCCCACACTGCAGCAGTGCTGACACCACCGC  17097

Query  721    CTCCTCCCCGTGGCGTGCAGCTAGGTGCAGGGGAGTCCAGTTCACAGCTCCAAGAGCACC  780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17096  CTCCTCCCCGTGGCGTGCAGCTAGGTGCAGGGGAGTCCAGTTCACAGCTCCAAGAGCACC  17037

Query  781    CATGTTTGCGTGGCTCTCTGCCAGCAGATGGATGATCTCCAGGTGGCCCTTGTAGGCTGC  840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17036  CATGTTTGCGTGGCTCTCTGCCAGCAGATGGATGATCTCCAGGTGGCCCTTGTAGGCTGC  16977

Query  841    TAGATGCAGGGGTGTCCAGCCCTGGTGGGTGGGCAGCTCAAGGCTGGCTCCGTACCTGAG  900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16976  TAGATGCAGGGGTGTCCAGCCCTGGTGGGTGGGCAGCTCAAGGCTGGCTCCGTACCTGAG  16917

Query  901    CAGCATCTTGCAGATCAGGTATTTGCCCCTGGCAGCTGCAGTGTGCAGTGGGCCGTAGCC  960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  16916  CAGCATCTTGCAGATCAGGTATTTGCCCCTGGCAGCTGCAGTGTGCAGTGGGCCGTAGCC  16857

Query  961    GCTCTGGTCAAGGGCATCAGGGACCGCTCCACTCTTCAGCA  1001
               |||||||||||||||||||||||||||||||||||||||||
Sbjct  16856  GCTCTGGTCAAGGGCATCAGGGACCGCTCCACTCTTCAGCA  16816
```

2. Target Sequence:

CTCGCGGCATCCCCCCTGCACCCCC**[A/G]**GCATCCCCCCTGCAGCCCCCCCAGC

rs25531 *SLC6A4*

BLAST® search results:

Homo sapiens solute carrier family 6 member 4 (SLC6A4), RefSeqGene on chromosome 17
Sequence ID: NG_011747.2Length: 48618Number of Matches: 1
Related Information
Range 1: 3432 to 4319GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1635 bits(885)	0.0	888/889(99%)	1/889(0%)

```
Query  1     TCTCCCGCCTGGCGTTGCCGCTCTGAATGCCAGCACCTAACCCCTAATGTCCCTACTGCA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3432  TCTCCCGCCTGGCGTTGCCGCTCTGAATGCCAGCACCTAACCCCTAATGTCCCTACTGCA  3491

Query  61    GCCCTCCCAGCATCCCCCCTGCAACCTCCCAGCAACTCCCTGTACCCCTCCTAGGATCGC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3492  GCCCTCCCAGCATCCCCCCTGCAACCTCCCAGCAACTCCCTGTACCCCTCCTAGGATCGC  3551

Query  121   TCCTGCATccccattatccccccttcacccctcgcggcatccccctgcaccccccagg  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct  3552  TCCTGCATCCCCCATTATCCCCCCCTTCACCCCTCGCGGCATCCCCCCTGCACCCCCA-G  3610

Query  181   catccccccctgcagccccccagcatctcccctgcaccccagcatccccccctgcagccc  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3611  CATCCCCCCTGCAGCCCCCCCAGCATCTCCCCTGCACCCCCAGCATCCCCCCTGCAGCCC  3670

Query  241   ttccagcatccccctgcacctctcccaggatctcccctgcaaccccccattatccccccctg  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3671  TTCCAGCATCCCCCTGCACCTCTCCCAGGATCTCCCCTGCAACCCCCCATTATCCCCCCTG  3730

Query  301   caccccctcgcagtatccccccctgcacccccccagcatcccccccatgcacccccggcatccc  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3731  CACCCCTCGCAGTATCCCCCCTGCACCCCCCCAGCATCCCCCCCATGCACCCCCGGCATCCC  3790

Query  361   ccctgcaccccTCCAGCATTCTCCTTGCACCCTACCAGTATTCCCCCGCATCCCGGCCTC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3791  CCCTGCACCCCTCCAGCATTCTCCTTGCACCCTACCAGTATTCCCCCGCATCCCGGCCTC  3850

Query  421   CAAGCCTCCCGCCCACCTTGCGGTCCCCGCCCTGGCGTCTAGGTGGCACCAGAATCCCGC  480
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3851 CAAGCCTCCCGCCCACCTTGCGGTCCCCGCCCTGGCGTCTAGGTGGCACCAGAATCCCGC 3910

Query  481  GCGGACTCCACCCGCTGGGAGCTGCCCTCGCTTGCCCGTGGTTGTCCAGCTCAGTCCCTC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3911 GCGGACTCCACCCGCTGGGAGCTGCCCTCGCTTGCCCGTGGTTGTCCAGCTCAGTCCCTC 3970

Query  541  TAGACGCTCAGCCCAACCGGCCGCACAGTTTTCAGGGGTCAGTTCCTCCAAGTACAAGGG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3971 TAGACGCTCAGCCCAACCGGCCGCACAGTTTTCAGGGGTCAGTTCCTCCAAGTACAAGGG 4030

Query  601  GCGGTGGCTTCTCTGGAGCTGCAAACTTGTCACTGCTATTTCCTTTCGGTCTTCTACTTC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4031 GCGGTGGCTTCTCTGGAGCTGCAAACTTGTCACTGCTATTTCCTTTCGGTCTTCTACTTC 4090

Query  661  CTATCGTTCCTGGCCTCCTCTTGGGGAGAGGTAGAGCCCTCTCCTTTCCGCCTCAGGGAC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4091 CTATCGTTCCTGGCCTCCTCTTGGGGAGAGGTAGAGCCCTCTCCTTTCCGCCTCAGGGAC 4150

Query  721  AACCCAAAGCAAGTACTGCATGTGCCCttttaaagttttaaataatttagcaaaaagg 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4151 AACCCAAAGCAAGTACTGCATGTGCCCTTTTTAAAGTTTTAAATAATTTTAGCAAAAAGG 4210

Query  781  atattaacattaaatcaatttttaaacttttttgaaaaaattATCAAAACTACATGCACAT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4211 ATATTAACATTAAATCAATTTTTAAACTTTTTGAAAAAATTATCAAAACTACATGCACAT 4270

Query  841  GGTTCAAAACAATAGGCTCCTGCTGGGCCCTTTCAGATAATTCAAATTG  889
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4271 GGTTCAAAACAATAGGCTCCTGCTGGGCCCTTTCAGATAATTCAAATTG  4319
```

3. Target Sequence:

GGTCAACTTGTCCCACTTAGATGGC**[A/G]**ACCTGTCCGACCCATGCGGTCCGAA

rs1799971 *OPRM1*

BLAST® search results:

Homo sapiens opioid receptor mu 1 (OPRM1), RefSeqGene on chromosome 6
Sequence ID: NG_021208.2Length: 243367Number of Matches: 1
Related Information
Range 1: 33662 to 34662GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1845 bits(999)	0.0	1000/1001(99%)	0/1001(0%)

```
Query  1      TGTGTTTGCACAGAAGAGTGCCCAGTGAAGAGACCTACTCCTTGGATCGCTTTGCGCAAA  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  33662  TGTGTTTGCACAGAAGAGTGCCCAGTGAAGAGACCTACTCCTTGGATCGCTTTGCGCAAA  33721

Query  61     ATCCAccccttttccctcctccctccttccagcctccgaatcccgcatggcccacgctc  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  33722  ATCCACCCCTTTTCCCTCCTCCCTCCCTTCCAGCCTCCGAATCCCGCATGGCCCACGCTC  33781

Query  121    ccctcctGCAGCGGTGCGGGGCAGGTGATGAGCCTCTGTGAACTACTAAGGTGGGAGGGG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  33782  CCCTCCTGCAGCGGTGCGGGGCAGGTGATGAGCCTCTGTGAACTACTAAGGTGGGAGGGG  33841

Query  181    GCTATACGCAGAGGAGAATGTCAGATGCTCAGCTCGGTCCCCTCCGCCTGACGCTCCTCT  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  33842  GCTATACGCAGAGGAGAATGTCAGATGCTCAGCTCGGTCCCCTCCGCCTGACGCTCCTCT  33901

Query  241    CTGTCTCAGCCAGGACTGGTTTCTGTAAGAAACAGCAGGAGCTGTGGCAGCGGCGAAAGG  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  33902  CTGTCTCAGCCAGGACTGGTTTCTGTAAGAAACAGCAGGAGCTGTGGCAGCGGCGAAAGG  33961

Query  301    AAGCGGCTGAGGCGCTTGGAACCCGAAAAGTCTCGGTGCTCCTGGCTACCTCGCACAGCG  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  33962  AAGCGGCTGAGGCGCTTGGAACCCGAAAAGTCTCGGTGCTCCTGGCTACCTCGCACAGCG  34021

Query  361    GTGCCCGCCCGGCCGTCAGTACCATGGACAGCAGCGCTGCCCCCACGAACGCCAGCAATT  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34022  GTGCCCGCCCGGCCGTCAGTACCATGGACAGCAGCGCTGCCCCCACGAACGCCAGCAATT  34081

Query  421    GCACTGATGCCTTGGCGTACTCAAGTTGCTCCCCAGCACCCAGCCCCGGTTCCTGGGTCA  480
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34082    GCACTGATGCCTTGGCGTACTCAAGTTGCTCCCCAGCACCCAGCCCCGGTTCCTGGGTCA  34141

Query  481      ACTTGTCCCACTTAGATGGCRACCTGTCCGACCCATGCGGTCCGAACCGCACCGACCTGG  540
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  34142    ACTTGTCCCACTTAGATGGCAACCTGTCCGACCCATGCGGTCCGAACCGCACCGACCTGG  34201

Query  541      GCGGGAGAGACAGCCTGTGCCCTCCGACCGGCAGTCCCTCCATGATCACGGCCATCACGA  600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34202    GCGGGAGAGACAGCCTGTGCCCTCCGACCGGCAGTCCCTCCATGATCACGGCCATCACGA  34261

Query  601      TCATGGCCCTCTACTCCATCGTGTGCGTGGTGGGGCTCTTCGGAAACTTCCTGGTCATGT  660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34262    TCATGGCCCTCTACTCCATCGTGTGCGTGGTGGGGCTCTTCGGAAACTTCCTGGTCATGT  34321

Query  661      ATGTGATTGTCAGGTAAGGAAAGCGCCAGGGCTCCGAGCGGAGGGTTCAGCGGCTTAAGG  720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34322    ATGTGATTGTCAGGTAAGGAAAGCGCCAGGGCTCCGAGCGGAGGGTTCAGCGGCTTAAGG  34381

Query  721      GGGTACAAAGAGACACCTAACTCCCAAGGCTCAATGTTGGGCGGGAGGATGAAAGAGGGG  780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34382    GGGTACAAAGAGACACCTAACTCCCAAGGCTCAATGTTGGGCGGGAGGATGAAAGAGGGG  34441

Query  781      AGGTAAACTGGGGGGACTCTGGAGGAGACCACGGACAGTGATTGTTATTTCTATGAGAAA  840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34442    AGGTAAACTGGGGGGACTCTGGAGGAGACCACGGACAGTGATTGTTATTTCTATGAGAAA  34501

Query  841      ACCTACTTTTCTGTTTTTTCTTCAACTGATAAAGAAAGAATTCAAAATTTCAGGAGCAGA  900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34502    ACCTACTTTTCTGTTTTTTCTTCAACTGATAAAGAAAGAATTCAAAATTTCAGGAGCAGA  34561

Query  901      GAAGTTGCTTTGGTAAAAGCTACAAATGTCTAGGGGTGGGGGGCGGAGGGAAGCTATAGC  960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  34562    GAAGTTGCTTTGGTAAAAGCTACAAATGTCTAGGGGTGGGGGGCGGAGGGAAGCTATAGC  34621

Query  961      ATAGACTTGGAGCGCTTCCTTATACTGAGCAAAGAGGGCTC  1001
                |||||||||||||||||||||||||||||||||||||||||
Sbjct  34622    ATAGACTTGGAGCGCTTCCTTATACTGAGCAAAGAGGGCTC  34662
```

4. Target Sequence:

CCAGCGGATGGTGGATTTCGCTGGC**[A/G]**TGAAGGACAAGGTGTGCATGCCTGA

rs4680 *COMT*

BLAST® search results:

Homo sapiens catechol-O-methyltransferase (COMT), RefSeqGene (LRG_1010) on chromosome 22
Sequence ID: NG_011526.1Length: 35236Number of Matches: 1
Related Information
Range 1: 26509 to 27509GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1845 bits(999)	0.0	1000/1001(99%)	0/1001(0%)

```
Query  1      AGAGGGCAGCTCTGTGTTAGGACACACTGGGGCCAGCCAGGAAGGGTGGAAAAGATAGGG  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26509  AGAGGGCAGCTCTGTGTTAGGACACACTGGGGCCAGCCAGGAAGGGTGGAAAAGATAGGG  26568

Query  61     ACCAGCGTGAGCATAGAGGCTAAGGGACCATGGGAGCTCCAAGCGCGCTCACAGTGGGGA  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26569  ACCAGCGTGAGCATAGAGGCTAAGGGACCATGGGAGCTCCAAGCGCGCTCACAGTGGGGA  26628

Query  121    CCAGGTCCTGGGGGCTGGGGACACCAGGGAGGTGAAATACCCCTCCAGCGGGTAGGGAGG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26629  CCAGGTCCTGGGGGCTGGGGACACCAGGGAGGTGAAATACCCCTCCAGCGGGTAGGGAGG  26688

Query  181    GTGGGCAGAGGAGGGCCAGCGGCCAGGCATTTGGGAGGGGCTCCTGCTCTTTGGGAGAGG  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26689  GTGGGCAGAGGAGGGCCAGCGGCCAGGCATTTGGGAGGGGCTCCTGCTCTTTGGGAGAGG  26748

Query  241    TGGGGGGCCGTGCCTGGGGATCCAAGTTCCCCTCTCTCCACCTGTGCTCACCTCTCCTCC  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26749  TGGGGGGCCGTGCCTGGGGATCCAAGTTCCCCTCTCTCCACCTGTGCTCACCTCTCCTCC  26808

Query  301    GTCCCCAACCCTGCACAGGCAAGATCGTGGACGCCGTGATTCAGGAGCACCAGCCCTCCG  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26809  GTCCCCAACCCTGCACAGGCAAGATCGTGGACGCCGTGATTCAGGAGCACCAGCCCTCCG  26868

Query  361    TGCTGCTGGAGCTGGGGGCCTACTGTGGCTACTCAGCTGTGCGCATGGCCCGCCTGCTGT  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26869  TGCTGCTGGAGCTGGGGGCCTACTGTGGCTACTCAGCTGTGCGCATGGCCCGCCTGCTGT  26928
```

```
Query  421    CACCAGGGGCGAGGCTCATCACCATCGAGATCAACCCCGACTGTGCCGCCATCACCCAGC  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  26929  CACCAGGGGCGAGGCTCATCACCATCGAGATCAACCCCGACTGTGCCGCCATCACCCAGC  26988

Query  481    GGATGGTGGATTTCGCTGGCRTGAAGGACAAGGTGTGCATGCCTGACCCGTTGTCAGACC  540
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  26989  GGATGGTGGATTTCGCTGGCGTGAAGGACAAGGTGTGCATGCCTGACCCGTTGTCAGACC  27048

Query  541    TGGAAAAAGGGCCGGCTGTGGGCAGGGAGGGCATGCGCACTTTGTCCTCCCCACCAGGTG  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27049  TGGAAAAAGGGCCGGCTGTGGGCAGGGAGGGCATGCGCACTTTGTCCTCCCCACCAGGTG  27108

Query  601    TTCACACCACGTTCACTGAAAACCCACTATCACCAGGCCCCTCAGTGCTTCCCAGCCTGG  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27109  TTCACACCACGTTCACTGAAAACCCACTATCACCAGGCCCCTCAGTGCTTCCCAGCCTGG  27168

Query  661    GGCTGAGGAAAGAcccccccAGCAGCTCAGTGAGGGTCTCACAGCTCTGGGTAAACTGCC  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27169  GGCTGAGGAAAGACCCCCCCCAGCAGCTCAGTGAGGGTCTCACAGCTCTGGGTAAACTGCC  27228

Query  721    AAGGTGGCACCAGGAGGGGCAGGGACAGAGTGGGGCCTTGTCATCCCAGAACCCTAAAGA  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27229  AAGGTGGCACCAGGAGGGGCAGGGACAGAGTGGGGCCTTGTCATCCCAGAACCCTAAAGA  27288

Query  781    AAACTGATGAATGCTTGTATGGGTGTGTAAAGATGGCCTCCTGTCTGTGTGGGCGTGGGC  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27289  AAACTGATGAATGCTTGTATGGGTGTGTAAAGATGGCCTCCTGTCTGTGTGGGCGTGGGC  27348

Query  841    ACTGACAGGCGCTGTTGTATAGGTGTGTAGGGATGGCCTCCTGTCTGTGAGGACGTGGGC  900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27349  ACTGACAGGCGCTGTTGTATAGGTGTGTAGGGATGGCCTCCTGTCTGTGAGGACGTGGGC  27408

Query  901    ACTGACAGGCGCTGTTCCAGGTCACCCTTGTGGTTGGAGCGTCCCAGGACATCATCCCCC  960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  27409  ACTGACAGGCGCTGTTCCAGGTCACCCTTGTGGTTGGAGCGTCCCAGGACATCATCCCCC  27468

Query  961    AGCTGAAGAAGAAGTATGATGTGGACACACTGGACATGGTC  1001
              |||||||||||||||||||||||||||||||||||||||||
Sbjct  27469  AGCTGAAGAAGAAGTATGATGTGGACACACTGGACATGGTC  27509
```

5. Target Sequence:

GGGCAGGGGGAGCGGGCGTGGAGGG**[C/T]**GCGCACGAGGTCGAGGCGAGTCC

rs1800955 *DRD4*

BLAST® search results:

Homo sapiens dopamine receptor D4 (DRD4), RefSeqGene on chromosome 11
Sequence ID: NG_021241.1Length: 10402Number of Matches: 1
Related Information
Gene-associated gene details
Range 1: 3980 to 4980GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1845 bits(999)	0.0	1000/1001(99%)	0/1001(0%)

```
Query  1     TGGGGTCCCACAGAGTGGTGCCCCCTTTTAGTGTCTTCTAGGCCCCTTAGTGACAGACTA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3980  TGGGGTCCCACAGAGTGGTGCCCCCTTTTAGTGTCTTCTAGGCCCCTTAGTGACAGACTA  4039

Query  61    CAGAAAATACCTCTCAGGTCACAGGTCACCCCTCTTTGGTGAAGAGTCCATAGAATTCTC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4040  CAGAAAATACCTCTCAGGTCACAGGTCACCCCTCTTTGGTGAAGAGTCCATAGAATTCTC  4099

Query  121   TGCTGCGCTTTGCAAGCACTTTCTCTTCTGCACGTTTGGAACCTACCCCGGCCTGTCGTG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4100  TGCTGCGCTTTGCAAGCACTTTCTCTTCTGCACGTTTGGAACCTACCCCGGCCTGTCGTG  4159

Query  181   TCTTTCTCCTGGCCTCCTCGCGAGCCGAACCTACTGTCCGGTCCCGGGACCCCCTGCCCA  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4160  TCTTTCTCCTGGCCTCCTCGCGAGCCGAACCTACTGTCCGGTCCCGGGACCCCCTGCCCA  4219

Query  241   GGGTCAGAGGGGCGCCTACCTAGCTCACGGTCTTGGGCCGGAGGGAATGGAGGAGGGAGC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4220  GGGTCAGAGGGGCGCCTACCTAGCTCACGGTCTTGGGCCGGAGGGAATGGAGGAGGGAGC  4279

Query  301   GGGGTCGACCGCTCAGCTGTCCGCCCAGTTTCGGAGGCGGCCACGCGAGGATCAACTGTG  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4280  GGGGTCGACCGCTCAGCTGTCCGCCCAGTTTCGGAGGCGGCCACGCGAGGATCAACTGTG  4339

Query  361   CAACGGGTGGGGCCGCGGCTGACCGTGGTGGTCGCGGGGGCTGAGGGCCAGAGGCTGCgg  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4340  CAACGGGTGGGGCCGCGGCTGACCGTGGTGGTCGCGGGGGCTGAGGGCCAGAGGCTGCGG  4399
```

```
Query  421   gggggggggcggcgggatgagctaggcgtcggcggttgagtcgggcgcggagtcgggggca  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4400  GGGGGGGGCGGCGGGATGAGCTAGGCGTCGGCGGTTGAGTCGGGCGCGGAGTCGGGGGCA  4459

Query  481   gggggagcgggcgtggagggbgcgcacgaggtcgaggcgagtccgcgggggaggcgggCA  540
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  4460  GGGGGAGCGGGCGTGGAGGGTGCGCACGAGGTCGAGGCGAGTCCGCGGGGGAGGCGGGCA  4519

Query  541   GAGCCTGAGCTCAGGTCTTTCTGCGTCTGGCGGAACGGGCCTGGGAGGGAGGTTTTGCCA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4520  GAGCCTGAGCTCAGGTCTTTCTGCGTCTGGCGGAACGGGCCTGGGAGGGAGGTTTTGCCA  4579

Query  601   GATACCAGGTGGACTAGGGTGAGCGCCCGAGGGCCGGGACGCACGCACGGGCCGGGTAGG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4580  GATACCAGGTGGACTAGGGTGAGCGCCCGAGGGCCGGGACGCACGCACGGGCCGGGTAGG  4639

Query  661   ATGGCGCTGGCGTCGATGCCCGCGCGCTTCAGGGCCTGGTCTGGCCGCCCCTCCATCCTT  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4640  ATGGCGCTGGCGTCGATGCCCGCGCGCTTCAGGGCCTGGTCTGGCCGCCCCTCCATCCTT  4699

Query  721   GTCGGTTTCTCGGGTCGCGGACCCCGCGCGGCGCCGGGCGATGCTGGCCTGCCCGTGGCC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4700  GTCGGTTTCTCGGGTCGCGGACCCCGCGCGGCGCCGGGCGATGCTGGCCTGCCCGTGGCC  4759

Query  781   ACCACCTCGCTTCATTCCCGTCTCTTTGGGCCGCCGCATTCGTCCACGTGCCCGTCTCTC  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4760  ACCACCTCGCTTCATTCCCGTCTCTTTGGGCCGCCGCATTCGTCCACGTGCCCGTCTCTC  4819

Query  841   CCTGCGCAAAATTCCAAGATGAGCAAATACTGGGCTCACGGTGGAGCGCCgcgggggccc  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4820  CCTGCGCAAAATTCCAAGATGAGCAAATACTGGGCTCACGGTGGAGCGCCGCGGGGGCCC  4879

Query  901   ccctgagccggggcgggtcgggggcgggaccagggtccggccggggcgtgcccgagggga  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4880  CCCTGAGCCGGGGCGGGTCGGGGGCGGGACCAGGGTCCGGCCGGGGCGTGCCCGAGGGGA  4939

Query  961   gggACTCCCCGGCTTGCGACCCGGCGTTGTCCGCGGTGCTC  1001
             |||||||||||||||||||||||||||||||||||||||||
Sbjct  4940  GGGACTCCCCGGCTTGCGACCCGGCGTTGTCCGCGGTGCTC  4980
```

6. Target Sequence:

CTGCTCTTGGCTCCTCCCCTCATCC**[C/T]**GCTTTTGGCCCAAGAGCGTGGTGCA

rs3813929 *5HTR2C*

BLAST® search results:

Homo sapiens 5-hydroxytryptamine receptor 2C (HTR2C), RefSeqGene on chromosome X
Sequence ID: NG_012082.2Length: 332977Number of Matches: 1
Related Information
PubChem BioAssay-bioactivity screening
Range 1: 4469 to 5463GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1796 bits(972)	0.0	992/1001(99%)	6/1001(0%)

```
Query  1     GACAAGGATGGGGAAGTGGGCCTTATAACAGGATTGTGGCCTTTGCGCACTCACCAAATG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4469  GACAAGGATGGGGAAGTGGGCCTTATAACAGGATTGTGGCCTTTGCGCACTCACCAAATG  4528

Query  61    TTTGACCCTGTGAGTGCCTCAGTTGCTACTGTTGGAAGAATGGGCAAGAGTCGGAACAGA  120
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  4529  TTTGACCCTGTGAGTGCCTCAGTTGCTACTGTTGGAGGAATGGGCAAGAGTCGGAACAGA  4588

Query  121   GACCCTTGAAGGGAGTTTCAAAGCTTGATGAAATTTGCAAGACTTGAGAATGCTGTTTGT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4589  GACCCTTGAAGGGAGTTTCAAAGCTTGATGAAATTTGCAAGACTTGAGAATGCTGTTTGT  4648

Query  181   TGAAATGAAATGTACAGGGgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtTTGGGGAGG  240
             ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Sbjct  4649  TGAAATGAAATGTACAGGG------GTGTGTGTGTGTGTGTGTGTGTTTGGGGAGG  4702

Query  241   GGTATGCTATGAATCTTTGAGGGTACATTCTTGAGAAAGCCTTCCCTCTCTCTCTATCCG  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4703  GGTATGCTATGAATCTTTGAGGGTACATTCTTGAGAAAGCCTTCCCTCTCTCTCTATCCG  4762

Query  301   GTGCCATGGCTGATCCTGGTTCCCCCTACTCTCTAGGCCTTGTGAATCAGATTAATCATC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4763  GTGCCATGGCTGATCCTGGTTCCCCCTACTCTCTAGGCCTTGTGAATCAGATTAATCATC  4822

Query  361   ACCCCCACCCCCATCTCCACCATGGGGTCTCGCGCCCCCTGCCAGCAGGCTCCAGATGCA  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4823  ACCCCCACCCCCATCTCCACCATGGGGTCTCGCGCCCCCTGCCAGCAGGCTCCAGATGCA  4882
```

```
Query  421   CTAAGAGACCGGTCCAAACAGGCCCGGGGGCCACGTAATGCTGAGTGCTGATTGGCTGCT  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4883  CTAAGAGACCGGTCCAAACAGGCCCGGGGGCCACGTAATGCTGAGTGCTGATTGGCTGCT  4942

Query  481   CTTGGCTCCTCCCCTCATCCBGCTTTTGGCCCAAGAGCGTGGTGCAGATTCACCCGCGCG  540
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  4943  CTTGGCTCCTCCCCTCATCCCGCTTTTGGCCCAAGAGCGTGGTGCAGATTCACCCGCGCG  5002

Query  541   AGGTAGGCGCTCTGGTGCTTGCGGAGGACGCTTCCTTCCTCAGATGCACCGATCTTCCCG  600
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  5003  AGGTAGGCGCTCTGGTGCTTGCCGAGGACGCTTCCTTCCTCAGATGCACCGATCTTCCCG  5062

Query  601   ATACTGCCTTTGGAGCGGCTAGATTGCTAGCCTTGGCTGCTCCATTGGCCTGCCTTGCCC  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5063  ATACTGCCTTTGGAGCGGCTAGATTGCTAGCCTTGGCTGCTCCATTGGCCTGCCTTGCCC  5122

Query  661   CTTACCTGCCGATTGCATATGAACTCTTCTTCTGTCTGTACATCGTTGTCGTCGGAGTCG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5123  CTTACCTGCCGATTGCATATGAACTCTTCTTCTGTCTGTACATCGTTGTCGTCGGAGTCG  5182

Query  721   TCGCGATCGTCGTGGCGCTCGTGTGATGGCCTTCGTCCGTTTAGAGTAGTGTAGTTAGTT  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5183  TCGCGATCGTCGTGGCGCTCGTGTGATGGCCTTCGTCCGTTTAGAGTAGTGTAGTTAGTT  5242

Query  781   AGGGGCCAACGAAGAAGAAAGAAGACGCGATTAGTGCAGAGATGCTGGAGGTGGTCAGTT  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5243  AGGGGCCAACGAAGAAGAAAGAAGACGCGATTAGTGCAGAGATGCTGGAGGTGGTCAGTT  5302

Query  841   ACTAAGCTAGAGTAAGATAGCGGAGCGAAAAGAGCCAAACCTAGCCGGGGGGCGCACGGT  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5303  ACTAAGCTAGAGTAAGATAGCGGAGCGAAAAGAGCCAAACCTAGCCGGGGGGCGCACGGT  5362

Query  901   CACCCAAAGGAGGTCGACTCGCCGGCGCTTCCTATCGCGCCGAGCTCCCTCCATTCCTCT  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5363  CACCCAAAGGAGGTCGACTCGCCGGCGCTTCCTATCGCGCCGAGCTCCCTCCATTCCTCT  5422

Query  961   CCCTCCGCCGAGGCGCGAGGTTGCGGCGCGCAGCGCAGCGC  1001
             |||||||||||||||||||||||||||||||||||||||||
Sbjct  5423  CCCTCCGCCGAGGCGCGAGGTTGCGGCGCGCAGCGCAGCGC  5463
```

7. Target Sequence:

ATCACATCTGGTGGAGTAATTTTCC**[A/G]**GTCACCTCTAATGTCAGTTCAGCCC

rs1137101 *LEPR*

BLAST® search results:

Homo sapiens leptin receptor (LEPR), RefSeqGene (LRG_283) on chromosome 1
Sequence ID: NG_015831.2Length: 227995Number of Matches: 1
Related Information
Range 1: 176766 to 177766GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1845 bits(999)	0.0	1000/1001(99%)	0/1001(0%)

```
Query  1       CTTTGGTATGTCTGaaaaaaaaaGCCTTTATTTCATCATTATTTTGAAAGCTGTTTTCGC  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  176766  CTTTGGTATGTCTGAAAAAAAAAGCCTTTATTTCATCATTATTTTGAAAGCTGTTTTCGC  176825

Query  61      TGGGTATAGGattttagaattgcagtttttcttttatttttagtacttcacttttacgtca  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  176826  TGGGTATAGGATTTTAGAATTGCAGTTTTTCTTTTATTTTAGTACTTCACTTTTACGTCA  176885

Query  121     ttatctttttgcttaTGTTATTCCTGATGATTAACCTGCTGTAATCTTTATCTTTGTTTT  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  176886  TTATCTTTTTGCTTATGTTATTCCTGATGATTAACCTGCTGTAATCTTTATCTTTGTTTT  176945

Query  181     TCTAATGTAGGGttttttttttttCAGATACCCTTTAAGCTGGGTGTCCCAAATAGTTTAC  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  176946  TCTAATGTAGGGTTTTTTTTTTTTCAGATACCCTTTAAGCTGGGTGTCCCAAATAGTTTAC  177005

Query  241     TTCAATTAGTATTTAGTATCCTGCTTTAAAAGCCTATCCAGTATTTTCATATCTGTTTTA  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177006  TTCAATTAGTATTTAGTATCCTGCTTTAAAAGCCTATCCAGTATTTTCATATCTGTTTTA  177065

Query  301     ATATTTAGCTCTTATTTTTCAATATAGGCCTGAAGTGTTAGAAGATTCACCTCTGGTTCC  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177066  ATATTTAGCTCTTATTTTTCAATATAGGCCTGAAGTGTTAGAAGATTCACCTCTGGTTCC  177125

Query  361     CCAAAAAGGCAGTTTTCAGATGGTTCACTGCAATTGCAGTGTTCATGAATGTTGTGAATG  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177126  CCAAAAAGGCAGTTTTCAGATGGTTCACTGCAATTGCAGTGTTCATGAATGTTGTGAATG  177185

Query  421     TCTTGTGCCTGTGCCAACAGCCAAACTCAACGACACTCTCCTTATGTGTTTGAAAATCAC  480
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177186   TCTTGTGCCTGTGCCAACAGCCAAACTCAACGACACTCTCCTTATGTGTTTGAAAATCAC  177245

Query  481      ATCTGGTGGAGTAATTTTCCRGTCACCTCTAATGTCAGTTCAGCCCATAAATATGGGTAA  540
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  177246   ATCTGGTGGAGTAATTTTCCAGTCACCTCTAATGTCAGTTCAGCCCATAAATATGGGTAA  177305

Query  541      GTTATGCACTAAAATGATGATAATAGGTCTAAACATCAGTCATATATAAAGGTTAAAAAT  600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177306   GTTATGCACTAAAATGATGATAATAGGTCTAAACATCAGTCATATATAAAGGTTAAAAAT  177365

Query  601      TGCTTACAAAAATATTTGCTAGCTTATCTCACTTTGCTTAACACTGTAATGATGGTAGAT  660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177366   TGCTTACAAAAATATTTGCTAGCTTATCTCACTTTGCTTAACACTGTAATGATGGTAGAT  177425

Query  661      GTAGTACTGGGGGTATTAAGAGTGGCTTCTAGAATGATTTAACAATGGTATGTATATCTC  720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177426   GTAGTACTGGGGGTATTAAGAGTGGCTTCTAGAATGATTTAACAATGGTATGTATATCTC  177485

Query  721      TGCCATTGTCACTTAAATTCTGTTTTGAAAACTGTTTTCTTTCAATCCTGGATCTATGTA  780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177486   TGCCATTGTCACTTAAATTCTGTTTTGAAAACTGTTTTCTTTCAATCCTGGATCTATGTA  177545

Query  781      ATGGATGTATATTGATTGGATATCACTTTTTCACATCTCAGATAACTATTTTTGAAAATA  840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177546   ATGGATGTATATTGATTGGATATCACTTTTTCACATCTCAGATAACTATTTTTGAAAATA  177605

Query  841      GTAGCATGTTTCTTGCCTGAATTTATTCCTTCAATAAATATTTCTTAGAGGCTCATGTTT  900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177606   GTAGCATGTTTCTTGCCTGAATTTATTCCTTCAATAAATATTTCTTAGAGGCTCATGTTT  177665

Query  901      GTCAGAGACTGCTCCAGGAGCTGGAAAAAGAGTGGGACATTAGACATAGTTCCCACCTCA  960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  177666   GTCAGAGACTGCTCCAGGAGCTGGAAAAAGAGTGGGACATTAGACATAGTTCCCACCTCA  177725

Query  961      GAGAGCAGGGACAAACAATAGTAGGCAGAGACAATGATAAA  1001
                |||||||||||||||||||||||||||||||||||||||||
Sbjct  177726   GAGAGCAGGGACAAACAATAGTAGGCAGAGACAATGATAAA  177766
```

8. Target Sequence:

AATTGGAAATCTGTAACGCAGCTTC**[C/T]**GTAAGCATGTGTGGGCAAAAAAGCA

rs3219151 *GABRA6*

BLAST® search results:

Homo sapiens gamma-aminobutyric acid type A receptor alpha6 subunit (GABRA6), mRNA
Sequence ID: NM_000811.2Length: 2519Number of Matches: 1
Related Information
Gene-associated gene details
UniGene-clustered expressed sequence tags
GEO Profiles-microarray expression data
NewGenome Data Viewer-aligned genomic context
Range 1: 1424 to 2335GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps
1681 bits(910)	0.0	911/912(99%)	0/912(0%)

```
Query  90    GCATCCTGACTCCAAATATCATCTGAAGAAAAGGATCACTTCTCTGTCTTTGCCAATAGT  149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1424  GCATCCTGACTCCAAATATCATCTGAAGAAAAGGATCACTTCTCTGTCTTTGCCAATAGT  1483

Query  150   TTCATCTTCCGAGGCCAATAAAGTGCTCACGAGAGCGCCCATCTTACAATCAACACCTGT  209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1484  TTCATCTTCCGAGGCCAATAAAGTGCTCACGAGAGCGCCCATCTTACAATCAACACCTGT  1543

Query  210   CACACCCCCACCACTCTCGCCAGCCTTTGGAGGCACCAGTAAAATAGACCAGTATTCTCG  269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1544  CACACCCCCACCACTCTCGCCAGCCTTTGGAGGCACCAGTAAAATAGACCAGTATTCTCG  1603

Query  270   AATTCTCTTCCCAGTTGCATTTGCAGGATTCAACCTTGTGTACTGGGTAGTTTATCTTTC  329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1604  AATTCTCTTCCCAGTTGCATTTGCAGGATTCAACCTTGTGTACTGGGTAGTTTATCTTTC  1663

Query  330   CAAAGATACAATGGAAGTCAGTAGCAGTGTTGAATAGCTTGCGGCCAGGACAACCTGAAT  389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1664  CAAAGATACAATGGAAGTCAGTAGCAGTGTTGAATAGCTTGCGGCCAGGACAACCTGAAT  1723

Query  390   TCTATAAGTTCTTGTTTTCTGTTTCCTATGTTTTCTTAAAAAATAGCATTGAGACTTGTG  449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1724  TCTATAAGTTCTTGTTTTCTGTTTCCTATGTTTTCTTAAAAAATAGCATTGAGACTTGTG  1783

Query  450   TAGATGCTTCTCAGAACATGAAATCAAATTGGAAATCTGTAACGCAGCTTCYGTAAGCAT  509
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct  1784  TAGATGCTTCTCAGAACATGAAATCAAATTGGAAATCTGTAACGCAGCTTCCGTAAGCAT  1843
```

```
Query  510   GTGTGGGCAAAAAAGCAATAATCCTACTCCTCAAAATAGAAAGTTGAAGATTGCTGAAAA  569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1844  GTGTGGGCAAAAAAGCAATAATCCTACTCCTCAAAATAGAAAGTTGAAGATTGCTGAAAA  1903

Query  570   ATATGACTTTTCTGTATGTTAGAGAAAAACTTTATGAGGATGAAATGGGTTCAAGATGAA  629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1904  ATATGACTTTTCTGTATGTTAGAGAAAAACTTTATGAGGATGAAATGGGTTCAAGATGAA  1963

Query  630   TTTGTCAACTTTTGTCTTCCATTGTTCAGTATTTTTAATTGTCACTGTAAATAACATTTA  689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1964  TTTGTCAACTTTTGTCTTCCATTGTTCAGTATTTTTAATTGTCACTGTAAATAACATTTA  2023

Query  690   CCACAAGGCAGATAAAATAAGAAATGCTGACACTTCCAAAGGTTGCCTTAAAATATGTTT  749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2024  CCACAAGGCAGATAAAATAAGAAATGCTGACACTTCCAAAGGTTGCCTTAAAATATGTTT  2083

Query  750   ATTTTGGCTTAGTTCCCGAGAGGGCAAAATATAAATACAGTCTAAATATTTATCAGTAGG  809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2084  ATTTTGGCTTAGTTCCCGAGAGGGCAAAATATAAATACAGTCTAAATATTTATCAGTAGG  2143

Query  810   TTAATACCAGCATGTTGGAGGCCTTTATGCTAGTAAAATGGCTTTCAGTGGCATTGTAAA  869
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2144  TTAATACCAGCATGTTGGAGGCCTTTATGCTAGTAAAATGGCTTTCAGTGGCATTGTAAA  2203

Query  870   GCCTACATTGAGCTTAGCCATTTGTTTTTAACCTCGCTGTGCTCTTTTACCTCAATAAAA  929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2204  GCCTACATTGAGCTTAGCCATTTGTTTTTAACCTCGCTGTGCTCTTTTACCTCAATAAAA  2263

Query  930   TGTGGTGTTTGTATACATATAAATTATACATAGCTCATAAATTATGTATGCATATGTACA  989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2264  TGTGGTGTTTGTATACATATAAATTATACATAGCTCATAAATTATGTATGCATATGTACA  2323

Query  990   TAGCTGTAGTTG  1001
             ||||||||||||
Sbjct  2324  TAGCTGTAGTTG  2335
```

Appendix B2: Alignment of Sequencing Primers to FASTA sequence as a Quality Control Check

1. Primer Design:

Forward	ACCTGGAGATCATCCATCTG
Reverse	AATTTCCATCTCGGCTCCTG

rs1800497 *DRD2*

(Kindly note the below FASTA sequence is in the reverse compliment orientation, due the manner in which the discovery sequence has been published in the NCBI database)

```
TCCAGGCGAG AGGCCCCAAG TAGTCTAAAT TTCTTTCTTT CTTTCTTTTT TATATGGAGT
 CTCGCTCTGT TGCCCAGGCT GGAGTGCAGT GGTGCGATCT CGGCTCACTG CAACCTCTGC
 CTCCTGGGTT CAAGGAATTC TCCTGCCTCA GCCTCCCTGG TAGTTGGGAT TACAGGCACG
 TGCCACCATA CCCAGCTAAA TTTTGTATTT TTAGCAGAGA CAGGGTTTTG CCATGTTGGC
 CAGGCTGGCC TCAAACTCTT GATATCAGGT GATCTGCCTG CCTCAGCCTC CCAAAGTGCT
 GGGATTACAG ACGTGAGCCA CCACGGCTGG CCAAGTTGTC TAAATTTCCA TCTCGGCTCC
 TGGCTTAGAA CCACCCAGAG TGGCCACTGA CGGCTCCTTG CCCTCTAGGA AGGACATGAT
 GCCCTGCTTT CGGCTGCGGA GGGCCAGTTG CAGGGGTGTG CAGCTCACTC CATCCTGGAC
 GTCCAGCTGG GCGCCTGCCT
 [C/T]
 GACCAGCACT TTGAGGATGG CTGTGTTGCC CTTGAGGGCG GCCAGGTGGG CGGGTGTCCA
 GCCCACCTTG TTGCGGGCGT GGACATTTGC GTGATGTTCT AGGAGGTTGA TGACACTCAG
 GAAGGTGCTC CTCTGGACCG CCAGGTGGAG GGGTGTCCAG CCTGACTGCT CTGCAGCATT
 GGGGTCAGCC CCACACTGCA GCAGTGCTGA CACCACCGCC TCCTCCCCGT GGCGTGCAGC
 TAGGTGCAGG GGAGTCCAGT TCACAGCTCC AAGAGCACCC ATGTTTGCGT GGCTCTCTGC
 CAGCAGATGG ATGATCTCCA GGTGGCCCTT GTAGGCTGCT AGATGCAGGG GTGTCCAGCC
 CTGGTGGGTG GGCAGCTCAA GGCTGGCTCC GTACCTGAGC AGCATCTTGC AGATCAGGTA
 TTTGCCCCTG GCAGCTGCAG TGTGCAGTGG GCCGTAGCCG CTCTGGTCAA GGGCATCAGG
 GACCGCTCCA CTCTTCAGCA
```

2. Primer Design:

SLC6A4 (rs25531)	Forward - V1	GTTGCAGGGGAGATCCTGGGAGAGG
	Reverse - V1	CCTCCTAGGATCGCTCCTGCATCC
SLC6A4 (rs25531)	Forward - V2	GGCGTTGCCGCTCTGAATGC
	Reverse - V2	GAGGGACTGAGCTGGACAACCAC
SLC6A4 (rs25531)	Forward - V3	GGTAGGGTGCAAGGAGAATGCTGGAG
	Reverse - V3	CTGCAACCTCCCAGCAACTCCCTGTAC
SLC6A4 (rs25531)	Forward - V4	CTGAAGAGGAATCGGCTCTGGGC
	Reverse - V4	CGATGTTCACTCCAAATGATGTGC

rs25531 *SLC6A4*

FASTA Sequence

```
TCTCCCGCCT GGCGTTGCCG CTCTGAATGC CAGCACCTAA CCCCTAATGT CCCTACTGCA
 GCCCTCCCAG CATCCCCCCT GCAACCTCCC AGCAACTCCC TGTACCCCTC CTAGGATCGC
 TCCTGCATCC CCCATTATCC CCCCCTTCAC CCCTCGCGGC ATCCCCCCTG CACCCCC
 [A/G]
 GCATCCCCCC TGCAGCCCCC CCAGCATCTC CCCTGCACCC CCAGCATCCC CCCTGCAGCC
 CTTCCAGCAT CCCCCTGCAC CTCTCCCAGG ATCTCCCCTG CAACCCCCAT TATCCCCCCT
 GCACCCCTCG CAGTATCCCC CCTGCACCCC CCAGCATCCC CCCATGCACC CCCGGCATCC
 CCCCTGCACC CCTCCAGCAT TCTCCTTGCA CCCTACCAGT ATTCCCCCGC ATCCCGGCCT
 CCAAGCCTCC CGCCCACCTT GCGGTCCCCG CCCTGGCGTC TAGGTGGCAC CAGAATCCCG
 CGCGGACTCC ACCCGCTGGG AGCTGCCCTC GCTTGCCCGT GGTTGTCCAG CTCAGTCCCT
 CTAGACGCTC AGCCCAACCG GCCGCACAGT TTTCAGGGGT CAGTTCCTCC AAGTACAAGG
 GGCGGTGGCT TCTCTGGAGC TGCAAACTTG TCACTGCTAT TTCCTTTCGG TCTTCTACTT
 CCTATCGTTC CTGGCCTCCT CTTGGGGAGA GGTAGAGCCC TCTCCTTTCC GCCTCAGGGA
 CAACCCAAAG CAAGTACTGC ATGTGCCCTT TTTAAAGTTT TAAATAATTT TAGCAAAAAG
 GATATTAACA TTAAATCAAT TTTTAAACTT TTTGAAAAAA TTATCAAAAC TACATGCACA
 TGGTTCAAAA CAATAGGCTC CTGCTGGGCC CTTTCAGATA ATTCAAATTG
```

Extend BLAST sequence for SLC6A4 gene

```
   1 gcctgtaatc ccagcacttt gggaggccga ggtgggtgga tcacaaggtc aggagatcga
  61 gaccatcctg gctaacaagg tgaaacccat ctctactaaa aatacaaaaa ttagccgggc
 121 atggtggtgg gtgcctgtag tcccagctac tcgggaggct gaggcaggag aatggcatga
 181 accgaggagg agcttgcagc aagccaagat tgcgccactg cactccagcc tgggcaacag
 241 agcaagactc tgtctcaaaa taaaataaaa taaaataata ataataataa taatatcaac
 301 ccctgaccta aactacttgt gttatctgta aattccaggc attgtatgag gaagcattgc
 361 aaaagtttct gttctgttag ctgatgcatg tagcccccag tcatgttccc cacgcttgct
 421 cgatctatca caacccttc acgtgtaccc cttagagtgg taagccctta aaagggccag
 481 gaatttcttt ttcagggagc tcagttcttg agacatgagt ccgctgatgc tcccagccga
 541 ataaagccac ttccttcttc aacctggcat ctgaagggtt ttgtctgagg ctcgtcctgc
 601 tacaccatgg ctcacagaga ttgcagatag atcgtggggt agtcacagta tctgtaaagt
 661 ctggggtgtc acctacaccc tgatttctcc tcacgtgtga gtatccttca agctgggcct
 721 ttccacatca gccactcctg ccttttcctc tcttttagaa aaaattatct gtgaagtctg
 781 aatatacaat gagttaagca gggttagcaa cagaaactta aatcgatgct aatctgacaa
 841 atacttactg attttaaaca attattactg tgcctctcaa tggggttgtg gaagcagcag
 901 gggcccatgg tgaagactga gggaattaat tatggggctt agatagaaag tgctataaaa
 961 gtcttagtta ttgttattgg cattagcatt tgcccagcaa gtgcttctgg gggtgtttgg
1021 ggtgttatca tagagtgagt ggtgtggtga ggctgctgcc gtaatacctg actgagagat
1081 aatgggagtt caagtcaaga ggatggtggg gctggggaga gggaggtggg agaccgggtg
1141 atggggtgct aggaggaggg tttagcaaag gcagaaggct ctggaggaat gaggggaagg
1201 agctgggaaa aatgaccaat gacaaaatga acctacttcc aaatgtttct ggccctttctt
1261 agggctgaac ctgatctgtt cctcccccaa acccccccgca ctcccaaaag gaccttgtct
1321 ggagttccct ttagcttgct tcatctggat ggggtgggcc tgctaggggc tcccatggag
1381 acctccgtag agatgagccc agggttcaca gtggcagcaa gtcctgcctt ggaggtgcca
1441 ctgaacatga atgtagagtc accagaacca tgccaggaaa aactgggggt cccaggaaaa
1501 accaggggtg tcggggaggt gctggtggga atggagcaag atgacactct aaggactgaa
1561 tattgtcagc attttcctgg gcctgagaaa aagcccacaa ggaggggtgg gccttcttct
1621 ttcgttattg tataaacagg gcagagacag gacagacaga tggcaccaag ggcttccgga
1681 gggagagcct gggattcgca tggttcggtc ctggctggct ccacagggga ctcactgctg
1741 gcctttgaga cttccaggaa gtgggttttc ttggtggtgt ttggggcagg tgtaaatatt
1801 tatgtgcccc tttcctctta aagtgaaagc ttgtatttat tctgagcccc agagcagctt
1861 ctaagcccca ctatttctgg gctcaagcag gtgaacaaag aaaaactatt gctgtgcggt
1921 gattgccagg cagcctcttc ccagctgccc aggacctgtc gcccttgaca tctccctggc
1981 tgtccctata caccaccagc accaagcttc gcaggtgcct gctgggtttc tggcttgcat
2041 ctcttccctg cggccaccag ctgacgcctg cattcacatt ctaaggaacc cgtctgtccc
2101 aaccaccacg acaaccccct tccccctac tcttacctgc agcccacgca ggtcacagcc
2161 attagatcct tcccctgaag cgctgcttcc ctcctgccgc actcagggct ccctcctgcc
```

2221 ttccatgccc aggacagggc ctcccggcct ctccaggtct ctggagcaga gaggttgtgc
2281 cttcctcatg cggtcaggct caggctggga aggagcggcc cagaggcggg gggaggcacg
2341 ggtgcggggc tgcctcttcc ccggtagtgg acgctgaagc ctgtccacct gaattggagg
2401 cggggcgggg cagtcgactg accgccaggc aggggagcag ggcgcgggct cagagggagg
2461 gagtgggagg ccgaggaagg acaaggtgct tgtggcgatg agtatcagaa ggcagagcgg
2521 agctgctcaa gagaggctgg ctctgtggag gtgaggacct gcaggtgggc ctggcctctc
2581 tgtgggatgg tacccagggc aggaggaaga gaaacacctg tttagagaga gggggagggt
2641 ctgggccagt tgtgtgcgga ctctgccatg gaggcagggg gaaccagggc ggggccgtct
2701 gtgtaattct aatgtcattt gctaatggtg gttgcccgca gctggggagc aagaggggag
2761 gagatgggac cacccggcca gagagaggag gggtcgaagg ccccccttgga ggaagggccc
2821 agcgcacagg gcctgcaagg agttggcccc ttccccaccc acagctgggc agtgcccact
2881 gagatcatcc tccctacttc acaaacagag acacaaaggc cagacagctt gaggttaagg
2941 agctcggtca agcccacacc acagtacagg gtgggggccg ggatttgacc caggtccaca
3001 gggtgcccac tctctaggca gcacctgcac agaacagggc agtgccatag ggaggggctg
3061 cgcgctgggc aggaatgcat cttcggctag aaagagtcta gaaggaagat cagaagtggc
3121 gagactgcac aattaagaag tgctaagaag tcacactgga cagacctggt cttggtagcc
3181 aagcgctgaa gtcacccctg cagccgtcct ccgctttggc gcctcttccc agcgtccctg
3241 ccccccctcct ttggccctcc tggaaaggac actttgcgtt ttctgttgcc cttgcctata
3301 cagcacaaac atgctcattt aagaagtgga acgtgggagg cagcagacaa ctgtgttcat
3361 ctgaaaggag gaggccccac tcccgtgcag gcgctgcccc tgggggtgaa attcccaagc
3421 ttgttgggga ttctcccgcc tggcgttgcc gctctgaatg ccagcaccta acccctaatg
3481 tccctactgc agccctccca gcatcccccc tgcaacctcc cagcaactcc ctgtacccct
3541 cctaggatcg ctcctgcatc ccccattatc ccccccttca cccctcgcgg catccccccct
3601 gcacccccag catcccccct gcagcccccc cagcatctcc cctgcacccc cagcatcccc
3661 cctgcagccc ttccagcatc cccctgcacc tctcccagga tctcccctgc aacccccatt
3721 atccccccctg caccctcgc agtatccccc ctgcaccccc cagcatcccc ccatgcaccc
3781 ccggcatccc ccctgcaccc ctccagcatt ctccttgcac cctaccagta ttcccccgca
3841 tcccggcctc caagcctccc gcccaccttg cggtccccgc cctggcgtct aggtggcacc
3901 agaatcccgc gcggactcca cccgctggga gctgccctcg cttgcccgtg gttgtccagc
3961 tcagtccctc tagacgctca gcccaaccgg ccgcacagtt ttcaggggtc agttcctcca
4021 agtacaaggg gcggtggctt ctctggagct gcaaacttgt cactgctatt tcctttcggt
4081 cttctacttc ctatcgttcc tggcctcctc ttggggagag gtagagccct ctcctttccg
4141 cctcagggac aacccaaagc aagtactgca tgtgcccttt ttaaagtttt aaataatttt
4201 agcaaaaagg atattaacat taaatcaatt tttaaacttt ttgaaaaaat tatcaaaact
4261 acatgcacat ggttcaaaac aataggctcc tgctgggccc tttcagataa ttcaaattgt
4321 caccaggttg gagtgcagtg gttcgatcac ggctcactgc agcctcgacc tcccgggctc
4381 agctgatcct ccacctcagc ctcctgagta gctgggaaca caagcgcgag caaccacgcc
4441 cggctaatta aaaaaatttt ttttctagag atggggtctt gctgtgttgc ccaggctggt
4501 cttgaattcc tgggctcaag caatcctccc gcctcagcct cccaaagcac tgtgctcctt
4561 tttgacgcag ctttgaactg tagctggtta acaaaatgag aaccagttct tcattccttc
4621 attgtggaag tctttattgt gagactctgg ggacggagag gaattagaca agggcctcta
4681 agctgagctc acatcccagc cggtcagtca gataaacgca tgggtatcga gtactgctag
4741 gtcccaggaa gaaagagaga gcagctttcg ggatggggac gatggggagg tgtccgaggt
4801 caagagaaag cggcacgagc agacccctgt gtgccgtcct gtgggcgcgg ggcggcaggg
4861 gaggcgcaca cctgctcctt tgtgcagcct ccccccctccc gcaaagttaa agagcaggaa
4921 agtcaggatt cctcgctcgg ccctgccctg ccggctgctc cgcgctccgc tcctccctgc
4981 gagcgtgtgt gtgtgtcggg ggtccctccc ctcctggctc tggggtcggg cgcgcacccc
5041 gccccgtagc gcggcccctc cctggcgagc gcaaccccat ccagcgggag cgcggagccg
5101 cggccgcggg gaagcattaa gtttattcgc ctcaaagtga cgcaaaaatt cttcaagagc
5161 tctttggcgg cggctatcta gagatcagac catgtgaggg cccgcgggta caaatacggc
5221 cgcgccggcg cccctccgca cagccagcgc cgccgggtgc ctcgagggcg cgaggccagc
5281 ccgcctgccc agcccgggac cagcctcccc gcgcagcctg gcaggtgggt ccgcttttcc
5341 tctccgcctc gaacccacgt ttctttccag accttcttcc ccgcctcggg gaggggggata
5401 gaaccgctgc gccccaccgc cctgcgagga ggcgaggagg tgcatgcgcc ccagcggtgg
5461 gcgccggatc ctgcccctgc gccctccacg ctcagcaaga gccagagctg aagctgaccg
5521 gccagagtgg gagacgagga acgtggagtg ctcgaagtgg gcgggcgtag ggggctcctt
5581 tgtctattgt tgcaggggct ttgcaacctc ttgaggaccc gtgaaagacc ccttaaagag
5641 ggtcttgagc aaagtgctgt cctgcgcttt gggaagagtt gcttgctttc gtttcaaccc
5701 atggtttatg tttgatcttt actttgctgt catcgcgtgc aggtggcttt atgcaaaggg
5761 agagttctgg ttgacacaaa tgcccagaca gctagagaaa tctctgagtg gcgcctaagt
5821 gggccgagtc tccgtgcgct tgggtggttt cactgtcaga gctgaggcgc tctcagacac

5881 ctcacaggcc attcagaatg tatgtgaaac tccaactcca acgactgtgc tacgtggttg
5941 aaggatataa ccatcatttg cttgaaaatt gtgtacattc aaacttgttc ttgaaaaaaa
6001 gctctctgga gctcaagagt ggatttgggt cttgggctaa ggagcagatt ttcttctgaa
6061 ctggggcttt tgcatattga attgttcctt ttccatttgg gactctgaag aatttgagag
6121 gggtagaatt tctgaatcag acattacaaa gactttgtgt atgtattttt accatcagtt
6181 ttgtccagaa aagtgaacct ggtcaatgga ttatttatga gcctgctagc taactctcat
6241 cataggaaga cgagagcaaa tatttaatgt gttttgcctt cagttacata ggaaaaatgt
6301 ctaagaaggt gacattggga cttgtgttta atgaaacaga gaactgtggg cagcgtcagt
6361 gttgggtttt taggagcgta gggagcacac agctttgact ctttgtccca ttacttgctt
6421 ctgtgtgaag ccactgaggc cccagggttc gggttttcct gccacgcact ctggggtggc
6481 agtgaccgtt ccaatatgga tgagtgagaa gcaggttctt atgaggtttg tgcaaattga
6541 gcaagccact tggggcatgg gtttttcctc ttttctttct ttctttcttt ttcccactaa
6601 agaacagatt gaaagtgctt cacaattaat caaaagtccc ctcaacactc tggtgatcca
6661 tctaagacct ctcagagata taaccaccag cacagatatt caaacccatt tttttcaatc
6721 tccttgttta agaaatatat atatactagg agtgggatgg gggaataatg ttgatgtttt
6781 aatggaatct gaaagacctc agaaattcct ctcgggggat gtttagagtc gtctgagcat
6841 tttgatgaaa cttggagagg catgttgaag agtttcgcaa ataggtgagg tgcctgagga
6901 atggggatgc ttggctggag gggggttagg agcatggatc agtctgtttc ttacagcagg
6961 tgcttgcttc ccaaacaccc tggatgtccg ctgccaaaca gaaacacagg taactccaga
7021 cattaagcca atacgtggca gctccttctc cttcctagtt gtggaatgaa tggtataatt
7081 ggagagaata tagaaggcag ggcttcctga agaagcttaa ttttaatttg gcttcatgaa
7141 tcaacaggat atggtcatga cttgactttt tctaaagatg tctggacagc cagagcaaaa
7201 attatacagt cacttttggt tgttgttgag aaggtaactc agaccttagt tattcccagt
7261 aacatatcag ctcctcgtgg tcctcagctc taaatgaggg gctgtggaga ggactgccta
7321 acacgatgtc cgagtcccct ccaactcact ctgcaaccat gaccctaacc tctccacttt
7381 aagaaaacat gagttgaaat gatggtgcaa aaatgccgga gggccggtac tctctaccca
7441 gatatcacct tcaaatcagg gtctcctggg gatcccctca ggtttcactc taaggcagag
7501 gctttgcaca ctctgacaca tgtaagtaag ttaacccctg caccttaacc acagccaacc
7561 aacaggggcc caataacccg actccccagc tgcccaaatg ccctttgcaa gatgcccagg
7621 gtgtggtgcg gtagaaatag cccccaggca gcacaaattt gggctgcacc gattttcctg
7681 tccagtccct gggctattaa cagccacctc cactccctga tgtttgcggt gttccatgga
7741 acatacgtgg gtgtggagcg aaacatctca taaaagcgcc tgtttctcca cacttgggac
7801 aagtgaggca gactcccttg ctcttacagg gaatctgttt ctgggcttaa ccaggcaaaa
7861 tgacattcca agggaagaaa actcccatct ttctaggagg gacccagggg aaaacagatg
7921 aaactagcct ttcacagaaa cctggcatca atgatcttcc ctgagcagaa aggcaaattg
7981 ctcttccagg tgggcacggg ctccaccatt ggcctgcaca gccccgtact tccaggtgga
8041 gctcacggtg ctgactttgg ttcacattca gaaatgttgc cctgtgagaa acgtctctgc
8101 aggggctgct cccttcagcc tgacttttgt ttgctctggg tgactgggaa gacgaatgct
8161 catttactga atgtcaacaa cgcgctttta aatgggagca agtgttttcc cagccctgtg
8221 tctaatgagg tgacaaaatg agtagcgata aagtggatgt gaggcagtct gactgcctta
8281 ttattatttc catctttcct ttggcccctc tggaacctgg cagcagctcc ccagcttcca
8341 tcccaaccct gaaccaaggg gatttgttaa tgggatttaa gccacaggct ctttagcacc
8401 caggactgga tggacctcca ttaatcccag aagcccctgg gctgcaaaag gagagacgca
8461 acccagtgat gtggagacag agaggcaatg tcccaggtct tcaggtactg acgagtgaca
8521 gacagcaagg ggagggtctc ctgggatgct gatgtccagt tcttcaggtc attctttgag
8581 gtagggaggg aagaggtgac aggaggggtt atgatgagac caggtggtga tctgggcgag
8641 cacttctcag caaagctggg gacacagcga acccattgtg tgctttggga gagcccagag
8701 ccccagactg ggtgggcagc tagcaggagg gggcaagtac gtgggccttt gccgtcttgg
8761 gttttggtgt gacctttctg tgtgtataca ggaaaaaaga gcatttcatt tcctttgtta
8821 ctgaactgtt gaataaaata agctgatttg tcttctctaa gggaactttt tcctctaagg
8881 ggccaagaat tagtgagtgg gagacctgga aggaccaggg aggggttctg ctccccgtag
8941 ctctgtgcat ggctggctct agctcaccca cggctcttga ggacaaagcc aggcctagaa
9001 cttgagtctc ttctactttc gctctcttct cttccctccg taatgatggg catgtcttct
9061 ttaaaaggcg aaccaaatta gctggccctt tctcttgttt cagagttggg gagctgggcc
9121 gaaggaacag ggagggaaga gcctggaggg gagaggggag gtaagctgtt ctggtcagca
9181 gggagctttg gctgactgga agtttcccag cactggctga gggctagaat tatttttct
9241 tttgaaaatt ttcaactgaa aaatcagaga tatggtccat tttcaccagt taacattttg
9301 ccaatttgtg tacaaagtca ctctaattat gcatatatct aatttttgca tatgcagtat
9361 atatatgatt tgatatatta tgtgtaatat ataagatcat taatataatt atgccattta
9421 tgtatctgta tgtatacata cagatatata tacacactat atatacacat atgtacatac
9481 agatacatac agatatatat atacatatgt acatacagat acatacagtg tatatatgtg

```
 9541 tctgtataca ctatatgtat atataaacta tatgtatctg tatgtatacc tacagatata
 9601 tatagtgtgt atatatatat ctgtaggtat acatacaggt atatatatac tatatatata
 9661 ggttcaacta atatatatag ttgaaccttt gagagtttgc agacttccag atacttcttc
 9721 cttaaatact tcaacacaga tttccaacaa acaaggacct tcttgtacgt aaccacagcg
 9781 ccattttcat gctggggaaa ctggagatag atacaatact attatcgcac atacagctca
 9841 tattcccatt tcccccgctg tctcaatgat gtccttttta gcatacagtg ttactctgtg
 9901 agactgtaaa ctaggggtca gaaagttttt cttgctctca tctcagtgca tatctgagaa
 9961 ctagactggt gcaatgatgt gtaaggatga actgattttt catgagtgtc tccaactgtg
10021 tgactcacac taaagcacat ctcagaaatt gttgacctaa gagtataggc acctttgagc
10081 agcaacagcc aaactccaaa caccttgtgg ctgaagaggt ggttgcacgt gtcctccgcg
10141 ccccctctcc acctggatgc acccgtctcc tccagcagcc tcacctcacc tgacctcacc
10201 tctccttctc ccctccatgc cccacccaga gaaatgggcc agtttccccct ctcttttctc
10261 ttctccactc ccattttatt ttgtagttcc acttcaaacc ccctttcccc cacttccctc
10321 ttccgccttg ctcagcagcc acaggtgaat ccacccttat ttcatagcac cctctccctg
10381 gggttgtttt ttcctctctt ctctctgttt tgggtccctt tgtggaccca atcatgtctg
10441 ttttacacaa aggtaatctt agtaacatac ctgtataaag gaaaagaaaa gctgtgtgtg
10501 gacatgttcc catgtgttct atctcgattt ttacttattg cataaattgt atttatcaca
10561 acagtaaagg agtgtttgga aggtgaagtg acactcctaa tcctgattaa tctaacacat
10621 catctgtttt gagtttctca agtctctttc ccagcctcgt ccacatttaa atatactatg
10681 gtaacagagt aaaagcagtt ttgttttctg caaaaagttg taagttgtga aattatgtgt
10741 cacaagcaat tcatgtcacc aacttcttct tcttcctctt cctcttcctc ctcctcctcc
10801 tcctcctcct cctcttcctc ttcctcttcc tcttcctcct cttcttcttc ttcttcttct
10861 tcttcttctt cttcttcttc ttcttcttct ttcttctttt tcttcttctt ctttttgag
10921 atggagtttc gctcctgttg tccaggctgg agtgcaatgg cccgatctca gctcaccaca
10981 acctctgcct cctgggttca agctattctc ctgcctcagc ctctaaaagg ctggtattac
11041 aggcatgcgc caccatgcct ggctaatttt tgtattttta gtagagacag ggtttctcca
11101 tgttgatcag gctggtctcg aactctgacc tcaggtgatc cgcccacctt ggcctcccaa
11161 agtgctggga ttacaggcat gtgacactgc gcccagccca tgtcatcaac ttctaaagaa
11221 cttggaccag gcgcagtggc tcatgctgta atctcagcac tttgggaggc tgaggcagga
11281 ggatcgctgg agtccaggag ttcaagacca gcctgggcaa catagcgaga tctcatctct
11341 actgtgtatt ttaaaattga acaacaaaca aaactaaaga cctcaaaagt gcctaaaaga
11401 cacctccgtg gaatctgtac aaggatctcc ctgctccttg ttcttctttc cttaaattta
11461 gttttcctta agaaccaaaa agaaggttaa aataaagaaa gaacttcgaa tcagtctcat
11521 tgtcagcccc ctccccaccc ctaagccgtc gtgtagcact taacaacgct gtgttctaca
11581 gctttcgttt cagatgtttt agtttatcac atttatttct gtgggttagg aaatatattc
11641 taagttccaa accaccagct cacagatgaa catttggagc ccagaccgct gatgggtgtg
11701 cctgccagta cttgtttcaa agtctgaaag cccatcccgg gctctttaat cctgggcatg
11761 tcagtctctt cctgtccata gtccatgaaa atgactgtta aggcagccct cctcctctgc
11821 ccacaagatg cacagccaat cttcccgatg ggtgaggcag ggagggacag ggccccttag
11881 aatagaaccc tgatcccatt tgacttaact ccagggagga gagaaagagc agcacctcac
11941 atggagctca gatctgttgg caccctcgcc cagtcaggaa acccaaaaca agccctccat
12001 gctgagagct gtacagggcg tgctccagtc tgcaagtatt tggactcagg gacgtgtagc
12061 catagcccct gcttccaggg gcacacgaag caaatgtata taaaaggata actaatgatg
12121 cctggtaatc aatggcgagt gacaaatgag aattgagtag aaagagtgta tcttgggctg
12181 gagtgctgag aggaggtgga tagaagctta tgccaggcct tggagggtgc aggataatta
12241 gaggaaggga agctacaggt ggggacgcac acggcactgt tggtgcctac tggcccctga
12301 ataacctgcg gtatttgtta aaaactgcag atccctgagt gcccacgtgg atctcttgaa
12361 ttagaatctc tgggagagca gcccaggagt ctgtttccca ggtgacccca agcactgatt
12421 gagagcagca agcatgaccc agtgaggccc gtgctgggag ccacagatca cttaagtccc
12481 aagtcttgtg ggcctcagtt tccctgctag ggtatgaagt agaaaggcac aatgggttgc
12541 ggatgtatgg tggtgatagt gacatcaact cactgtcatt tattgatgct tccctgtgct
12601 agacactgag ataagcatgt gtcatacttt aattcccacg acagccttta ggaagtagcg
12661 gttttaacag tctgctttta cacaggagga aatggaagct aagagaggtt aagtgaccta
12721 ctcaaagtca cacagccatt aagtggtgga ccagggattc cagtgcagac ctgaaacaga
12781 aaggagagct atgggttgtg aacagaggaa ctaatgaaag atgactctgg aaggaagtta
12841 gtctgggctt ggctgtttat attcactgga ggagggctgt tcacatctgc tatgctaaat
12901 tcatttcttt ctagtaaagt caaaataaat aatccagtat ccttttggcc agcctctcag
12961 ggggaccttg accccatcac aaggggcata tcctctagat tttatactct gtgtagagag
13021 aagcccaggg ctatcctctc ccaggcaggg aaagttctgg acagaaattc tttttctttc
13081 tcatgataac tggcccaaat ggccctgagg attttatcct atttaattcc tctccaaaaa
13141 attaaaccag gctgggtgca gtggatcatg cctgtaatcc cagcactttg ggaggccaag
```

13201 gcgggaggat tgctcatagg ccaggagttc aagaccaacc tgggcaacat ggtgaaaccc
13261 catctctgaa aaaaatacaa aaaattagct gggtgtggtg gcacacacct gtaatctcag
13321 ctactcggga gcctgaggtg ggaggattcc ttgagcccgg gaggttgagg ctgcagagag
13381 ccatgatcat gccactgcac tccagcctgg atgacagtgc aagactctgt ctcttaaaaa
13441 aacaacacac acacacacac aaacacacac acaaaggaac agaaaaaaca ctaaactgac
13501 tctgtcatat gaggaatcct ttggataagg agatggctgt gatagtggag ttcatttcta
13561 ccctttttgt ccaaaaaaaa aacaagcaga ttttctctat tctcttcatt tctatcatga
13621 aatctcttca attcattgga aaggcaaggc aggagtccct atggcttccg acctgtgccc
13681 ccatgaccct gcacacttcg ggaacacagc agtgccttca gcagatgcaa acgtgagccc
13741 ccattttaca ctgcagagcc tgcaaggaag gatcagccaa cagccaaggt cctgtgagag
13801 aaactgtgca atagcagcaa aggaagactt aaagcctggg ggcaactgtg gtgcaaaaca
13861 tgtgagctat ctaaacacgc cgcggcgagc cagcggcagc ctctgttcac cacgggatgt
13921 gtggcagcac cagcctctcc agaagcatcg tgtttttcct agcagttggt tttctttttc
13981 agattttagg cttggcctgt ggactccatg caggcaggac tgtgcctgcc ttgtttcttt
14041 tttctttctt tttctttttt ttttgagaca gagtctcact ctatcgccca ggctggagcg
14101 cagtttggct gactgtaacc tctgcctccc aggttcaggc gattctcatg catcagtctt
14161 ctgagtagct gggattacag acatgcacca ccacgcccgg ctcatttttt gtatttttat
14221 ttttatttat ttatttgaga cagagtttca ctcttgtcac tcaggctgaa gtgtagtggc
14281 tcaatctcag ctcactgcaa cctccgcttc ctgggttcag gcaattctcc tgcctcagcc
14341 tcacaagtag ctgagattac aggcacccgc caccacaccc agctactttt tgtatttta
14401 gtagagacag gatttcacca tgctggccac gctggtctca aactcctgac ctcaggtgat
14461 ccgcccgcct tggcttccca aagtgctggg attacaggcg tgagccagct ggcccggctt
14521 gcctgccttg tttctgagga gccgagcaca tgtgggccta cgcacatctt tgctgaggga
14581 gtgcttctcc atgaagaagg ccagattcca tgtttactgc ttcgggcggc accagcgttc
14641 cataattccc atcactcttg ggagatgccc tgggaacttc agcctccctc tgaatagagg
14701 atacgaactg ccagttggct ctatgacctg taactaatcc tatctcaacg ttaaacccac
14761 agttcatttg tttttgtttg ttttgagaaa gaaagttgag tttcctgaga aatgaactca
14821 ctttgatctt gacacatgaa gtcaaagttc aacatttctg tagtgtcttt ctgcttagaa
14881 ggcactgata ggcatcaagt agatggagcg tcacaggcat ttccaaaggg ctgggtgttt
14941 atgaccaatc gggataatta catgcttgtt caggtacggg ttgctcctgg ctatagagac
15001 aaagtgtgtg gaaagaggaa tggagattcg atcagaagtt taaatagaaa cagtcctttt
15061 ttacttggag tcagggattg ggggaatgga gtgcattagc ttcaggagag ttaagagaaa
15121 gtcaaccact cccattggct ggtcaactgc acccctcatg tgcagagacc taggggacg
15181 aagcttctgt ctgccagctg tcattccaca gcagaggtgg ggtccacggc accaccctca
15241 gcctgtcacc agctcagcct ggaacaacag cacacatcca agatgcaaag gaaggcttcc
15301 ttctggggct ttttctggac atggagaaac tcaaagactg cctggagaag gtcacagctg
15361 taatattgac ctcgtttgca atctttccac ccagcaggat aaaaatacct tttaaaaaaa
15421 gaaagcttag aagagggcat tctcagaaac caaccttctt aacaactcac acttctctat
15481 tcctgggcca gcttgttttc agccattgtt atgtgcaagt cttgatagaa aaagccagcc
15541 agggctctgc agaatcagaa aaaataggaa gagagtctgg ggctccacag gcaggagctt
15601 gcaggcagaa agggccccgg gcacatggca tcttctcctg gcagacccga ttagtcaccg
15661 agaggagttt tcctctacct ggcagaggca gaggagtaag ttaagaaaag gagatgtgtg
15721 tctcctctga gtggctgtaa ctttgtgttt gcaccccaga ttctatgatg agggtggggg
15781 ttcagtagga gaaccccaga aagcacagct ggatagagag agacagcctg tggggagagg
15841 ctgcccggtc atgagttgtc tgagggatgc tctgggtttt cttgcactgg ggccatttca
15901 gcctgttgat gtaaaatttt ttaattaaaa aattaagaag aactagggga tatttcagga
15961 aacaagacct ctgtaatggc tcagaacaca aatagaaaca gcgtattact cttccaggag
16021 ctgcttacca tagttgccag gaattcagga ctggatgcct tcggggctga acccctgagg
16081 actcctgaga acacacgttg ccctcaaaaa tctacctacc tgcacacatt ttctaccctc
16141 tctgcaggac gcctcaagga agaaagtaaa caaaatccag aacaaatcat ttggcccctg
16201 aggcaggagg gtaactgctc actgctgctg ctaaatgtct atttcctctc cccaaagtat
16261 tgcttctgag atggaccgca tttcccttcc gtagaccctc tgggcctggg tagagaggcc
16321 agggaacgct tgggtgatgc aattggagct tctatttaag tgagaggaaa gtggcagtac
16381 agaaaataga agaggttctt gcagaaaatt gtggggccag gccaggcatg gtgtctcaca
16441 cctgtgatct cagctctttg ggaggccaag gcgggtggat catgtgaggt caggagttca
16501 agaccatcct ggccaacata gtgaaacccc atctctacta aaaatataaa aattagccgg
16561 gcatggtgga gcacgcctgt agtcccagct actcggaggc tgaggcacaa gaatcgcttg
16621 aacccgggag gcagaggttg cagtgagcca aaattgtgcc actgcacacc agcctgggca
16681 acagagcaag actgtttcaa aaaaaaaaaa acaaaactgt ggggccagtg ggaggaagag
16741 caagggaagg agcagctctg tcaactcggt gccctccagc cgcacaggga aaagctcttc
16801 ctccctgcca tttggccttc aagttcgtca gttgtttaac gtacaaaaca tgtcattgat

16861 ataggccgga gggaagtaca gaggcctttt taaacaacac aaccactgac taccaagttc
16921 aggtttaatt tagctgaata tttaaatatg ttgacatctc atcatgtaat aaaataataa
16981 tatcccttgg cctggtaaat ggttacctat ggtgggtgag aggatgtgtg gttattttaa
17041 aagcctatca agcctagttg agggccgaag gaaggggggaa gtgggggggag gccaaggcca
17101 ggcagtagca taaatggtga gcagggtgag gttatggaga ccgctcttgt cagaattcat
17161 atccttctga gtggtgtttg cattcttgag cctggggtgg gtggcggagg gcccaacaag
17221 tgcggtacag cccaagacgt ctacctcccc ttacacacac acagccccta gctacatgaa
17281 gctgatttat ctcttgttgc tgtagatatt tttggcaaca atcgatccct ttaaaatttt
17341 tattttccac tttacagaaa aaatgcagca aacgccaaac cctagtgact gacattgcct
17401 ggtgttcatg tctaatcctt gctcatgttt atagccatag tcattcttgt gaagctacta
17461 ttttgtgtcc tgatttttca cttactatgt ttttgtcttc acatgtgtag gtagctaagt
17521 tgcagtcacc aaaaacccca aataggtagg agacatccat gcgccatttc tattagccag
17581 ttcctagcca tcctcctaac tgttggatct ggagggttag cctcagatca ttgcagttat
17641 aaacagtcca tccctggcta tcttttctag aagggtccag tgtcatctca gctaggcaga
17701 aacgtatttc aactcccaaa ctgtttcaag agtttgaaaa gctcaggccc agaagaaaca
17761 gagtgagcag ttgaacagtc ttacaactgc tgcttacaac tttataacag gccataggtt
17821 taaagtccaa aagggctcaa atttctaccc cacgtttgac gactcagacc tcaagccaga
17881 gctgagctga cttccctgcg cccaggccac ggcggtgaaa tgaaggcaca gcagccccgg
17941 ccaggccctc ccagtggagg cacaggggga cggcgacacc aaccacccca gcatcagtaa
18001 cctgcacact cttctcccta ggtctcctgg aggcaaggcg accttgcttg ccctctcttg
18061 cagaataaca aggggcttag ccacaggagt tgctggcaag tggaaagaag aacaaatggt
18121 gagcagcagt gccgtttaag gaccagtcgg gcttgcaaat gcataatggg aacgcagaaa
18181 ggccatgtgc agatagactg gacacagccg gaggcagctc aaatctgccc caaccttcct
18241 aggggctaaa agacgagtga cagaatctgt agctgatgag ctggggtgcc tcacgaaggg
18301 tggccttttt acctctgtcc tttgcttttt tccttttaaa atcatagctg tagttgatat
18361 aaaatataac tttgggtgtg cagcttaaat gatggtgagg actgtcctgc ctgctgaaag
18421 gcagaaacgc ctttgcttgt cattgtttcc ccaacatagc atccaagaca aaggggtgc
18481 tctgtgacgc tttggagcct ctctcacatt cttctgagct gctggcaaac cacaagatga
18541 gagtccagtg gctcaaaggg aagcgagtgc tttgctctcc agtccagcct gtgcaaacct
18601 tggtgatttg ctgtctattt gtgttggttt tattgtttcc ttttgaaatt acttgattgt
18661 catttgatgc cctgtatagg ttgagaaatt tagttgaatt tatggactgc catgtagcaa
18721 ataggatgtc tcacatttaa cagatgcatt tgtaatcccg ttaactaatg tagtttaaga
18781 tgtatgatga gttaatacca aagatgagta aaatgtctaa ctttttctc cttcctctgt
18841 gtgtctttcc aacagagtca atcccgacgt gtcaatcccg acgatagaga gctcggaggt
18901 gatccacaaa tccaagcacc cagagatcaa ttgggatcct tggcagatgg acatcagtgt
18961 catttactaa ccagcaggat ggagacgacg cccttgaatt ctcagaagca gctatcagcg
19021 tgtgaagatg gagaagattg tcaggaaaac ggagttctac agaaggttgt tcccacccca
19081 ggggacaaag tggagtccgg gcaaatatcc aatgggtact cagcagttcc aagtcctggt
19141 gcgggagatg acacacggca ctctatccca gcgaccacca ccaccctagt ggctgagctt
19201 catcaagggg aacgggagac ctggggcaag aaggtggatt tccttctctc agtgattggc
19261 tatgctgtgg acctgggcaa tgtctggcgc ttcccctaca tatgttacca gaatggaggg
19321 ggtcagtatc acaggctgcg agtagaggct gtgacccagg gtgggctgtg acccggagtg
19381 ggctgtgacc cggggtgggc tgtgacccgg gtgggctgcg acctggggtg ggctgtgacc
19441 tgggatgggc tgtgacccgg gtgggctgtg acctggggtg ggctgtgacc cgggtgggct
19501 gtgacctggg gtgggctgtg acccgggtgg gctgtgacct gggatgggct gtaggtcctc
19561 ttgagaggcc agaagacaga ttatgtcttt tcagtcttca ctggcgtaag actaggaaca
19621 tgatgactta gactttcgag ctggtagaaa agtcaaatct cgccaggcgt ggtggctcac
19681 acctgtaatc ccagcacttt gggaggctga ggcaggtgga tcacctgagg tcaggagttc
19741 aagaccagcc tgaccaacat ggcaaaaccc cgtctccact aaaaatataa aaattagctg
19801 ggtatggtgg tgggcgcctg tagtcccagc tacttgggag gctgaggcag gagaatcact
19861 tgaacctggg aggtggaggt tgcagtgagc cgagatggca ccactgcact ccagcctagg
19921 cgacagagcg agactctgtc tcaaaaaaaa aaaaaaaaaa aaaaagtcaa atctccagca
19981 agccaaaagc tgaagacagc gtgtctgagg ctctccgggt tgtgaagctg ggtgaagccc
20041 tcccaaacct gacctcacag atctccagca tcattagcgc tccatttatt gaagggcctc
20101 cttctcatcc tgcattctcc aggctctgtc ctcccccagc cagaagtgag ggaatggccg
20161 tcagccctgt gctgagtccc ttcactcctg ctggcaagtt ctgctgggaa gagtccacga
20221 gggtaggcct gcgtaaggga aaatttggag gtggcttcag agtcccagat gcttctgatc
20281 ctttggacct gtgtgaggtc agagaagggc atcagcgtta agtcctaaat cttaattacc
20341 agtctgggca gaatttaagc tggattagat ccacagcatg acattaatac caacctcaaa
20401 ccacgcatgg ctggccctcc tatctcctcc tgggatgtta gatttttgag gccagaggcc
20461 ctatcttgtt ccttgtaccc ctaaagcctc tcaagtagtt aacttgtaga ttataaaaaa

20521 tagtagctga ctgtaggcat gaagatggga acctgggggа gagcagagca gtggggaaaa
20581 gtgaatgggt ttggggcctc tggacttgtt atcagcacaa gaaaataaca agaagtgggt
20641 gtcagatgcg tcaagcttga ttttcatggg aactctgaca acagatttgg gagggcagct
20701 acaaagaggt ttggggatca tccatttgtg tcaccctcac ttctcccgtg acctaacaca
20761 gagtggcctg gatttactct cttttaaaga tttctgtcca tgtagttgca aaggtagttc
20821 caggggggccc ctgggaagct cagcgcttcc tgcctaggag gagcaaaagt cataatcacc
20881 ttatctcctg ctccagagga ctcaccttta gcctgggcat tatgtaaggg gctgggtcac
20941 tggtgacatt tccccgacac tcattcccca gcgtaaagct gatcaagcaa actctttgga
21001 ggaaggccat cacgagaaca catggtttta ttctcgagcc gggtttaccc agtgtgactg
21061 taaattacat ggacctttct gatttgcaca gacctcaggt ttctgggttc agagcatttt
21121 acagcaggga aaattaggga aaatcacagc tgatccaccc tttaggctaa aaactcaggc
21181 tgcaggttag cattccccaa cttggtattt ttaaattcag cctcagccac aagatccagg
21241 agacaccagc atctccagtt aggtagcagc tctgtgcttg cttgttctgt gaaaaactac
21301 cctgtccagc cactcctgca aacggcgctt gtcagggagc attctggctc aatggaatta
21361 tggaagtgta gccaaatatc aataatcatg tgtctggagc catagataat actgatgaac
21421 aaagatactg atttggtgac tcacagagag ttgcagctgt tgctacagat tgtcacaggt
21481 gatctgttcc aagggtccca ggtactggct gcatcggaat cacctaggga gcttgaacac
21541 aatttttgtg ttgtgcccat gcttggagat tccgcttcag gaggtctggg gtagacccag
21601 gaatctaggt gactatgatg tgtggctaca tttgggaaac tcagctctcc tcaactgtgt
21661 tcttttatag ggaggaaatt ggggctcaga gaggggaagt ggcccagcgt tctccaggta
21721 gtcataagtg gatgtgatcc tgcctgttct gtcgactccc aaaccaggct ttcttcccaa
21781 ttgctctctt ggcctcctct gaaacctcat tttagtcatg gtgctccccc tacgactgct
21841 tctagcaaag cactgagacc agaggagcac agcatgaggc cactcagccc agggtggctc
21901 ctgacactca catcatccgg cgactgtgat tctcagtcac agattggttg ggctatcctg
21961 ggacagacgt gggcagggag tgaaattgtc tttcatctgc ctcaggggca ttcctcctcc
22021 cctacaccat catggccatt tttgggggaa tcccgctctt ttacatggag ctcgcactgg
22081 gacagtacca ccgaaatgga tgcatttcaa tatggaggaa aatctgcccg attttcaaag
22141 gtaactgaag ggctcggtgg ggtgggtaag tggagcctct caggaaggga tttccatcag
22201 tgggtgggag ctggggacct tcttctggaa tcagtcaggc cttaggcatc cctgggagac
22261 tgcttgagta attaaggttt tgactgcgta agtggtgtga aattgtcaaa gagaaacaga
22321 ataaacggca atgagaagta aactaatgta aacattttaa agcattaaca ggacatgtac
22381 acattttcta aaacatgtta ccttgcaaat acatatgagt gactaacggc agggccactt
22441 ttcccacggt ggcctggagt ccttggaatg gaatggaggg cagcacccac gtgtacccct
22501 ctgctgggct gcggggctct ccagtgcccc gtttcagcct tttgccgttg ggccaccctt
22561 aaactcggcc tctgtctctc catctcaccc actgcccagg gattggttat gccatctgca
22621 tcattgcctt ttacattgct tcctactaca acaccatcat ggcctgggcg ctatactacc
22681 tcatctcctc cttcacggac cagctgccct ggaccagctg caagaactcc tggaacactg
22741 gcaactgcac caattacttc tccgaggaca acatcacctg gaccctccat tccacgtccc
22801 ctgctgaaga attttacacg taagtgcacg taagtgaggg gttggcctgt taggggcagg
22861 ccacacccct ggggtttggc ttcttaaaaa gaaggccagg gccaagggct cctcacccct
22921 ttctgagaga tgaagctatc aggggtggga gtagcagaat ttggagctgg gaaactctac
22981 ctcaagtgct ccacggcctg gccctgctgt gcctccttgc actttctcat gcgccaagtc
23041 agtttgctcc aatgcgccca agcccgtccc tcccttcaca ttccaaaatt actccccttc
23101 ttccggggtc ctcctctcag ccaaatgtca tggaggcgaa gcatcctagt actttgatag
23161 ttcaaaacta ttcaaatggt ttaatgtgtc aacttcgtaa gtgggatcta ccctttggaa
23221 gaagatgctt gaggaaggta gttttcttta atgctggaat ttcaggcaat ggcagcccaa
23281 agagagagct actttggggt agggaaggga gccaactccc cacacttaca catccctcat
23341 gttcccaaaa gccagtcaca ttggtcaaag gagttcaggt cccatccagc aaaggagaat
23401 ccaggaagtc tgcctttagg ttgaacactc ttcccagggt gttggccaaa tgtgctttgt
23461 ggtccttcag gacaggcagg gccaatccca gctcttcctg ggcctcagtg accacttggt
23521 cctagcagca ccgtgtccat ttgtcccggg caccctcaaa ggagcagcac agatgtgtgc
23581 ccaggtcact ctacagagcc tctcaggggc ctttctttct ccttcacccc aggcgccacg
23641 tcctgcagat ccaccggtct aaggggctcc aggacctggg gggcatcagc tggcagctgg
23701 ccctctgcat catgctgatc ttcactgtta tctacttcag catctggaaa ggcgtcaaga
23761 cctctggcaa ggtgaggaag gctctggctg ctgacccagg cctgcccagg ggcccacagg
23821 ctctcaaact caaaacccag tagcctggac agacagcact gtttccttcc acaccaaaca
23881 cattccatga tgcaccctct ctctcttcac cgcatttgtt agaaaattct agaagtgggt
23941 tgggcatggt ggctcacgtc tctaatccca gcactttggg aggctgaggt gggtggatca
24001 tttgaggcca ggagttcaag accagcctgg ccaacacatc gaaacctcat ctctacttaa
24061 acaaacaaac aaacaaacaa acaaacaaac aaaaaaaaac aatagccagg tgtggtggca
24121 ggtgcctgta atcccagcta ctcgggaggc tgaggcaaga gaatcacttg aacctgggag

24181 gtggagattt ctgtgagcca agatcatgcc attacactcc tgggcagcag agtgaaattc
24241 tgtctcacta aaaaaaaaaa ggaaaagaaa aaagaatatt ttctagaaat gggccaatgc
24301 caagaaccat tacatgattt ttttttttca agattccttc aacccactgg tagtaccaca
24361 attaataatc gtggcaatga cttactggaa acccatcatg gaccatgttc agtgcttttt
24421 atacactggc cagtgctatt gtttcctgct agatactccc tcaaccttac agatgcatag
24481 tgggctcaga ggtagttctc ctggccaacc tgaatccgat ccctgtgtga ctccaagggt
24541 tgtgatcttt ctgccacacc acctcgccct cctttctcaa ggtcttcaag aataaaagac
24601 actgacatcc attcacccac ctccctcccc tcccctcccc tcccccgctc ccttctcctt
24661 ctcccctccc ctcccctccc ctccctggaa cagcatggtg ataacccttc ttccttacag
24721 gtggtgtggg tgacagccac cttcccttat atcatccttt ctgtcctgct ggtgaggggt
24781 gccaccctcc ctggagcctg gaggggtgtt ctcttctact tgaaacccaa ttggcagaaa
24841 ctcctggaga caggggtaag atgatgaggg aaatatgtgt acctgtctgt cattttacta
24901 aacgaagctg gtcctcctcc aaataggaca gtagggacaa ctgaaatgtg gcttcagtgt
24961 gcaagtctgc aatgatactt taaccaagaa atacaagcat taaagtacag ttgaggccac
25021 tggatgggtt tatgcctgaa tcataggaag agaaatcaca tttgttggca cttggtttat
25081 ggctatgcct ggaatgagcc ttaaaagctg tttttctgcc tctgcactta gccacatgga
25141 ggcaggagtg gcgaccctgt tggtgccagg gtaagcctgt gtctcccccc caccccatgc
25201 ccttacccct gctgtgttcc aggtgtggat agatgcagcc gctcagatct tcttctctct
25261 tggtccgggc tttggggtcc tgctggcttt tgctagctac aacaagttca acaacaactg
25321 ctaccagtga gtatctgggg tccccaagca aagcttgaaa gtggcgtgag tgaaagtcag
25381 ctagtgtgga ggaaggtgat tggaccgacg gcctcgaatc tggccaggct taggccttga
25441 taccactgcc cctcaggggt gctgcaggcc cctgaccagc tcacccatgg ccagagcact
25501 catccaggcc gtggcttcac ctgggccact caggctgaga ggggccccat gcccttcggt
25561 tttgtcagcg ctcctgtata taaaaataat aagaggctgg acatggtggc tcacgcctgt
25621 aatcccagca ctttgggagg ccaaggcagg tggatcacct gaggtcagga gttccaggct
25681 aacctggcca acatggcgaa accctgtctg tactaaaaat acaaagatta gccaggcatg
25741 gtgttgcaca cctgtaagca cagccacttg ggaggctgag gcacgagaag ttgcagtgag
25801 ctgagatctt gccactacac tccagcctga gtgacagagc aagattccag aaagaaagaa
25861 agagagagag agaaagagag agagagaaat aaagaaagaa agaaagaaag aaaaaagaaa
25921 gaaagaaaga aagaaaggaa ggaaggaaag aaaaagaaag aaagaagaaa gaaggaagga
25981 aagaaagaaa gaaagaaaga gagagacggg gtgcggtggc tcccacctgt aatctcaaca
26041 ctttgggagg ccgaggcggg tggatcacga ggtcaggaga tcgataccat cctggctaac
26101 acggtgaaac cccatctcta ctaaaaatac aaacaaatta gccgggcgtg gtggcgggcg
26161 cctgtagtcc cagctactca ggaggctgag acaggagaat ggcgtgaacc caggaggtgg
26221 agtttgcggt gagccgagat tgtgccactg cactccagcc tgggcgacag agtgagattc
26281 catcaaaaaa aaaaaaaaaa aagaaagaaa gaaagagaaa ttgcaaagca agtttacaaa
26341 aagctgcaac atacaattaa gtagttaggt ttaactagtc aatgtgttta acaataaaaa
26401 ttgtaatgtg attatttgta ttatgtataa acaaaaagta tgatgatcac attattttga
26461 ctaacacaaa ataccataat ttttcttttc ttttcttttt ttttttttt ttttgagacg
26521 gagtttcgct cttgttgtcc aggctggagt gcaatggcgc aatctcagct caccacaacc
26581 tctgcctccc gggttcaagc gattctcctg cctcagcctc ccgagtagct gggatgacag
26641 gcatgcgcca ccacgactgg ctaaatttgt atttttagta gagatggggt ttctccatgt
26701 tggtcaggct ggtctccaac tcccgacctc aggtgatcca cctgcctcag cctcccaaag
26761 tgctgggatt acaggcgtga gccactgtgc ccagccaaaa tgccataatt tttgcaacaa
26821 aatcctattt cagaaatttt aacagaacag gaaatgttac accaattatg gttttaaata
26881 tcacccactg ggccaggtgc agttgctcaa acctgtaatc ttagcccttt gggaggctga
26941 ggcaggagga tcacttgagc ctaagagttt gagtccagcc tgggcaacat ggtgagatcc
27001 catctctaca aaaaatttaa aaattagctg agtgtggtgg tgcatgccta tggtcccagc
27061 tcttcaggag gctgaagcag gaggattgct tgagcccggg aggttaaggc tgcagtgagc
27121 tgtgttcatg ccactgcact tcagcctgtg caacagaggg agaccctatc tcaaaaataa
27181 ataaataaat aaataaataa atgtcaccca tcatctttag acttaatttt aattgatcac
27241 agatggcagc atccatcaat ctctttgtgc tatgcataag ataattgtaa aactgatttt
27301 tttaaaggtg ggagggagat tcacatgtac cctaagcaat gtggccttgc actgggacac
27361 caccctgcag ttgtatgatg aggctgtcat ttaaaagtgt gagatcagga gtttgagacc
27421 agcctggcca acatggggaa accccatctc tactaaaaat acaaaaatta gccaggcgtg
27481 gtcgtgggcg cctgtaatcc cagctacttg ggaggcggaa gcaggagaat cacctgaacc
27541 cgggaggcag agggtacagt gagctgagat caagccactg cactccagcc tgggcaatag
27601 agcaaggctc tgtctcaaaa aaaaaaaaaa aaaagaaaaa aaaatgaggg tatctctctg
27661 gctggcagtc tgtgacttag tttgaagata aagcctaaat ctgaggatct cacaaatgtg
27721 aggctcaggc cctgtcctgg cccccactcc aacctgggcc ctgctccaga acaaggctgg
27781 tttgaaatgg agaggtctgt gctagcagga ttgcctttta aagcacccct gactgctctt

27841 ctccctccac catgtgtggt ggtcatggca gtcagagctt caggcctcgg cagctgggtt
27901 ctgggagaag gcgaagggac tccgggacag gcttgaggtt ccttgagttc agcctgcttg
27961 tccttgctgt ggccgccctg tctaaggcag ggcaggcaca cggacagatg ctccttagac
28021 ccctgatctt ggaactgtct caggcggccc cttgggtttt cccctccaga gatgccctgg
28081 tgaccagcgt ggtgaactgc atgacgagct tcgtttcggg atttgtcatc ttcacagtgc
28141 tcggttacat ggctgagatg aggaatgaag atgtgtctga ggtggccaaa gacgcaggta
28201 ggacctcggg ttctatgcag gtcccttgct cctactgagt tgctttgctt tgaatctttg
28261 taaagatcta acttttgatt tatttctaaa tgttcctgaa attttccaaa gaaggctggg
28321 catggtggag ttcaagacca gcctgggcaa catagtgaga ccctgcctct ataaaaaaat
28381 aaatgagctg ggcatggtag tgcacaccta tagtcccagc tactcaggag gctgaggtgg
28441 gaagattgct tgagcctggg aggtcaaggc tacagtgagc catgatcaca ccactgtact
28501 tcagcctggg tgacagaata agaccctgtt tccaaaacat tttttattca aagatttgag
28561 agaagctcga tgccttagag atctttccat gtgttcaggc catgcaggga cagggcaact
28621 cccttggcac ccagttctgg ggcctccaaa atctatgggc ttgaaatgat gactctgatg
28681 atggctttaa aaccagtggg tttgcttcct gtgatggtct tgtgatggct caaagtctag
28741 taagcaaata tcatttatcc ttagaactgc cccttgcaga aaaagaaaat ctagccaata
28801 tgaccactct attttaccgt gaagtttttg tggaaaaaga aaagagcaac agagggagaa
28861 aaaacagatg atggcatcta gtgggcgagg tcccagggcc ttcttgttac aagcatgtta
28921 gcatgcttgt atctttggcc tgattcggtg ctttgaccaa ccccataccc caaacgaggg
28981 caccccctagc taacccatga gaggctgaag acatgtccct cacactttca aaagggactt
29041 taaaatagaa agaggtcttt attccaaagc ctctgtgact ctacggaaat caatttaaca
29101 tttcctactg ttctgaagct gaagacatca gaggggggcta actgtcatgg actttagaaa
29161 tatcaatcct ggtcctgaga aatcccatgg aattgttata tggtggaaac tcttattctc
29221 atgtgcagct ttgccagtct ctgggtcttg gatctgttac ccacccccag ggttaatgga
29281 gggcacacag ttttcaagat ggaagcccca cccttcctgc ctcccaggaa ctggggtggc
29341 cacgtcagac ggagtagagt gtgaggacgc acttggtgtg tgtcaaaacc tgattcacac
29401 ataaggtctt gtgatgagaa ttgtaactgt tgttgtggct gagttttcct ctgtgacatc
29461 tttgtaggac aggtcttgtc aaccacctcc tcctctctcc cctctgtctc aggtcccagc
29521 ctcctcttca tcacgtatgc agaagcgata gccaacatgc cagcgtccac tttctttgcc
29581 atcatcttct ttctgatgtt aatcacgctg ggcttggaca gcacggtgag gaaatgggaa
29641 ggggaaatgg gctgtgggag gatgaagagg aaaggaggag gactcagctg gactcaattc
29701 cctcccatcc ctggccatcc caggcagctt tggctactcc cttctggtgc tccagacccc
29761 cggctgactc ctcggcattt accaccctct ggtacattcg agtcattaat gggataaagt
29821 gggtcaagcg ccaggcgctt atgaggccca aggctcacat atgtgaactg ttattaagca
29881 ctgtcccttg gcccaattat ggcgcctaca ggccggggcg aggtctctcc tgggtctctg
29941 ctttgccaag ggttcagttc agggcctgga attattagtg tggagtgagt ctcagctttg
30001 agttggcctt tctggtgcaa acggagtctc cccgagagac ctcatatatt gaagttctgt
30061 ccacgtgtgc tattttgttt cctggctctc cttagtgatg ggctttaaat gacagtgtaa
30121 aacagggaaa tactctttga gatgcttaca tggtccagcg agccggggaa ctggtgatga
30181 tcagcttgac cagtcactta gcaaccttac ccctccctcc tgttcactgt cctgaaggcc
30241 acagccttgt cccaaagcct ccaccgtgac tgcttgcctg tacctctcag tttgcaggct
30301 tggagggggt gatcacggct gtgctggatg agttcccaca cgtctgggcc aagcgccggg
30361 agcggttcgt gctcgccgtg gtcatcacct gcttctttgg atccctggtc accctgactt
30421 ttgtgagtat catgcccctc agctttgagt tggcctcccc tagggcagca gaacgctaca
30481 ggaagctgtt cgtctccttc catctcacac agtgccctgg ctttgggacc aaaggagggg
30541 aggttgttac tcccattaaa aggtttgtca tcaaggagaa gaaaagttgt ttcttgtgca
30601 tctacttgca gattccgtgc atagtatcca ggggctgtca gcattttacc gcagatccca
30661 ctagatttca ggcttcaaaa actatctgga gaaaattagg atttgctttt ccagcctatt
30721 tccactcttt agagctgctt tcttgaggat actgccctga cgttcttcac gtcatgtgtg
30781 taatgccgct tggtctcctc cagaatttgc aatggtaaag aagtcacacc ctctctagag
30841 atccttttt ttttttttg gtagctgaag taaatactct aaggcagatt ttgcaaatgt
30901 gcatctttaa gccgtgaccc cgatggccct ggagtggggg tcaaggggac tgagtggaac
30961 cccagtgagt catgatttag gacagaaagg caaaactgtg accaaggccc ttaatgtgtc
31021 tggtatcctg ggcccttctg gattcagtct ttagaggcag ggattctcca gggtcgtgtc
31081 cacagcaggt gttcaataag ttgaagaaag tttaaaatag atcccaaggc cctttggagt
31141 ttgtctctta aaatttagtg gggtgaagaa atgagaagaa tgaactacat aggatgggaa
31201 gaggcgtctg agtgggtaaa acctgcccac actgccactg caaaggccag tccaagaaca
31261 gtccaagaaa cagtttcttt ttttttttt tttggacaga gtcttgctct gtcccccagg
31321 ctggagtgca atggtgctat ctcagctcac tgcaacctct gcctctcgtt caagcaattc
31381 tcccgcctca gcctcctgag tagctgggat taaaggcgtg tgccaccacg cctggctaat
31441 tttcaccata ttggccaggc tggtctcaaa ctcctgacct caggtgatac gcccgcctca

31501 gcctcccaaa gtgctgggat gacaggcatg agccaccgca cctggcctaa gaggcagttt
31561 ctggggtact cacgttcggt cccaggtggg atttttccag aatgattcta agtgaggaag
31621 cccctgcagg tctctgtagt tgggaggggt ggctccttgg agggccctca cccaggctct
31681 gctctcccct ccagggaggg gcctacgtgg tgaagctgct ggaggagtat gccacggggc
31741 ccgcagtgct cactgtcgcg ctgatcgaag cagtcgctgt gtcttggttc tatggtaaga
31801 cgcttccttc agccagctgg gtcccttcta caccatcagc cagcacagta gcaaagggct
31861 gatttcacca aaaaaggctg tgaggggcag tttacaagat gtgacctccc gatgatgggt
31921 ctgtaaaaga gtctcggcca ggcgcggtgg ctcacgcctg taatcccagc attttgggag
31981 gctgacacag gtggatcacc tgagatcagg agttcgagac cagcctggcc aacatggtga
32041 aactccatct ctactaaaaa tacaaaaaaa aaaaaaaatt agccgggcat gatggcgcaa
32101 gcctgtaatc tcagctactc aggaggctga ggcaggagaa ttgcttgaat ctgggaggca
32161 gaggttgcag tgagccaaga tcgcgccact gcactccagt ctggacaaca gagtgagagt
32221 ctgtctcgaa acaaacaaaa aacccaaaaa caaacaaaaa taaaccactt ttgcatgtta
32281 tttgttacac atattatcaa ggaggggggca cattttgggt tgcacacagg gctgcttttg
32341 ttttgttttg ttttgttttt taaagaagaa atcttcccac ctccttggca attcagacga
32401 ccaagttcaa gtgtgtgaat ctccatagac ctgccctagg aactgctggg acctctcggg
32461 gaagcaacca cagagcacct tccccttctc aatttctcac tcttctttgg gagttctcag
32521 ttctcacttc cttaccagtt acttttttata gaagaacctt gatcttgatc tggcatatga
32581 atgggcaata ttgtggttaa aaatttccct tgaaagtttc cagtgtcatg gtcaaagtca
32641 gtgtctgtca ggcgggttca ccattcctgg acaaaattca actgttttat aataacaatg
32701 aaatcagata cacactcagc tctcaggaag gcggtgtggc accaagaagg gggtgtgtgt
32761 ctcttcccag atgcccatcc atcctcccaa atgtccatcc atccttccag atgcccatcc
32821 gtcctcccac cctatgtaga aagagtcttg cttgctataa aaatgcagtt ttgtgacaag
32881 tcccccagcc ccactttcgg accagattgt tgtcgctgta ctcctcctag agagaaattc
32941 cagagctacc ctgtgtgtcc ggggtaggaa ttccggcttg tcatttcccc ttgtcattcc
33001 cagtgctgat cctgtgactc tgactctccc agagaatcag ggcacagcag cacagcccat
33061 gtcctttctt agtctctgcc tctcttcccc ttgggcggtt ccaggcatca ctcagttctg
33121 cagggacgtg aaggaaatgc tcggcttcag cccggggtgg ttctggagga tctgctgggt
33181 ggccatcagc cctctgtttc tcctggtgag ttgccatttc ccttctcctc cttgccctgg
33241 cgaaagactt cccccagaca cataaaaaat gaattgtaat ggctgggcac ggtggctcat
33301 ccttgtaatc ccaacacttt gggagtctga ggcaggcggc caacatggca aaaccccgtc
33361 tctacttaaa aaaatacaaa aattagccgg gtgtggtgat gggcacctgt aatcccagct
33421 actcaggaag ctgaggcagg agaatggctt gaacccagga ggcggaggtt gcagtgagct
33481 gaaattgcgc cactgcactc cagcctgggt gacaaagtga gactctattt cagaaaagaa
33541 ataaatagta ttgaagaata acatatacag aaaagtatta atatacagct cgtaagtgta
33601 catctcaaag aaatgttata aattgagcac acctatgtag ccaaaaccat accaggagac
33661 aaaaacatga ccagcacttc agaagcccct cttgtgcccc ctcccagcca ctatgcccac
33721 catggttccc cacccttctg actcctaaca ccatggagta gtctgggctt cagtcactct
33781 ggcagtgata gtgatgcata acatttatgg agcgctaaat gccaggccag acttcaggct
33841 caacttacac tcacttttct ataaatcctc acaatagttc tatgaagtgg ttatagttct
33901 ggcactccca ttttacattt aaggagattg aggctgggtg cagtggctca tgcctgttat
33961 cccagcactt tgggaggctg aggtggttgg atcacctgag gtcaggagtt caagaccagc
34021 ctggccaaca tggtgaaacc ccgtctctac taaaaatata aaaattagct gggcatggta
34081 ctaggtgcta ctcaggaggc tgaggcagga gaattgcttg aacctgggag gcggaggttg
34141 cagtgagctg agatcgcgcc attgcactcc agcctgggcg atagagcaag actctgtctc
34201 aaaaaaaaaa aaaaaaaaaa aaaagaaaag aaaagaaaaa aggagattga ggcttagcaa
34261 tgtaaagtga agggtccaaa gtcacataac aaaaagaact gttcaagacc agggaaggat
34321 ctgctatgct cttcatgggg gccctggatg gggtactagt acttctgtgt ttacaatttt
34381 gtatgcatct tcttcaagat gatacaaagc tctctaagca aaaatccacg tcaaaagtag
34441 aatccattta ttaccaaagg acctgggtga gaaagttgag tcccaagggt tgtccacgtg
34501 ttgagatgtg ttcaatattc actttgtcct tggttcacgt gcccctgtcc aagatggata
34561 acagcctgct ggcctgccaa gtgtgtgctg aagtagcctg gggtaaactg atgtcatgca
34621 caccagcagg cactattctt cagttctttc atttatctgt aactttgttt tttttgtgtt
34681 ttttttttt ttttgtctt ttagagccag aaatgttgtt agattcttcc ttcagtccac
34741 ctccagccta ttaggagtat atgggctatc ccattttgtt ttaggacatc tctttggcat
34801 ccagctgttc tgtgtcccct ttcaaaacag ctgggtgcag tggttcacac ctgtaatccc
34861 agcattttgg gaggccgagg caggtggatc ccttgaggtc aggagttcaa gaccagcctg
34921 ggcaacatgg cgaaaccctg tctttactaa aaatacaaaa attagcagga cgtgaacaag
34981 actcagtcta aacaaacaaa caaacaaaca aaaaacaaac atgaaaacag ccctttttgt
35041 attctgtgtt ttgactttct tggggtaacc ctttatttttt tgtacttctt ttgtacccac
35101 acatcttgac attaacttat ctatgatgga caacatgctt gttcttgtct ctgtgacact

35161 gcatgtgtct ctctgtaccc cagaattcct ctttgacttc ctttccccca gaccctggag
35221 gtttccaaaa tcaaagccag tgtttgcacc aagcagcaca tgttcccagg ctgcagtgac
35281 tctggtgcaa gctctgggct cctgagatgt tggcctcgga gctccttatg accttatcag
35341 cttgctgatg catttaggca catttaaaat gcgtgtgttt ttaaatccag aagttttagt
35401 tgttttcagt gagagggttg ttggaaatcg cctagcccat cctaccagaa gcagaagccc
35461 tgacataaat gggttattaa tttatattat taccagtact tctcaataat ctacttaaat
35521 atttttattt gtttttcaaa ggaagaagag tcattggatt aattccttga gggcagaaag
35581 atgtttgttt gtttgtttgt ttacttccat atttattccc cacgcagtga ggaaatgatg
35641 tctaggggag acttgaaggc aagtgaaaca ttcactgtgg gtataaaatt taagaaggca
35701 ccaaagagct caggactcag gataaatatt attttcacac aatatttttt aaaaataaaa
35761 attaattcaa aaagatccat gattaacaaa gtatcaacat tttcaataaa cacaggagcc
35821 agccctgtac ttgtacatcg tacttcactc gctttaccct catctggccc tgttagatcc
35881 tgtcttttaa aaaattgtaa taaaatatta tttattttga ttttggctcc tgctgaaatg
35941 tcgtgcctga ggtgagtgcc tccatgtgcc ctggggcttg gggctgttcg tttctgtcaa
36001 tttcctgctg tcagactggg tacactcttg atgcacagtc attgagtttc tttgcccagg
36061 aaaacccacg tgcagctttg gatgtgggcg cactactcat aattcccatg atgtttgcca
36121 tactcacccc tctacagcca agtgccctag tgcctgagag aggcaaatct gtcttccact
36181 tctacgttca cttaacagat cccctgctct gggcttctta aggctgcttt aggtaaacaa
36241 caagtagctt actaaaaata ttctttttca ctgcagccaa gttaaaaaaa agaaaatatt
36301 gtgtacgagt agaggccagg atgtgtgagg gggtttctga agtagcactt gccccaaccg
36361 gaggttgagt cccccaatga ctgagtggga tttcacacac ctgacattaa gagacaagca
36421 gccgggatgg ctcctaacca gatcttcttt ccctgttccc aaaccatctt tcttcatagc
36481 cggctctggg gatgaggagc ctgggttagg aggaaggttt gcaattgacc aggttcctgt
36541 tttgaaggct tccacctaga cttaagatag caccgctcag aagatggatg tgtgtttagc
36601 aatttcccat tttattacct gcagacaaag aaaaaaaaag atataaatag atgtttaacc
36661 acacaaataa ttcacattac tgtcttcttt atgactgtga aataataaaa taaaattaaa
36721 tcaacagcaa taacaatttc ggcacagtgg ttactaggca aaatctggaa accaaattga
36781 aataagaaaa ctcataaact gggtgttggg agcacactgt ggaatttttc ttggtcaaaa
36841 ttctcacaga gtgttcaagt gaaaggtgta ttaagctgct atttttcctt ccataattcc
36901 tttcatgatg aggacctgtg gcaggcagac attcattgtt gagatgccta gtggcagagc
36961 catccttgtt gaaactccag gcgatgcagc ttaatgcaga aggcagagag ctcaggaaaa
37021 tgggagaagg acacaccccc aggagcctga gtcctctctg ggtcttgaga accttcttat
37081 ctttcctgga catcaggcac cttccgtctg accaaaaaca accacagcca gaaacgcaag
37141 gcaatgaaaa caaagcagag agaaatatga agcctggggc ctggcggggg gctgcggaga
37201 ggtcacccag tttattcctc tgccttctgg caagacaagg cccaaattac tctcaaggtt
37261 gccagtctct ttgcctgggc agaggcctgt gctgaagtgg gagctgcagg cttcagcccc
37321 gactgtaggg gcagctgtag ggctgctgca acatccccct ataggaagtg agggctgggg
37381 aggaaaatca gggaaagcac ctggctggtg tagtcttgca gcttgggtac agaagttggt
37441 aattctgtgt gtcagagttc aatccaggat atatcacatc ttgggaattt cctgaaagta
37501 tcttagcttt agcatttggt gtttgtggta tcggaaacat ctctatctga gtggatattg
37561 ttaaggtttt taatgtttct tggaatttct gttttagttc atcatttgca gttttctgat
37621 gagcccgcca caactacgac ttttccaata taattatcct tactggagta tcatcttggg
37681 ttactgcata ggaacctcat ctttcatttg catccccaca tatatagctt atcggttgat
37741 catcactcca gggacattta aagaggtacg tgctagttag tgtgtatatg tgtgtgttca
37801 gactactaat tgttttggga gaaaaggatt atcttgttat gctctattca aaatgagatt
37861 caccaggagg catgcaaccc tatatatgta tgtacctgtg ccatggcgaa gacctctaat
37921 ggctgtcact gttatcttaa ggcatctccc aaagcctcaa atgcaagggc aaaagagtga
37981 gctgatcatt tgggtgcacc agagcattca gcttgggctg tgcataggtt ttcaaactta
38041 tcagacaact accctgcctc ctgcttgcca ctgggcatgg atatcagacc tgtgatgtct
38101 ggagcaatac acaaatttta agtggcaatt atggtgggaa atgtccatgg tcacttttct
38161 aacagcaatg tgaaaagtat tattttgaaa atcatcacaa aataaggatg ttgcaacagt
38221 atcatagttt tagatcaaaa ttattagcta tctaacaata gcatgatttt agctcaagcc
38281 attggagtaa aaatgttttg ccaggatagc aagattcatt tcggtaatga atgcgctgta
38341 gaaatgttaa taacagtacc ttacaatttg tatgaatgcc attttcaaag catctgcatg
38401 tcgattatcc tattctgtta tcaaagtacc ttgggaggta aatgggcagg cattattacc
38461 agacaaggac actgaggtgg ggaataatcc tgctcgaagt ttccttattt ccggctgggc
38521 acagtggctc acacctgtaa tctctgcatt ttaggaggcc aaggcatgtg gatcacctga
38581 ggtcaggagt tcaagaccag cctggccaac atggtgaaat cctgtctcta ctaaaattac
38641 aaaaattagc tgcgcatggt ggcacgtgcc tgtagtccca gctactcagc ataatctata
38701 atagcctgag aacggaggtg ggctaatttg agatcttacg caaaatttag acatgagtcc
38761 atcacaatag tcatcctaag acatttcata agaaatgaga agcaagagaa cataggcagt

38821 taaacatcag ctctttcttt tttttcagac aaggtctcac tctgtacccc aggctggagt
38881 gcagtggtat aatcatggct cactgcagcc tcaaactcct gggctcaggt gatcctttca
38941 cctcagcctc ctgactagct gggactacaa gcgcctggct aaatttttaa atttttaata
39001 gaaacaagat ctcactatgt tgcccaggct ggtctcaaac tcctgagctc aatcctcctg
39061 cctcagcttc ccaaagtgct gggattgccg gcatgagcca ctgcacccgg ccaaatatca
39121 gccctttctg agaacaactg cgttcatttc ttttttgaaa caggaaaact gtattctggg
39181 tgaacagaac tctgagctcc gtagttacta atttcctaca gccaaaccag ccctctcatg
39241 accgcactcc ccctaaagag ccttgggtcc cgtgctgctt cccgcacact ctcctaccca
39301 gagtcattga acgcattcct tccggtcatc tgtcttggca gttgtgatga tccatgagct
39361 ctgttgcttc aaatccagag gccaaggttg gaggtgggaa ctgtgttagg ttaataaatg
39421 tcattcttta gtcacaggct acacctcaga tcctggccat ctttctcaca aattccccca
39481 gttggccaca agggaatgcc acccaaaagt gtgtcactaa cagagctcag gtctgttagc
39541 aatagaggta agactaattg taggtagggc gtcttcacga ggggccacca gatggggagc
39601 agcgggttac ccggaagcag tgtgatacac ccctgtcggt gaggaccatg gaaaaggccc
39661 tagtcggctc tggcaggtac tcatcttatc tgggcctcgg tttcctcctc tttcaaatgt
39721 gctcatttga ccatgttaac atttcccaaa gtgtgttcca ggggctgtca caggtattac
39781 tgctggagaa aaagtttatg tagtcaggta atttttggag actgctggac taaggcaagt
39841 taaacaggtt tcttggccgg gcacagtggg tcacacctgt aatcccagca ctttggcacg
39901 ccgaggtgga tgggtcacct gaggtcagga gttcgagact agcctagcca acatggtgaa
39961 acccccgtct ctactaaaaa tgcaaaaatt agccgggtgt ggtggcgggc acctgtagtt
40021 ccagcttctc aggaggccga ggcatggaaa ttgcttgagc ctgggaggca gagattgcag
40081 tgagctgaga tcatcgagcc actgcactcc accctgggca acagagcgag actctgtctc
40141 aaaaaaatta aataagtaaa taaacaggtt tgtttacctg aggacttctc aaaaaaactt
40201 taatatgccc atgtgcatgt ttaatctctg agaagaataa agcatgctgt gtttttgcc
40261 aaaggtctag aattcttcac aattctttat gaataccaga cttgctagtc tgtaactttc
40321 tatgtgtgag ggagggaagc ctcatataaa atccttttgg gccctggcca tatgcaaggt
40381 tggagtttgt ggagtgagcc accgcttact ggtgtaattt gaagggtacc atacacaccc
40441 acggtctgga agtcttgcgg ttggagtcat ccactggctt aggtcatttt tgcagaaact
40501 actaataatc cacccaaaga tactggaaaa tggtccctaa tgagtacacg tttatgaaga
40561 tatgaagagg aatccatttc ttagattaac ttgttatcat cttgcagaag gccaggcagg
40621 gcctggcaaa tttccagcac aacggtagcc tggcaactag aatgctggca aagcacggaa
40681 aacagacgtg ggcctggcat cctgggtggt cagtcaccaa cagggcagca gcccatgtgg
40741 gtaggtaagg gcacaaaaac caagtgtttg actttatgac gtcgttcatt gcgggccttg
40801 ccccaccaaa cacctcggtg atctcttctc tctctgagct atggtgatag ctccacttaa
40861 gacaagttct aatagttgag catcgtacca tgcacctgta ggtccagcta cttgggaggc
40921 tgaggcagga ggatcccttg agcccagggg ttcaagactg cagtgaggca tgattgtgcc
40981 actgcactcc agcctgggca acagagcaag accctgactc ttaaaaaaaa aaaaaaagga
41041 ctcgttatag tgatctctgt ctagttctca actaacaagc cataattctg tttaatcagg
41101 tatagttcac atgttagtct gtgggaatgc atggtttctg acaagttgct ttccgtgact
41161 cggcatgctc tggtagctga tgactccttt ttaaacggat gactcctttt taaactccct
41221 gcaatgactc ctttttaaac ggatttctag tgtgaaagtt cagtacacag atactaaggt
41281 catgaggagc acaaaatgtg caattttgtg cattttgccc tatagtgcat ttcattaact
41341 gtaggactgg aacttttga attatagagg ccctatttct taaagaacaa actcccagaa
41401 tgagaaaata gaaaagcttg aaggcaactc cagacatttt cagtccaaaa taaaaatgac
41461 aggggcagaa cttcacctaa agatcagtaa atcaaaggat gctttcagat ccaaattgtc
41521 tgcttatgat acatacagac ttttggccaa caaagatttt attaagcaac ttaataaaga
41581 tttttcaaag cctctgcagt gtggctgttt ttttttttt ctcctctata tccctgagct
41641 gctttttgcc actctgatgc tgagttttac aaggccagtt gtatatttca cacctaagcc
41701 caagtatatt tcagatccat cttaaagctt ctgtgtacct cagctgtttc tatagtagac
41761 cttctcagcc acttctctgt gatgagtact acatcaacac aatgttgcat tagaagtgat
41821 gatttcaata atcgcatctt ttacagttta actatgtgat aatttatact ctcatatgaa
41881 ctaaccagta ttgctttaca ggtgaattca ttaaattagt aacatcctgg ctgttacagt
41941 ggctgatagc actgggcagt atggctgcag tgtgtctgca ggagacaagg agcctgggtg
42001 aggggatggg ggccaaaacc cagcaattac tctgctcccc aagccctttg tggccgcaga
42061 agtctcctct gtccagagaa aagggtaccc tttcccaagt ccttcaaaag catgttattg
42121 gccagcagag attctgttga aggctcatca ttttcttcca ttagtgttat aatgattttc
42181 ttagtttaca gatgtgtata gtttctgaat ataccatttt aattcatatt tgagaataca
42241 taaattggga attcttttcc taaggaatgt cagtgagact attccaactc gctcttagat
42301 gttattaaag tgttatttaa gctttgtttt aatattaatg ttgactattt ttgcaagttt
42361 taaaaattac aaggatgttt ataacattgt attttcttcc caatagcgta ttattaaaag
42421 tattacccca gaaacaccaa cagaaattcc ttgtggggac atccgcttga atgctgtgta

42481 acacactcac cgagaggaaa aaggcttctc cacaacctcc tcctccagtt ctgatgaggc
42541 acgcctgcct tctcccctcc aagtgaatga gtttccagct aagcctgatg atggaagggc
42601 cttctccaca gggacacagt ctggtgccca gactcaaggc ctccagccac ttatttccat
42661 ggattcccct ggacatattc ccatggtaga ctgtgacaca gctgagctgg cctattttgg
42721 acgtgtgagg atgtggatgg aggtgatgaa aaccacccta tcatcagtta ggattaggtt
42781 tagaatcaag tctgtgaaag tctcctgtat catttcttgg tatgatcatt ggtatctgat
42841 atctgtttgc ttctaaaggt ttcactgttc atgaatacgt aaactgcgta ggagagaaca
42901 gggatgctat ctcgctagcc atatattttc tgagtagcat atataatttt attgctggaa
42961 tctactagaa ccttctaatc catgtgctgc tgtggcatca ggaaaggaag atgtaagaag
43021 ctaaaatgaa aaatagtgtg tccatgcaag cttgtgagtc tgtgtatatt gttgtttcag
43081 tgtattctta tctctagtcc aatattttgg gcccattaca aatatatgaa ttccccaaat
43141 ttttcttaca ttaacaaatt ctaccaactc aattgtgtat ggaggttatt atttgaaggg
43201 tacaatcact acaacatgct ctgccaccca ctccttttcc agtgacacta cttgagccac
43261 acactttcct ttacaggcca gcctctggcg tttgctgcac ctcattgcca ccttcctgtc
43321 tctctgtgct aaacattcag gacagtgttc cacaggcaga tctggcctat ttcattagtc
43381 accatggctt ggctgtgaag tacgttgaag gtggatcttg tcacatgccc cttcagtgtt
43441 cacctggccc tctggtttaa gttctgtctg ccttacgtga ctgagtttga ctgtccaggt
43501 tgctttgctc ggtgaagaga ggagggtaaa tcggattctc gtttagcact gggttataca
43561 gatctggcac cctaacctaa accaaggcat cttcactcca agagcagttg gagagtctgg
43621 gttagcctta cgtggacctc gccgctcgct ggcggtcacg attgtgagcc ctccagataa
43681 tttttaaggt tgagtctaag taaggctgct tgggaaatgg tcagctaagt aaatcacctt
43741 tcatttcaca taaggccctt aatatagata agtaaatttg gcctttggtg tctcgtgact
43801 ctcagaggcg taggtagagg agcaaattaa tatttgcagc atgggaattc cttatcagaa
43861 ttttgagggg aataaatcct catcagagac aaaaggactt aatcatctgg ccacctatca
43921 cttcagttct ctgtataaat gaaatttaat tctaacaacc ttataaaaag aaggtccaga
43981 cagcagagga aacatcctgt ccaattctag gttttcctcc cttggcctcc tttccccagc
44041 attgtctacc ctggcccact tcctgcattc tccccatgcc ctgctatttc tgattctttg
44101 cttctcctag cgagatactt tccttatatg atagctgctg agaagtttcc cagaactgct
44161 agaggaaaag aagtggggaa tttaggaaat atccctcact gacctaactc cattatcttc
44221 actctttcct tcttcctgcc acctcatgcc cattctcttt actgtctagc atgctgaaag
44281 aaggaagtga tctaaatgcc agcgtgttca gtggtaaata ttagttggtg caaaagaaaa
44341 accatgatta cttttgcact aacctaatag ctttgcaaat tttaagaact tgctttatga
44401 agatattcgg atatggattc tccccacccc acatacttag acattgttca aatatactac
44461 ttttaaaaaa acacctttтc aaacagaatt agcgttttgc caagtctggt attaatggaa
44521 ttgtacagga gctttgaaag ttttcaaact ttattaaact aaaaaaaaaa aatcgaaaat
44581 ctctgtctgt tccgcatagt atgcatttat ttgaccccta tttatcaata ctatgatggg
44641 gtttttttтt tttaaagaaa atttaagagt aggtaggtgg attttaaaat aatattttaa
44701 agacctttta tatctatatg tagcatttat agaaaaataa aaactaaaaa tagaattgaa
44761 ttgtaacatt atttaaggac tgaagttttt tttcttgtat cagtaagaaa tacccaagag
44821 gctgggtgca gtgactcaca cctgtaatcc cagcactttg ggaggctgca gtgggaggat
44881 cacatgacat caggagtttg agaccagctt ggccaacata gtgaaacgcc gtctctatta
44941 aaaatacaga aaattagctg aatgtggtgg caggcgccta taattcctgc tacttgggag
45001 gctgaggcag gagaattgct tgaacccagg aggcagaggt tgcagtgagc caaacgttcc
45061 actgcattcc agcctggatg acaagagcga aactccgtct caaaaaaaaa aaaaaaaatt
45121 gttatgctta gttccatgga aagactattc tgaagcttta agtcttcttt ttctattttc
45181 catagtattg ccccttcccc acttcattgc ttaactgtct ctaaatttta atgataataa
45241 tattttaaag gtcagataat gccattagag gcagagacaa acctggggac caagctaatt
45301 tctctgttat tgagagcgct aaagacaggt gaactggagt tttatttcct ctgctagagt
45361 gaaataatac cctctccacc aggcacaatg gctcacacct gtaattccag cactttggga
45421 ggccaaggtg ggcggatcac ttgaggtcag gagttcgaga ccagcctggc caacatggtg
45481 aaaccccatc tctactaaaa atacaaaaat tagccagcca tggtggcatg ctcctgtaat
45541 cccagctact tgggagactg aggtgggata atcacttgaa cctgggaggc ggaggttgca
45601 gtgagctgag attgtgccac tgcactccag cctgggtgac agagcaagac tccatctaaa
45661 aaacaaaaac aaaaacaaaa aaaccctctc aattggttta ataccacgat agaaaagata
45721 aatattttag gatggaatct taaatatgtc tgtccttttg tttcatatag ttgaaatcaa
45781 ttcagtattg tttctacatt aggtgtttca aaaacagtca cctttgcaac agaagagctc
45841 ttttgttgaa aatgattcac aatatatttc agttggaatg tcaggtggta tttctcttac
45901 caacacctca cagaatctat gccaagctgc tctaccacca atgcaagtat ttttttaaag
45961 ctactaaaat aaactttctt caccagccaa ctctaatttg gagacagtgt gcattggtga
46021 aggacctcgt cagaataagt ttgaatgtct actgaactaa gaagaatttt gtcgtttggg
46081 ggagagaata gatggcatca gtccttcaat tctgtaactg aagactccaa ttatagtaga

46141 taagaattgt gtctagcaat ttttaaacac tgacagtcca aacaaaaata tttggtgagg
46201 aaggatgtcc catatttttg cttaaatatc ataaacaaat atgagcattt actaattttt
46261 taaatggcat tttgaaagat tattcttatt tacactctaa aattaaaggt gtactttatc
46321 ttaagaaaat gatatattaa aaattcatat tttaaaagat aaaattgggg aatttacagt
46381 tattttgtga atggccttta aactatgatt tgatctatat acaacttttc agaatacttt
46441 tgattgtgtg ttggacatat ctaaaattaa ttttatctgg cagaattaaa cctaaattta
46501 tttgtataaa agagtcccct ttctaaaatt catacagtgc ccttgtattc atttatgcag
46561 tgtttataaa tatttcaagt tagattgtga gatattaata aatgatttat gctgtgaact
46621 gttactgtaa taactcttaa tatcatgatt tgttgtatca ttaaaaaaac acaaaattct
46681 aaactcttaa ggttaaaaat cttaacctta tagtcaaaat atttaaacat ctccctaaac
46741 atcactgcat agcacttaca aaaactacat tgaattagat gtattcaatt tccataaagt
46801 taaataagtt aatataattc tttgtaattt gtgaagagct ttagtttata atgttcttcc
46861 ccatatattt catgggatca ccagaagtct aggtgacagc tgtattttcc caggtgagga
46921 tggggctggg atgaagagcc tgaagctcaa agaggctgat tgatttgccc aagttcacac
46981 acctggaacg ttccaggaca ggctttatgc ctataggttt tctaattcca gcccaaaggt
47041 tatttacatg tttagcttta aaattagcat ccatagaggc ttgagaaatc caaagagtat
47101 tcaactgcag aatgggtttt tttctaaact gcaaggatgt tctatctcat tttaaagtgt
47161 gatagggaag aataattaac ataggggaag aacctgttac tgatatatca ttacatgtgt
47221 gctgtcccca gccagagaca gcagcagtcc ccagcctttt tggcaccggg accggtttca
47281 tggaagaaaa tttttctacg gatagtgggg agcagagggg tatttcagga tgaaactatc
47341 ctacctcagt tcatcagaca ttagattctc ataagaagca cacaacctag atccctcgca
47401 tgcataattc acgacagggt tcacactctt gtgagaatct aatgctgccg ctgatgggac
47461 aggaggtgaa gttcaggtgg taatgagctg tgtggtccag ttcctaacag acctcctgcc
47521 ctacagttaa agactctcca gttaccatca ccacactaaa tctaatcctt gtactagatt
47581 caaattatct gatttacatt gacttgttta aagaggtatt aaggtaggga ttcccaaaat
47641 aaaggataac ataagcaaaa gacgggtgac aaaaagaaag aaaaacaaga gcgagaaaat
47701 aaaatgggac cagaaatgag gctacatgtc acaccagaat gactttcaca tgtgctgttg
47761 gtatgccaca aatctgatcc taggcttcat tagagataac tgttgcatcc agttatgaga
47821 tgccatttat tatcactaag gcttagtggt cactctgggt cctgctgctt ttccagggat
47881 attgctatta tttaaaataa atgctaatgt tgtcagtgct ttatagagag ggcaactcca
47941 gtattaaagt gtaatatttc aattacagga atctctttgg agtctcagct ccagttttgg
48001 aacatatcac taagtaataa ctaaagatct caaaatgatt accagtattt ttttttacac
48061 actttttttt ttctttcttt tttttagaca aagtgtcctg tcacccaggc tggagtgcag
48121 tggcgtgatc ttggctcact gtagccttga cttcctggac tcaagtgatc ctcccacctc
48181 agcctgctga gtagctggga ctacaggcat gcaccaccat gactggctaa ttctttgcat
48241 ctttaataga gaccgagttt tgccatgttg ctcaggctgg tctcatactc ctgggctcaa
48301 gcaatctgcc caccatggcc tcccaaagta ctgggattac aggcataagc cactgcaccc
48361 agtgacatgg tttttataaa ctaaaaaatt caatatgtaa aatggtggag aggatagtta
48421 atgctgagat ctgaataaat ggacttgaat aaaaagtaga atatatctta tagattctaa
48481 ctacaaaaat gaaaatgaga catgtatctc tgacaaatga agacattctc ctacataacc
48541 acaattaaac catcatagtt aggaaagtaa cactgctgca tgactgtcaa ctcatcttta
48601 gaccctgttc aatgatct

3. Primer Design:

OPRM1 (rs1799971)	Forward	GCTATACGCAGAGGAGAATG
	Reverse	ACATGACCAGGAAGTTTCCG

rs1799971 *OPRM1*

TGTGTTTGCA CAGAAGAGTG CCCAGTGAAG AGACCTACTC CTTGGATCGC TTTGCGCAAA
ATCCACCCCT TTTCCCTCCT CCCTCCCTTC CAGCCTCCGA ATCCCGCATG GCCCACGCTC
CCCTCCTGCA GCGGTGCGGG GCAGGTGATG AGCCTCTGTG AACTACTAAG GTGGGAGGGG
GCTATACGCA GAGGAGAATG TCAGATGCTC AGCTCGGTCC CCTCCGCCTG ACGCTCCTCT
CTGTCTCAGC CAGGACTGGT TTCTGTAAGA AACAGCAGGA GCTGTGGCAG CGGCGAAAGG
AAGCGGCTGA GGCGCTTGGA ACCCGAAAAG TCTCGGTGCT CCTGGCTACC TCGCACAGCG
GTGCCCGCCC GGCCGTCAGT ACCATGGACA GCAGCGCTGC CCCCACGAAC GCCAGCAATT
GCACTGATGC CTTGGCGTAC TCAAGTTGCT CCCCAGCACC CAGCCCCGGT TCCTGGGTCA
ACTTGTCCCA CTTAGATGGC
[A/G]
ACCTGTCCGA CCCATGCGGT CCGAACCGCA CCGACCTGGG CGGGAGAGAC AGCCTGTGCC
CTCCGACCGG CAGTCCCTCC ATGATCACGG CCATCACGAT CATGGCCCTC TACTCCATCG
TGTGCGTGGT GGGGCTCTTC GGAAACTTCC TGGTCATGTA TGTGATTGTC AGGTAAGGAA
AGCGCCAGGG CTCCGAGCGG AGGGTTCAGC GGCTTAAGGG GGTACAAAGA GACACCTAAC
TCCCAAGGCT CAATGTTGGG CGGGAGGATG AAAGAGGGGA GGTAAACTGG GGGGACTCTG
GAGGAGACCA CGGACAGTGA TTGTTATTTC TATGAGAAAA CCTACTTTTC TGTTTTTTCT
TCAACTGATA AAGAAAGAAT TCAAAATTTC AGGAGCAGAG AAGTTGCTTT GGTAAAAGCT
ACAAATGTCT AGGGGTGGGG GGCGGAGGGA AGCTATAGCA TAGACTTGGA GCGCTTCCTT
ATACTGAGCA AAGAGGGCTC

4. Primer Design:

COMT (rs4680)	Forward	AAAAGATAGGGACCAGCGTG
	Reverse	TTTTCCAGGTCTGACAACGG

s4680 *COMT*

AGAGGGCAGC TCTGTGTTAG GACACACTGG GGCCAGCCAG GAAGGGTGGA AAAGATAGGG
ACCAGCGTGA GCATAGAGGC TAAGGGACCA TGGGAGCTCC AAGCGCGCTC ACAGTGGGGA
CCAGGTCCTG GGGGCTGGGG ACACCAGGGA GGTGAAATAC CCCTCCAGCG GGTAGGGAGG
GTGGGCAGAG GAGGGCCAGC GGCCAGGCAT TTGGGAGGGG CTCCTGCTCT TTGGGAGAGG
TGGGGGGCCG TGCCTGGGGA TCCAAGTTCC CCTCTCTCCA CCTGTGCTCA CCTCTCCTCC
GTCCCCAACC CTGCACAGGC AAGATCGTGG ACGCCGTGAT TCAGGAGCAC CAGCCCTCCG
TGCTGCTGGA GCTGGGGGCC TACTGTGGCT ACTCAGCTGT GCGCATGGCC CGCCTGCTGT
CACCAGGGGC GAGGCTCATC ACCATCGAGA TCAACCCCGA CTGTGCCGCC ATCACCCAGC
GGATGGTGGA TTTCGCTGGC
[A/G]
TGAAGGACAA GGTGTGCATG CCTGACCCGT TGTCAGACCT GGAAAAAGGG CCGGCTGTGG
GCAGGGAGGG CATGCGCACT TTGTCCTCCC CACCAGGTGT TCACACCACG TTCACTGAAA
ACCCACTATC ACCAGGCCCC TCAGTGCTTC CCAGCCTGGG GCTGAGGAAA GACCCCCCCA
GCAGCTCAGT GAGGGTCTCA CAGCTCTGGG TAAACTGCCA AGGTGGCACC AGGAGGGGCA
GGGACAGAGT GGGGCCTTGT CATCCCAGAA CCCTAAAGAA AACTGATGAA TGCTTGTATG
GGTGTGTAAA GATGGCCTCC TGTCTGTGTG GGCGTGGGCA CTGACAGGCG CTGTTGTATA
GGTGTGTAGG GATGGCCTCC TGTCTGTGAG GACGTGGGCA CTGACAGGCG CTGTTCCAGG
TCACCCTTGT GGTTGGAGCG TCCCAGGACA TCATCCCCCA GCTGAAGAAG AAGTATGATG
TGGACACACT GGACATGGTC

5. Primer Design:

DRD4 (rs1800955)	Forward	CCCTTAGTGACAGACTACAGAAA
	Reverse	TAGTCCACCTGGTATCTGGCAAA

rs1800955 *DRD4*

TGGGGTCCCA CAGAGTGGTG CCCCCTTTTA GTGTCTTCTA GGCCCCTTAG TGACAGACTA
CAGAAAATAC CTCTCAGGTC ACAGGTCACC CCTCTTTGGT GAAGAGTCCA TAGAATTCTC
TGCTGCGCTT TGCAAGCACT TTCTCTTCTG CACGTTTGGA ACCTACCCCG GCCTGTCGTG
TCTTTCTCCT GGCCTCCTCG CGAGCCGAAC CTACTGTCCG GTCCCGGGAC CCCCTGCCCA
GGGTCAGAGG GGCGCCTACC TAGCTCACGG TCTTGGGCCG GAGGGAATGG AGGAGGGAGC
GGGGTCGACC GCTCAGCTGT CCGCCCAGTT TCGGAGGCGG CCACGCGAGG ATCAACTGTG
CAACGGGTGG GGCCGCGGCT GACCGTGGTG GTCGCGGGGG CTGAGGGCCA GAGGCTGCGG
GGGGGGGGCG GCGGGATGAG CTAGGCGTCG GCGGTTGAGT CGGGCGCGGA GTCGGGGGCA
GGGGGAGCGG GCGTGGAGGG
[C/T]
GCGCACGAGG TCGAGGCGAG TCCGCGGGGG AGGCGGGCAG AGCCTGAGCT CAGGTCTTTC
TGCGTCTGGC GGAACGGGCC TGGGAGGGAG GTTTTGCCAG ATACCAGGTG GACTAGGGTG
AGCGCCCGAG GGCCGGGACG CACGCACGGG CCGGGTAGGA TGGCGCTGGC GTCGATGCCC
GCGCGCTTCA GGGCCTGGTC TGGCCGCCCC TCCATCCTTG TCGGTTTCTC GGGTCGCGGA
CCCCGCGCGG CGCCGGGCGA TGCTGGCCTG CCCGTGGCCA CCACCTCGCT TCATTCCCGT
CTCTTTGGGC CGCCGCATTC GTCCACGTGC CCGTCTCTCC CTGCGCAAAA TTCCAAGATG
AGCAAATACT GGGCTCACGG TGGAGCGCCG CGGGGGCCCC CCTGAGCCGG GGCGGGTCGG
GGGCGGGACC AGGGTCCGGC CGGGGCGTGC CCGAGGGGAG GGACTCCCCG GCTTGCGACC
CGGCGTTGTC CGCGGTGCTC

6. Primer Design:

5HTR2C (rs3813929)	Forward	TCCAGATGCACTAAGAGACC
	Reverse	GCTAGGTTTGGCTCTTTTCG

rs3813929 *5HTR2C*

GACAAGGATG GGGAAGTGGG CCTTATAACA GGATTGTGGC CTTTGCGCAC TCACCAAATG
TTTGACCCTG TGAGTGCCTC AGTTGCTACT GTTGGAAGAA TGGGCAAGAG TCGGAACAGA
GACCCTTGAA GGGAGTTTCA AAGCTTGATG AAATTTGCAA GACTTGAGAA TGCTGTTTGT
TGAAATGAAA TGTACAGGGG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TTTGGGGAGG
GGTATGCTAT GAATCTTTGA GGGTACATTC TTGAGAAAGC CTTCCCTCTC TCTCTATCCG
GTGCCATGGC TGATCCTGGT TCCCCCTACT CTCTAGGCCT TGTGAATCAG ATTAATCATC
ACCCCCACCC CCATCTCCAC CATGGGGTCT CGCGCCCCCT GCCAGCAGGC TCCAGATGCA
CTAAGAGACC GGTCCAAACA GGCCCGGGGG CCACGTAATG CTGAGTGCTG ATTGGCTGCT
CTTGGCTCCT CCCCTCATCC
[C/T]
GCTTTTGGCC CAAGAGCGTG GTGCAGATTC ACCCGCGCGA GGTAGGCGCT CTGGTGCTTG
CGGAGGACGC TTCCTTCCTC AGATGCACCG ATCTTCCCGA TACTGCCTTT GGAGCGGCTA
GATTGCTAGC CTTGGCTGCT CCATTGGCCT GCCTTGCCCC TTACCTGCCG ATTGCATATG
AACTCTTCTT CTGTCTGTAC ATCGTTGTCG TCGGAGTCGT CGCGATCGTC GTGGCGCTCG
TGTGATGGCC TTCGTCCGTT TAGAGTAGTG TAGTTAGTTA GGGGCCAACG AAGAAGAAAG
AAGACGCGAT TAGTGCAGAG ATGCTGGAGG TGGTCAGTTA CTAAGCTAGA GTAAGATAGC
GGAGCGAAAA GAGCCAAACC TAGCCGGGGG GCGCACGGTC ACCCAAAGGA GGTCGACTCG
CCGGCGCTTC CTATCGCGCC GAGCTCCCTC CATTCCTCTC CCTCCGCCGA GGCGCGAGGT
TGCGGCGCGC AGCGCAGCGC

7. Primer Design:

LEPR (rs1137101)	Forward	CCTGCTTTAAAAGCCTATCCAGT
	Reverse	ACCCCCAGTACTACATCTACCAT

rs1137101 *LEPR*

```
CTTTGGTATG TCTGAAAAAA AAAGCCTTTA TTTCATCATT ATTTTGAAAG CTGTTTTCGC
 TGGGTATAGG ATTTTAGAAT TGCAGTTTTT CTTTTATTTT AGTACTTCAC TTTTACGTCA
 TTATCTTTTT GCTTATGTTA TTCCTGATGA TTAACCTGCT GTAATCTTTA TCTTTGTTTT
 TCTAATGTAG GGTTTTTTTT TTTCAGATAC CCTTTAAGCT GGGTGTCCCA AATAGTTTAC
 TTCAATTAGT ATTTAGTATC CTGCTTTAAA AGCCTATCCA GTATTTTCAT ATCTGTTTTA
 ATATTTAGCT CTTATTTTTC AATATAGGCC TGAAGTGTTA GAAGATTCAC CTCTGGTTCC
 CCAAAAAGGC AGTTTTCAGA TGGTTCACTG CAATTGCAGT GTTCATGAAT GTTGTGAATG
 TCTTGTGCCT GTGCCAACAG CCAAACTCAA CGACACTCTC CTTATGTGTT TGAAAATCAC
 ATCTGGTGGA GTAATTTTCC
 [A/G]
 GTCACCTCTA ATGTCAGTTC AGCCCATAAA TATGGGTAAG TTATGCACTA AAATGATGAT
 AATAGGTCTA AACATCAGTC ATATATAAAG GTTAAAAATT GCTTACAAAA ATATTTGCTA
 GCTTATCTCA CTTTGCTTAA CACTGTAATG ATGGTAGATG TAGTACTGGG GGTATTAAGA
 GTGGCTTCTA GAATGATTTA ACAATGGTAT GTATATCTCT GCCATTGTCA CTTAAATTCT
 GTTTTGAAAA CTGTTTTCTT TCAATCCTGG ATCTATGTAA TGGATGTATA TTGATTGGAT
 ATCACTTTTT CACATCTCAG ATAACTATTT TTGAAAATAG TAGCATGTTT CTTGCCTGAA
 TTTATTCCTT CAATAAATAT TTCTTAGAGG CTCATGTTTG TCAGAGACTG CTCCAGGAGC
 TGGAAAAAGA GTGGGACATT AGACATAGTT CCCACCTCAG AGAGCAGGGA CAAACAATAG
 TAGGCAGAGA CAATGATAAA
```

8. Primer Design:

GABRA6 (rs3219151)	Forward	CAGTGTTGAATAGCTTGCGG
	Reverse	CTAGCATAAAGGCCTCCAAC

rs3219151 *GABRA6*

```
ATATTTGTCA ATGGTGAAAG AGTGAATAAA TAAGCAATTA AGCAATATCT ATTCTTTCAT
 TTGGGCTTAA TATTTGTCTT TTTTCCACAG CATCCTGACT CCAAATATCA TCTGAAGAAA
 AGGATCACTT CTCTGTCTTT GCCAATAGTT TCATCTTCCG AGGCCAATAA AGTGCTCACG
 AGAGCGCCCA TCTTACAATC AACACCTGTC ACACCCCCAC CACTCTCGCC AGCCTTTGGA
 GGCACCAGTA AAATAGACCA GTATTCTCGA ATTCTCTTCC CAGTTGCATT TGCAGGATTC
 AACCTTGTGT ACTGGGTAGT TTATCTTTCC AAAGATACAA TGGAAGTCAG TAGCAGTGTT
 GAATAGCTTG CGGCCAGGAC AACCTGAATT CTATAAGTTC TTGTTTTCTG TTTCCTATGT
 TTTCTTAAAA AATAGCATTG AGACTTGTGT AGATGCTTCT CAGAACATGA AATCAAATTG
 GAAATCTGTA ACGCAGCTTC
 [C/T]
 GTAAGCATGT GTGGGCAAAA AAGCAATAAT CCTACTCCTC AAAATAGAAA GTTGAAGATT
 GCTGAAAAAT ATGACTTTTC TGTATGTTAG AGAAAAACTT TATGAGGATG AAATGGGTTC
 AAGATGAATT TGTCAACTTT TGTCTTCCAT TGTTCAGTAT TTTTAATTGT CACTGTAAAT
 AACATTTACC ACAAGGCAGA TAAAATAAGA AATGCTGACA CTTCCAAAGG TTGCCTTAAA
 ATATGTTTAT TTTGGCTTAG TTCCCGAGAG GGCAAAATAT AAATACAGTC TAAATATTTA
 TCAGTAGGTT AATACCAGCA TGTTGGAGGC CTTTATGCTA GTAAAATGGC TTTCAGTGGC
 ATTGTAAAGC CTACATTGAG CTTAGCCATT TGTTTTTAAC CTCGCTGTGC TCTTTTACCT
 CAATAAAATG TGGTGTTTGT ATACATATAA ATTATACATA GCTCATAAAT TATGTATGCA
```

Appendix B3: Sequencing information from Inqaba Biotechnology

REF: SA2018/53140

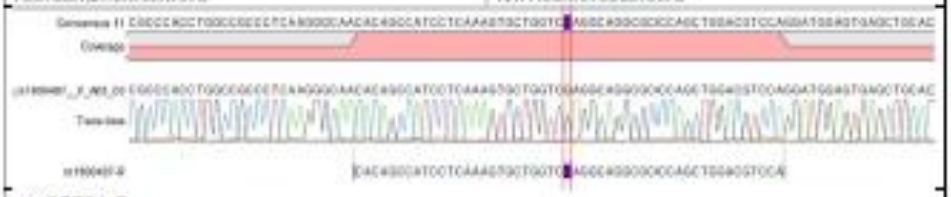
All genotypes were determined with Sanger Sequencing utilizing Brilliant Dye V3.1 (Nimagen) and the ABI 3500 XL (ThermoScientific).

Mutation Screened	Genotype
rs1800497-R	G/G
rs1799971-R	A/A
rs4680-R	G/A
rs1800955-Y	C/C
rs3813929-Y	C/C
rs1137101-R	A/G
rs3219151-Y	C/T
rs25531-R	Typing unsuccessful**

**Genotyping of SNP rs25531 was unsuccessful using standard Sanger methodologies. The failure was likely due to the high GC content and repetitive nature of the region.

Dr Christiaan Labuschagne
Genomics Manager
Inqaba Biotechnical Industries (Pty) Ltd P.O. Box 14356, Hatfield 0028, South Africa
Phone: +27 12 343 58 29
E-mail: christiaan.labuschagne@inqababiotec.co.za
Web: www.inqababiotec.co.za

REF: SA2018/53140

>rs1800497-R CACAGCCATCCTCAAAGTGCTGGTCRAGGCAGGCGCCCAGCTGGACGTCCA	
Genotype	G/G
Forward Primer: ACCTGGAGATCATCCATCTG	Reverse Primer: AATTTCCATCTCGGCTCCTG
	
>rs25531-R CTCGCGGCATCCCCCCTGCACCCCCRGCATCCCCCCTGCAGCCCCCCCAGC	
Genotype	Undetermined
Forward Primer: V1-GTTGCAGGGGAGATCCTGGGAGAGG V2-GGCGTTGCCGCTCTGAATGC V3-GGTAGGGTGCAAGGAGAATGCTGGAG V4-CTGAAGAGGAATCGGCTCTGGGC	Reverse Primer: V1-CCTCCTAGGATCGCTCCTGCATCC V2-GAGGGACTGAGCTGGACAACCAC V3-CTGCAACCTCCCAGCAACTCCCTGTAC V4-CGATGTTCACTCCAAATGATGTGC
>rs1799971-R GGTCAACTTGTCCCACTTAGATGGCRACCTGTCCGACCCATGCGGTCCGAA	
Genotype	A/A
Forward Primer: GCTATACGCAGAGGAGAATG	Reverse Primer: ACATGACCAGGAAGTTTCCG
	
>rs4680-R CCAGCGGATGGTGGATTTCGCTGGCRTGAAGGACAAGGTGTGCATGCCTGA	
Genotype	G/A
Forward Primer: AAAAGATAGGGACCAGCGTG	Reverse Primer: TTTTCCAGGTCTGACAACGG
	
>rs1800955-Y GGGCAGGGGGAGCGGGCGTGGAGGGYGCGCACGAGGTCGAGGCGAGTCC	
Genotype	C/C
Forward Primer: GGCTTAGTGACAGACTACAGAAA	Reverse Primer: TAGTCCACCTGGTATCTGGCAAA

SOUTH/EN AFRICA
Inqaba Biotechnical Industries (Pty) Ltd
PO Box 14335, Hatfield, 0028
Pretoria, South Africa
Tel: +27 12 345 5829
Fax: +27 86 677 8493
E-mail: info@inqababiotec.co.za

EAST AFRICA
Inqaba biotec East Africa Ltd
PO Box 1845
Nairobi 00606, Kenya
Tel: +254 735 970 695
Fax: +27 86 677 8409
E-mail: info@inqababiotec.co.ke

WEST AFRICA
Inqaba biotec West Africa Ltd
PMB 5330, Oyo Road, Ibadan 200001
Oyo State, Nigeria
Tel: +234 805 892 7272
Fax: +27 86 677 8409
E-mail: info@inqababiotec.org

>rs3813929-Y
CTGCTCTTGGCTCCTCCCCTCATCCYGCTTTTGGCCCAAGAGCGTGGTGCA

Genotype	C/C
Forward Primer: TCCAGATGCACTAAGAGACC	Reverse Primer: GCTAGGTTTGGCTCTTTTCG

>rs1137101-R
ATCACATCTGGTGGAGTAATTTTCCRGTCACCTCTAATGTCAGTTCAGCCC

Genotype	A/G
Forward Primer: CCTGCTTTAAAAGCCTATCCAGT	Reverse Primer: ACCCCCAGTACTACATCTACCAT

>rs3219151-Y
AATTGGAAATCTGTAACGCAGCTTCYGTAAGCATGTGTGGGCAAAAAAGCA

Genotype	C/T
Forward Primer: CAGTGTTGAATAGCTTGCGG	Reverse Primer: CTAGCATAAAGGCCTCCAAC

SOUTHERN AFRICA
Inqaba Biotechnical Industries (Pty) Ltd
PO Box 34325, Hatfield, 0028
Pretoria, South Africa
Tel: +27 12 345 5829
Fax: +27 86 677 8409
E-mail: info@inqababiotec.co.za

EAST AFRICA
Inqaba biotec East Africa Ltd
PO Box 1849
Nairobi 00606, Kenya
Tel: +254 735 570 655
Fax: +27 86 677 8409
E-mail: info@inqababiotec.co.ke

WEST AFRICA
Inqaba biotec West Africa Ltd
PMB 5320, Oyo Road, Ibadan 200001
Oyo State, Nigeria
Tel: +234 805 082 7232
Fax: +27 86 677 8409
E-mail: info@inqababiotec.ng

Dr Christiaan Labuschagne
Genomics Manager
Inqaba Biotechnical Industries (Pty) Ltd P.O. Box 14356, Hatfield 0028, South Africa
Phone: +27 12 343 58 29
E-mail: christiaan.labuschagne@inqababiotec.co.za
Web: www.inqababiotec.co.za

Appendix B4: Summary of Genotyping results

No	Age	Weight (kg)	Length (m)	BMI	BMI Class	Race	Gender	COMT	DRD2	DRD4	GABRA6	HTR2C	LEPR	OPRM1	SLC6A4
								rs4680	rs1800497	rs1800955	rs3219151	rs3813929	rs1137101	rs1799971	rs25531
M01	24	72,10	1,71	24,66	Normal	W	F	A/A	A/G	C/T	C/T	C/T	A/A	A/G	A/A
M02	37	65,00	1,59	25,71	Overweight	B	F	A/G	G/G	C/T	T/T	C/T	A/G	A/A	A/A
M03	58	55,30	1,56	22,72	Normal	W	F	A/A	G/G	C/T	C/C	C/C	A/G	A/G	A/A
M05	26	55,80	1,60	21,80	Normal	W	F	A/G	G/G	C/C	T/T	C/T	A/G	A/A	A/A
M06	25	75,30	1,53	32,17	Class I	W	F	A/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
M07	47	61,90	1,72	20,92	Normal	W	F	A/G	G/G	C/C	C/C	C/C	A/G	A/G	A/A
M08	32	68,80	1,61	26,54	Overweight	W	F	A/A	A/G	C/T	C/T	C/T	G/G	A/A	A/A
M09	39	98,90	1,89	27,69	Overweight	W	M	A/A	A/A	C/C	T/T	C/C	A/G	A/G	A/A
M10	37	81,20	1,72	27,45	Overweight	W	M	G/G	G/G	C/T	C/C	T/T	G/G	A/A	A/A
M11	30	79,10	1,60	30,90	Class I	W	F	A/G	G/G	C/T	T/T	C/C	A/A	A/A	A/A
M12	33	93,00	1,82	28,08	Overweight	W	M	A/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
M13	22	75,60	1,71	25,85	Overweight	W	F	A/G	A/G	C/T	C/T	C/T	A/G	A/A	A/A
M14	25	81,60	1,81	24,91	Normal	W	M	A/G	G/G	C/C	T/T	C/C	G/G	A/G	A/A
M15	50	58,10	1,57	23,57	Normal	W	F	A/A	G/G	C/T	C/T	C/T	A/A	A/A	A/A
M17	33	59,70	1,72	20,18	Normal	B	M	G/G	G/G	C/T	C/T	C/C	G/G	A/A	A/A
M18	27	56,60	1,59	22,39	Normal	B	F	G/G	A/G	C/T	C/C	C/C	G/G	A/A	A/A
M19	35	54,50	1,67	19,54	Normal	B	M	G/G	A/G	C/T	C/C	C/C	G/G	A/A	A/A
M20	34	79,50	1,79	24,81	Normal	B	M	A/G	G/G	T/T	C/T	C/C	G/G	A/A	A/A
M21	34	56,50	1,56	23,22	Normal	B	F	A/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
M22	21	68,20	1,70	23,60	Normal	B	M	A/A	A/G	C/T	C/T	C/C	A/A	A/A	A/A
M23	25	49,20	1,60	19,22	Normal	I	F	A/G	A/A	C/T	C/T	C/C	A/G	G/G	A/A
M24	29	55,10	1,66	20,00	Normal	B	F	A/G	A/G	C/T	C/T	C/C	G/G	A/A	A/G
M25	65	102,80	1,73	34,35	Class I	W	M	A/G	G/G	C/T	C/C	C/C	G/G	A/A	A/A
M26	34	63,30	1,75	20,67	Normal	B	M	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
M28	45	77,60	1,70	26,85	Overweight	B	M	A/G	A/A	T/T	T/T	C/C	A/G	A/A	A/A
M29	32	80,90	1,60	31,60	Class I	W	F	G/G	A/G	C/C	C/T	C/C	A/G	A/A	A/A
M30	30	70,00	1,65	25,71	Overweight	W	F	A/A	A/A	C/T	C/C	C/C	A/A	A/G	A/A

M31	31	59,60	1,69	20,87	Normal	W	F	A/A	A/G	C/T	C/T	C/C	A/A	A/A	A/A
M32	48	69,00	1,74	22,79	Normal	W	F	A/G	A/G	C/T	C/T	C/C	G/G	A/A	A/A
M33	53	70,50	1,63	26,53	Overweight	W	F	A/G	G/G	C/T	C/T	C/C	A/G	A/G	A/A
M34	22	71,10	1,72	24,03	Normal	W	F	A/G	G/G	C/T	C/T	C/T	A/A	A/G	A/A
M35	31	121,80	1,60	47,58	Class III	W	F	A/A	G/G	C/T	C/T	C/C	A/G	A/A	A/A
M36	25	67,10	1,65	24,65	Normal	W	M	A/G	G/G	C/T	C/C	C/C	G/G	A/A	A/A
M37	33	57,00	1,62	21,72	Normal	B	F	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
M38	52	55,80	1,59	22,07	Normal	W	F	A/A	G/G	C/T	T/T	C/C	A/G	G/G	A/A
M39	48	64,10	1,62	24,42	Normal	W	F	A/A	G/G	T/T	C/T	C/C	A/A	A/A	A/A
M40	30	61,40	1,74	20,28	Normal	W	F	G/G	G/G	C/C	T/T	C/T	A/A	A/A	A/A
M41	25	68,90	1,80	21,27	Normal	W	F	G/G	A/G	C/T	C/T	C/T	A/G	A/A	A/A
M42	35	72,40	1,73	24,19	Normal	W	F	A/A	G/G	C/T	C/T	C/C	A/G	A/G	A/A
M43	27	58,20	1,62	22,18	Normal	W	F	A/G	A/G	C/T	C/T	C/C	G/G	A/A	A/A
M44	26	67,10	1,64	24,95	Normal	W	F	A/G	A/G	C/T	C/T	C/C	A/A	A/A	A/A
M45	21	70,70	1,59	27,97	Overweight	W	F	A/G	G/G	C/T	T/T	C/C	A/A	A/A	A/A
M46	27	69,00	1,74	22,79	Normal	W	F	A/G	G/G	C/T	T/T	C/C	A/G	G/G	A/A
M47	64	68,40	1,65	25,12	Overweight	W	F	A/A	G/G	C/T	C/T	C/C	A/G	A/G	A/A
M48	26	65,10	1,65	23,91	Normal	W	F	A/A	G/G	T/T	C/T	C/T	A/G	A/G	A/A
M50	30	59,90	1,64	22,27	Normal	W	F	A/G	A/G	T/T	T/T	C/C	A/A	A/A	A/A
M51	22	56,20	1,65	20,64	Normal	W	F	A/G	A/A	C/T	C/T	C/T	A/A	A/A	A/A
M52	22	63,70	1,52	27,57	Overweight	W	F	A/A	G/G	T/T	C/T	C/T	A/A	A/A	A/A
M54	21	65,50	1,65	24,06	Normal	W	F	A/G	G/G	C/C	T/T	C/T	G/G	A/A	A/A
M55	60	80,80	1,60	31,56	Class I	B	F	A/G	A/G	C/T	C/T	C/C	A/G	A/A	A/A
M56	50	84,10	1,85	24,57	Normal	W	M	A/G	G/G	T/T	T/T	C/C	A/A	A/A	A/A
M57	44	70,70	1,52	30,60	Class I	W	F	A/G	A/G	C/T	T/T	C/C	G/G	A/A	A/A
M58	34	57,20	1,58	22,91	Normal	B	F	G/G	A/G	T/T	T/T	C/C	G/G	A/A	A/A
M59	36	72,10	1,77	23,01	Normal	W	F	A/A	G/G	C/T	T/T	C/C	A/A	A/A	A/A
M60	58	67,40	1,72	22,78	Normal	W	F	A/G	G/G	C/C	C/T	C/C	A/A	A/A	A/A

M61	37	54,50	1,58	21,83	Normal	W	F	A/A	G/G	C/T	C/T	C/C	G/G	A/G	A/A
M62	31	63,00	1,62	24,01	Normal	W	F	A/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
M64	26	63,70	1,71	21,78	Normal	B	F	A/G	G/G	T/T	T/T	C/C	G/G	A/A	A/A
M66	27	63,10	1,78	19,92	Normal	B	M	G/G	A/G	C/T	T/T	C/C	G/G	A/A	A/A
M67	35	51,10	1,67	18,32	Normal	B	M	G/G	G/G	T/T	C/T	C/C	G/G	A/A	A/A
M68	21	53,50	1,71	18,30	Normal	B	M	A/A	A/A	T/T	C/C	C/C	A/G	A/A	A/A
M69	39	79,10	1,63	29,77	Overweight	B	F	G/G	A/G	C/C	C/T	C/C	A/G	A/A	A/A
M70	32	75,20	1,56	30,90	Class I	B	F	G/G	A/A	T/T	C/C	C/C	A/G	A/A	A/A
M71	27	64,00	1,74	21,14	Normal	B	M	A/G	A/G	C/C	T/T	C/C	A/G	A/A	A/A
M72	31	72,60	1,64	26,99	Overweight	B	M	G/G	A/G	C/T	T/T	C/C	A/A	A/A	A/A
M73	38	63,70	1,79	19,88	Normal	B	M	A/G	A/A	T/T	T/T	C/C	A/G	A/A	A/A
M74	27	73,90	1,59	29,23	Overweight	B	F	G/G	A/G	T/T	C/T	C/C	G/G	A/A	A/A
M75	39	84,10	1,60	32,85	Class I	B	F	G/G	G/G	T/T	T/T	C/C	A/A	A/A	A/G
M76	37	72,10	1,78	22,76	Normal	B	M	G/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A
M77	20	61,00	1,60	23,83	Normal	W	F	G/G	A/A	C/C	C/T	C/C	A/A	A/A	A/A
M79	21	81,50	1,70	28,20	Overweight	W	F	A/G	G/G	C/T	C/T	C/C	A/A	A/A	A/A
M80	49	73,70	1,63	27,74	Overweight	W	M	A/G	A/G	C/T	C/T	C/C	A/A	A/A	A/A
M81	63	71,80	1,68	25,44	Overweight	W	F	A/G	G/G	C/T	C/T	C/C	A/G	A/G	A/A
M82	43	78,60	1,77	25,09	Overweight	W	F	G/G	G/G	T/T	T/T	C/C	A/G	A/G	A/A
M83	44	80,30	1,52	34,76	Class I	W	F	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A
M84	61	65,40	1,69	22,90	Normal	W	F	A/A	A/G	C/T	C/T	T/T	A/A	A/A	A/A
M85	72	73,30	1,65	26,92	Overweight	W	F	A/G	A/G	C/T	T/T	C/C	A/G	A/A	A/A
M86	46	110,30	1,85	32,23	Class I	W	M	A/G	G/G	C/T	T/T	C/C	A/G	A/A	A/A
M89	50	74,50	1,67	26,71	Overweight	W	F	A/A	A/A	C/T	C/C	C/C	A/G	A/A	A/A
M90	21	63,70	1,67	22,84	Normal	W	F	G/G	G/G	C/T	C/C	C/C	A/A	A/A	A/A
M91	46	75,00	1,72	25,35	Overweight	W	F	A/G	A/G	C/T	T/T	C/T	A/G	A/A	A/A
M92	54	77,60	1,78	24,49	Normal	W	M	A/A	A/G	C/C	C/C	C/C	A/A	A/A	A/A
M94	23	78,10	1,63	29,40	Overweight	W	F	G/G	G/G	C/T	T/T	C/C	A/G	A/A	A/A

M95	23	77,20	1,72	26,10	Overweight	W	M	A/G	A/G	C/T	C/C	C/C	A/A	A/A	A/A
M96	36	97,50	1,75	31,84	Class I	W	M	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
M97	24	74,00	1,72	25,01	Overweight	W	F	A/G	G/G	T/T	C/C	C/C	A/G	A/G	A/A
M98	34	72,00	1,68	25,51	Overweight	W	F	G/G	G/G	C/C	C/C	C/C	A/A	A/A	A/A
M99	23	65,00	1,56	26,71	Overweight	W	F	A/A	A/G	C/T	C/T	C/T	A/G	A/A	A/A
M100	19	92,00	1,88	26,03	Overweight	W	M	A/A	G/G	T/T	T/T	C/C	G/G	A/A	A/A
M101	59	75,00	1,61	28,93	Overweight	W	M	A/G	A/G	C/T	C/T	C/C	G/G	A/G	A/A
M102	54	80,00	1,73	26,73	Overweight	W	F	A/A	G/G	T/T	C/C	C/T	A/G	A/G	A/A
M103	34	91,00	1,79	28,40	Overweight	W	F	A/G	G/G	C/T	C/T	C/T	A/G	A/A	A/A
M104	38	68,00	1,62	25,91	Overweight	W	F	G/G	A/A	C/T	T/T	C/C	A/A	A/A	A/A
M105	53	62,00	1,57	25,15	Overweight	W	F	A/G	A/G	T/T	C/T	C/T	A/G	A/A	A/A
M106	51	67,00	1,89	18,76	Normal	W	M	A/G	A/G	T/T	T/T	C/C	A/G	A/A	A/A
M107	28	86,00	1,83	25,68	Overweight	W	M	A/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
M108	26	56,60	1,86	16,36	Normal	W	F	A/A	G/G	C/T	T/T	C/C	A/A	A/A	A/A
M109	21	54,40	1,72	18,39	Normal	W	F	A/A	A/G	C/T	T/T	T/T	G/G	A/A	A/A
M110	24	52,30	1,73	17,47	Normal	W	F	A/A	A/G	T/T	C/T	C/C	A/A	A/A	A/A
M111	24	75,60	1,60	29,53	Overweight	W	F	A/G	A/G	T/T	C/C	C/C	A/A	A/A	A/A
M112	20	56,00	1,74	18,29	Normal	W	F	A/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
M113	26	66,90	1,68	23,70	Normal	W	F	G/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
M114	21	59,20	1,52	25,62	Overweight	W	F	A/A	G/G	C/T	T/T	C/C	A/G	A/A	A/A
M115	23	73,70	1,74	24,34	Normal	W	F	G/G	G/G	C/C	T/T	C/T	A/G	A/A	A/A
M116	40	87,20	1,72	29,48	Overweight	W	F	A/G	G/G	C/C	C/C	C/C	A/G	A/A	A/A
M117	26	53,50	1,60	20,90	Normal	W	F	A/G	A/G	C/T	T/T	C/T	A/A	A/A	A/A
M118	23	81,00	1,66	29,39	Overweight	W	F	G/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
B0001	61	80,50	1,71	27,53	Overweight	W	F	G/G	G/G	C/C	T/T	C/C	A/G	A/A	A/A
B0002	41	77,90	1,68	27,60	Overweight	W	F	G/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
B0003	50	60,00	1,79	18,73	Normal	W	F	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
B0004	46	61,00	1,60	23,83	Normal	W	F	A/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A

B0005	50	87,00	1,82	26,26	Overweight	W	M	A/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
B0007	31	113,90	1,70	39,41	Class II	W	F	A/G	A/A	C/C	C/T	C/C	A/G	A/A	A/A
B0008	51	112,30	1,70	38,86	Class II	W	F	A/G	G/G	T/T	T/T	C/T	A/G	A/A	A/A
B0009	26	80,00	1,75	26,12	Overweight	W	M	A/G	G/G	C/T	T/T	C/C	A/G	A/A	A/A
B0010	26	56,00	1,64	20,82	Normal	W	F	A/G	A/G	C/T	T/T	C/T	A/A	A/G	A/A
B0011	24	56,00	1,67	20,08	Normal	W	F	A/A	A/G	C/T	C/T	C/C	G/G	A/A	A/A
B0012	55	133,70	1,62	50,94	Class III	W	F	G/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
B0013	39	82,00	1,72	27,72	Overweight	W	F	A/G	G/G	C/T	T/T	C/C	G/G	A/A	A/A
B0014	27	52,00	1,58	20,83	Normal	W	F	A/A	A/G	C/T	T/T	C/C	A/G	A/G	A/A
B0015	47	74,20	1,70	25,67	Overweight	W	F	A/G	A/G	T/T	C/T	C/T	A/G	A/A	A/A
B0016	21	64,00	1,66	23,23	Normal	W	F	A/G	A/G	T/T	C/T	C/C	A/G	A/A	A/A
B0017	49	78,00	1,65	28,65	Overweight	W	F	G/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
B0018	45	84,60	1,79	26,40	Overweight	W	F	A/G	G/G	T/T	T/T	C/T	A/G	A/G	A/A
B0020	46	69,00	1,68	24,45	Normal	W	F	A/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
B0021	46	106,50	1,64	39,60	Class II	B	F	A/A	G/G	C/C	C/T	C/C	A/A	A/A	A/A
B0022	46	82,00	1,75	26,78	Overweight	W	M	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A
B0023	30	86,00	1,82	25,96	Overweight	W	M	A/G	G/G	T/T	T/T	C/C	A/G	A/G	A/A
B0024	30	172,50	1,82	52,08	Class III	W	F	A/A	G/G	T/T	T/T	C/C	A/G	A/A	A/A
B0025	30	184,50	1,89	51,65	Class III	W	M	A/G	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0026	28	153,30	1,78	48,38	Class III	W	M	G/G	G/G	T/T	C/C	C/C	A/A	A/A	A/A
B0027	29	114,00	1,79	35,58	Class II	W	F	A/G	A/G	C/T	C/T	C/C	A/G	A/G	A/A
B0028	33	115,80	1,67	41,52	Class III	W	F	A/G	A/G	T/T	C/C	C/C	G/G	A/A	A/A
B0030	49	103,70	1,72	35,05	Class II	W	F	A/A	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0031	39	115,30	1,75	37,65	Class II	W	M	A/G	G/G	C/C	C/T	C/C	A/G	A/G	A/A
B0032	44	118,00	1,79	36,83	Class II	W	M	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
B0033	28	55,00	1,65	20,20	Normal	W	F	G/G	G/G	C/T	C/T	C/C	G/G	A/G	A/A
B0035	47	59,00	1,68	20,90	Normal	W	F	A/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
B0036	50	115,00	1,80	35,49	Class II	W	M	A/G	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0037	24	81,00	1,75	26,45	Overweight	W	M	A/G	A/G	C/T	C/T	C/C	G/G	A/A	A/A

B0038	27	62,00	1,62	23,62	Normal	W	F	A/G	G/G	C/T	C/T	C/C	A/A	A/A	A/A
B0039	51	90,00	1,64	33,46	Class I	W	F	A/G	G/G	C/C	T/T	C/C	A/G	A/A	A/A
B0040	78	63,00	1,68	22,32	Normal	W	F	G/G	G/G	C/T	C/T	C/C	A/A	A/A	A/A
B0041	74	97,00	1,85	28,34	Overweight	W	M	A/G	A/G	C/T	C/C	C/C	A/A	A/A	A/A
B0042	56	48,70	1,63	18,33	Normal	W	F	A/G	G/G	C/T	C/T	C/T	A/A	A/G	A/A
B0043	68	84,00	1,79	26,22	Overweight	W	F	G/G	A/G	C/T	C/C	T/T	A/G	A/G	A/A
B0044	21	58,00	1,72	19,61	Normal	W	F	A/G	G/G	C/C	C/T	C/C	A/A	A/A	A/A
B0046	61	60,00	1,73	20,05	Normal	W	F	A/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
B0047	31	60,00	1,70	20,76	Normal	W	F	A/G	A/G	C/C	C/C	C/C	A/G	A/A	A/A
B0049	35	75,00	1,78	23,67	Normal	W	M	A/A	G/G	C/C	C/C	C/C	A/A	A/A	A/A
B0050	35	98,00	1,82	29,59	Overweight	W	M	A/A	G/G	T/T	C/T	C/C	G/G	A/A	A/A
B0051	27	50,00	1,51	21,93	Normal	C	F	A/G	G/G	T/T	C/T	C/T	A/A	A/A	A/A
B0052	44	110,00	1,80	33,95	Class I	I	M	A/G	G/G	T/T	C/C	C/C	A/A	A/G	A/A
B0053	47	82,00	1,65	30,12	Class I	B	M	G/G	A/G	C/C	C/C	C/C	A/A	A/A	A/A
B0054	56	95,50	1,76	30,83	Class I	W	M	A/G	G/G	C/T	T/T	C/C	A/A	A/A	A/A
B0055	50	65,50	1,74	21,63	Normal	W	M	A/G	G/G	C/C	C/T	C/C	G/G	A/A	A/A
B0056	38	81,00	1,72	27,38	Overweight	W	M	A/G	G/G	C/C	T/T	C/C	A/A	A/G	A/A
B0057	31	94,00	1,89	26,32	Overweight	W	M	A/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A
B0058	30	86,00	1,96	22,39	Normal	W	M	A/G	G/G	C/C	C/T	T/T	G/G	A/A	A/A
B0059	53	85,00	1,72	28,73	Overweight	W	M	G/G	G/G	T/T	C/T	C/C	G/G	A/G	A/A
B0060	50	99,00	1,76	31,96	Class I	W	M	A/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
B0061	50	119,70	1,75	39,09	Class II	C	M	A/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
B0062	30	75,80	1,64	28,18	Overweight	W	F	A/G	G/G	C/T	T/T	C/C	A/G	A/A	A/A
B0063	57	85,00	1,70	29,41	Overweight	W	F	A/G	G/G	C/C	T/T	C/C	A/G	A/A	A/A
B0064	47	72,50	1,67	26,00	Overweight	C	F	A/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A

B0065	49	73,00	1,64	27,14	Overweight	W	F	A/A	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0066	23	84,00	1,80	25,93	Overweight	W	M	A/G	G/G	C/T	C/C	C/C	A/G	A/A	A/A
B0067	27	58,00	1,67	20,80	Normal	I	F	A/G	A/A	T/T	C/C	T/T	A/A	A/A	A/G
B0068	30	86,00	1,70	29,76	Overweight	W	M	A/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
B0071	43	86,00	1,60	33,60	Class I	W	F	A/A	A/G	C/C	C/T	C/T	A/G	A/A	A/A
B0072	24	65,00	1,89	18,20	Normal	W	F	A/A	A/G	C/C	T/T	C/C	A/A	A/A	A/A
B0073	31	94,40	1,78	29,79	Overweight	W	M	A/G	A/G	C/T	C/T	C/C	A/A	A/A	A/A
B0074	27	65,00	1,68	23,03	Normal	W	F	G/G	G/G	C/T	C/C	C/C	A/G	A/A	A/A
B0075	31	118,00	1,92	32,01	Class I	W	M	A/A	G/G	T/T	C/T	C/C	A/G	A/A	A/A
B0076	26	90,00	1,65	33,06	Class I	W	F	A/G	G/G	C/C	T/T	C/T	A/A	A/A	A/A
B0077	38	73,00	1,72	24,68	Normal	W	F	A/A	A/G	T/T	C/C	C/T	A/G	A/A	A/A
B0080	30	80,00	1,70	27,68	Overweight	W	F	A/A	G/G	C/T	T/T	C/C	A/G	A/A	A/A
B0081	44	87,80	1,64	32,64	Class I	C	F	G/G	A/G	C/C	C/C	C/T	A/A	A/A	A/A
B0082	34	125,00	1,96	32,54	Class I	W	M	G/G	G/G	C/C	C/C	T/T	A/G	A/A	A/A
B0083	55	75,00	1,74	24,77	Normal	W	F	A/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
B0084	31	69,60	1,70	24,08	Normal	W	F	A/G	G/G	C/C	T/T	C/T	A/G	A/A	A/A
B0085	37	75,00	1,73	25,06	Overweight	W	M	A/G	A/G	C/C	C/C	T/T	G/G	A/A	A/A
B0086	24	54,00	1,69	18,91	Normal	W	F	A/A	G/G	C/T	C/T	C/C	G/G	A/A	A/A
B0087	51	110,00	1,65	40,40	Class III	W	F	A/G	G/G	C/C	C/C	C/C	A/G	A/A	A/A
B0088	55	88,50	1,65	32,51	Class I	W	F	A/G	G/G	T/T	C/T	C/T	A/G	A/A	A/A
B0089	63	113,00	1,68	40,04	Class III	W	F	A/A	G/G	C/C	C/T	C/C	A/G	A/A	A/A
B0090	57	165,00	1,78	52,08	Class III	W	M	A/A	A/G	T/T	C/T	T/T	A/G	A/A	A/A
B0091	69	85,00	1,80	26,23	Overweight	W	F	A/A	A/G	T/T	C/C	C/C	A/G	A/A	A/A
B0093	60	106,00	1,72	35,83	Class II	W	M	A/A	A/G	T/T	C/T	C/C	A/G	A/A	A/A
B0094	59	73,00	1,73	24,39	Normal	W	F	A/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
B0096	18	63,00	1,42	31,24	Class I	W	M	A/A	G/G	C/T	C/T	C/C	G/G	A/A	A/A
B0097	68	72,00	1,78	22,72	Normal	W	M	A/A	G/G	T/T	T/T	T/T	A/G	A/A	A/A
B0098	24	98,00	1,68	34,72	Class I	W	F	A/A	G/G	C/T	C/T	C/T	A/G	A/A	A/A
B0099	24	100,75	1,82	30,42	Class I	W	M	A/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A

B0100	30	105,00	1,78	33,14	Class I	W	M	G/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
B0101	42	115,00	1,76	37,13	Class II	W	M	A/A	G/G	T/T	C/T	C/C	A/G	A/G	A/A
B0102	52	102,40	1,60	40,00	Class III	W	F	A/G	G/G	C/T	T/T	C/C	A/A	A/A	A/A
B0105	25	60,00	1,60	23,44	Normal	B	M	G/G	A/G	T/T	T/T	C/C	G/G	A/A	A/A
B0106	28	53,00	1,58	21,23	Normal	W	F	G/G	G/G	C/T	C/T	C/C	G/G	A/A	A/A
B0107	31	53,00	1,56	21,78	Normal	W	F	G/G	A/G	C/T	C/T	C/C	A/A	A/A	A/A
B0108	49	51,00	1,53	21,79	Normal	W	F	A/A	A/G	T/T	T/T	C/T	A/A	A/A	A/A
B0109	27	108,10	1,67	38,76	Class II	C	F	G/G	G/G	C/T	C/C	C/C	G/G	A/A	A/A
B0110	23	52,00	1,69	18,21	Normal	W	F	G/G	A/G	C/T	C/T	C/C	A/G	A/G	A/A
B0111	23	59,90	1,65	22,00	Normal	W	F	A/G	A/G	T/T	C/T	C/C	G/G	A/A	A/A
B0112	40	54,00	1,67	19,36	Normal	W	F	G/G	G/G	C/T	C/C	C/C	A/G	A/A	A/A
B0113	44	60,00	1,67	21,51	Normal	W	F	A/A	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0114	56	56,00	1,60	21,88	Normal	W	F	A/G	G/G	C/T	C/C	C/C	A/A	A/A	A/A
B0115	45	53,00	1,65	19,47	Normal	B	F	A/A	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0116	39	54,00	1,63	20,32	Normal	W	F	A/A	A/G	C/T	C/T	T/T	A/A	A/G	A/A
B0117	49	55,00	1,65	20,20	Normal	W	F	A/G	G/G	C/T	C/C	C/C	A/G	A/A	A/A
B0121	46	65,00	1,76	20,98	Normal	W	F	A/G	G/G	C/T	T/T	C/T	A/G	G/G	A/A
B0122	44	60,00	1,65	22,04	Normal	W	F	A/G	G/G	C/T	C/T	C/T	A/G	A/A	A/A
B0123	30	58,00	1,67	20,80	Normal	W	F	A/A	G/G	T/T	C/T	C/C	G/G	A/A	A/A
B0124	27	70,00	1,73	23,39	Normal	W	F	A/G	G/G	C/T	T/T	C/C	A/G	A/A	A/A
B0125	57	67,00	1,73	22,39	Normal	W	F	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A
B0126	31	72,20	1,72	24,41	Normal	W	F	A/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A
B0127	55	85,00	1,59	33,62	Class I	W	F	G/G	A/G	T/T	T/T	C/C	A/A	A/A	A/A
B0128	27	63,00	1,76	20,34	Normal	W	F	G/G	G/G	C/T	T/T	C/C	G/G	A/A	A/A
B0129	27	62,00	1,70	21,45	Normal	W	F	A/G	A/G	C/T	C/T	C/T	A/A	A/G	A/A
B0130	30	56,80	1,64	21,12	Normal	W	F	G/G	A/G	T/T	C/C	C/C	A/A	A/A	A/A
B0131	32	57,00	1,65	20,94	Normal	W	F	A/A	G/G	C/C	C/T	C/T	G/G	A/A	A/A
B0133	26	89,00	1,55	37,04	Class II	B	F	A/A	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0134	45	143,00	1,78	45,13	Class III	B	F	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A

Appendix B5: Samples run in duplication

No	COMT	DRD2	DRD4	GABRA6	HTR2C	LEPR	OPRM1	SLC6A4
	rs4680	rs1800497	rs1800955	rs3219151	rs3813929	rs1137101	rs1799971	rs25531
B0001	G/G	G/G	C/C	T/T	C/C	A/G	A/A	A/A
Repeat	G/G	G/G	C/C	T/T	C/C	A/G	A/A	A/A
B0002	G/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
Repeat	G/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
B0003	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
Repeat	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
B0004	A/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A
Repeat	A/G	G/G	C/C	C/T	C/C	A/G	A/A	A/A
B0005	A/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
Repeat	A/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
B0007	A/G	A/A	C/C	C/T	C/C	A/G	A/A	A/A
Repeat	A/G	A/A	C/C	C/T	C/C	A/G	A/A	A/A
B0008	A/G	G/G	T/T	T/T	C/T	A/G	A/A	A/A
Repeat	A/G	G/G	T/T	T/T	C/T	A/G	A/A	A/A
B0009	A/G	G/G	C/T	T/T	C/C	A/G	A/A	A/A
Repeat	A/G	G/G	C/T	T/T	C/C	A/G	A/A	A/A
B0010	A/G	A/G	C/T	T/T	C/T	A/A	A/G	A/A
Repeat	A/G	A/G	C/T	T/T	C/T	A/A	A/G	A/A
B0011	A/A	A/G	C/T	C/T	C/C	G/G	A/A	A/A
Repeat	A/A	A/G	C/T	C/T	C/C	G/G	A/A	A/A
B0012	G/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
Repeat	G/G	G/G	T/T	C/T	C/C	A/G	A/A	A/A
B0013	A/G	G/G	C/T	T/T	C/C	G/G	A/A	A/A
Repeat	A/G	G/G	C/T	T/T	C/C	G/G	A/A	A/A
B0014	A/A	A/G	C/T	T/T	C/C	A/G	A/G	A/A

Repeat	A/A	A/G	C/T	T/T	C/C	A/G	A/G	A/A
B0015	A/G	A/G	T/T	C/T	C/T	A/G	A/A	A/A
Repeat	A/G	A/G	T/T	C/T	C/T	A/G	A/A	A/A
B0016	A/G	A/G	T/T	C/T	C/C	A/G	A/A	A/A
Repeat	A/G	A/G	T/T	C/T	C/C	A/G	A/A	A/A
B0017	G/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
Repeat	G/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
B0018	A/G	G/G	T/T	T/T	C/T	A/G	A/G	A/A
Repeat	A/G	G/G	T/T	T/T	C/T	A/G	A/G	A/A
B0020	A/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
Repeat	A/G	G/G	C/T	C/T	C/C	A/G	A/A	A/A
B0021	A/A	G/G	C/C	C/T	C/C	A/A	A/A	A/A
Repeat	A/A	G/G	C/C	C/T	C/C	A/A	A/A	A/A
B0022	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A
Repeat	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A
B0023	A/G	G/G	T/T	T/T	C/C	A/G	A/G	A/A
Repeat	A/G	G/G	T/T	T/T	C/C	A/G	A/G	A/A
B0032	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
Repeat	A/G	A/G	T/T	C/C	C/C	A/G	A/A	A/A
B0033	G/G	G/G	C/T	C/T	C/C	G/G	A/G	A/A
Repeat	G/G	G/G	C/T	C/T	C/C	G/G	A/G	A/A
B0042	A/G	G/G	C/T	C/T	C/T	A/A	A/G	A/A
Repeat	A/G	G/G	C/T	C/T	C/T	A/A	A/G	A/A
B0043	G/G	A/G	C/T	C/C	T/T	A/G	A/G	A/A
Repeat	G/G	A/G	C/T	C/C	T/T	A/G	A/G	A/A
B0044	A/G	G/G	C/C	C/T	C/C	A/A	A/A	A/A
Repeat	A/G	G/G	C/C	C/T	C/C	A/A	A/A	A/A
B0133	A/A	A/G	C/T	C/T	C/C	A/G	A/A	A/A
Repeat	A/A	A/G	C/T	C/T	C/C	A/G	A/A	A/A
B0134	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A

Repeat	A/G	A/G	C/T	C/C	C/C	A/G	A/A	A/A
M108	A/A	G/G	C/T	T/T	C/C	A/A	A/A	A/A
Repeat	A/A	G/G	C/T	T/T	C/C	A/A	A/A	A/A
M109	A/A	A/G	C/T	T/T	T/T	G/G	A/A	A/A
Repeat	A/A	A/G	C/T	T/T	T/T	G/G	A/A	A/A
M110	A/A	A/G	T/T	C/T	C/C	A/A	A/A	A/A
Repeat	A/A	A/G	T/T	C/T	C/C	A/A	A/A	A/A
M111	A/G	A/G	T/T	C/C	C/C	A/A	A/A	A/A
Repeat	A/G	A/G	T/T	C/C	C/C	A/A	A/A	A/A
M112	A/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
Repeat	A/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
M113	G/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
Repeat	G/G	G/G	T/T	C/T	C/C	A/A	A/A	A/A
M114	A/A	G/G	C/T	T/T	C/C	A/G	A/A	A/A
Repeat	A/A	G/G	C/T	T/T	C/C	A/G	A/A	A/A
M115	G/G	G/G	C/C	T/T	C/T	A/G	A/A	A/A
Repeat	G/G	G/G	C/C	T/T	C/T	A/G	A/A	A/A
M116	A/G	G/G	C/C	C/C	C/C	A/G	A/A	A/A
Repeat	A/G	G/G	C/C	C/C	C/C	A/G	A/A	A/A
M117	A/G	A/G	C/T	T/T	C/T	A/A	A/A	A/A
Repeat	A/G	A/G	C/T	T/T	C/T	A/A	A/A	A/A
M118	G/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A
Repeat	G/G	G/G	T/T	T/T	C/C	A/G	A/A	A/A

Appendix B6: Statistical Data

. for var site bmi_class bmi_pool race gender comp_n - slc6a4_n : tab X

```
->  tab site

       site |      Freq.     Percent        Cum.
------------+-----------------------------------
     Bianca |        116       52.02       52.02
     Marais |        107       47.98      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab bmi_class

  bmi_class |      Freq.     Percent        Cum.
------------+-----------------------------------
     Normal |        107       47.98       47.98
    Over_Wt |         65       29.15       77.13
    Class I |         27       12.11       89.24
   Class II |         13        5.83       95.07
  Class III |         11        4.93      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab bmi_pool

   bmi_pool |      Freq.     Percent        Cum.
------------+-----------------------------------
     Normal |        107       47.98       47.98
    Over_Wt |         65       29.15       77.13
   Cl_I-III |         51       22.87      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab race

       race |      Freq.     Percent        Cum.
------------+-----------------------------------
      White |        184       82.51       82.51
      Black |         31       13.90       96.41
   Coloured |          5        2.24       98.65
     Indian |          3        1.35      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab gender

     gender |      Freq.     Percent        Cum.
------------+-----------------------------------
       Male |         66       29.60       29.60
     Female |        157       70.40      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab comp_n

     comp_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        G/G |         48       21.52       21.52
        A/G |        118       52.91       74.44
        A/A |         57       25.56      100.00
------------+-----------------------------------
      Total |        223      100.00
```

```
->  tab drd2_n

     drd2_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        G/G |        128       57.40       57.40
        A/G |         82       36.77       94.17
        A/A |         13        5.83      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab drd4_n

     drd4_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        T/T |         71       31.84       31.84
        C/T |        108       48.43       80.27
        C/C |         44       19.73      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab gabra6_n

   gabra6_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        C/C |         50       22.42       22.42
        C/T |        107       47.98       70.40
        T/T |         66       29.60      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab htr2c_n

    htr2c_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        T/T |         11        4.93        4.93
        C/T |         38       17.04       21.97
        C/C |        174       78.03      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab lepr_n

     lepr_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        G/G |         42       18.83       18.83
        A/G |        118       52.91       71.75
        A/A |         63       28.25      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab oprm1_n

    oprm1_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        A/A |        186       83.41       83.41
        A/G |         33       14.80       98.21
        G/G |          4        1.79      100.00
------------+-----------------------------------
      Total |        223      100.00

->  tab slc6a4_n

   slc6a4_n |      Freq.     Percent        Cum.
------------+-----------------------------------
        A/G |          3        1.35        1.35
        A/A |        220       98.65      100.00
------------+-----------------------------------
      Total |        223      100.00
```

. for var comp_n - slc6a4_n : tab X bmi_class, row col exact nolog \ tab X bmi_pool, row col exact nolog

```
->  tab comp_n bmi_class, row col exact nolog

+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
| column percentage |
+-------------------+
```

comp_n	bmi_class: Normal	Over_Wt	Class I	Class II	Class III	Total
G/G	23	14	8	1	2	48
	47.92	29.17	16.67	2.08	4.17	100.00
	21.50	21.54	29.63	7.69	18.18	21.52
A/G	54	37	15	7	5	118
	45.76	31.36	12.71	5.93	4.24	100.00
	50.47	56.92	55.56	53.85	45.45	52.91
A/A	30	14	4	5	4	57
	52.63	24.56	7.02	8.77	7.02	100.00
	28.04	21.54	14.81	38.46	36.36	25.56
Total	107	65	27	13	11	223
	47.98	29.15	12.11	5.83	4.93	100.00
	100.00	100.00	100.00	100.00	100.00	100.00

Fisher's Exact = 0.680

```
->  tab comp_n bmi_pool, row col exact nolog

+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
| column percentage |
+-------------------+
```

comp_n	bmi_pool: Normal	Over_Wt	Cl_I-III	Total
G/G	23	14	11	48
	47.92	29.17	22.92	100.00
	21.50	21.54	21.57	21.52
A/G	54	37	27	118
	45.76	31.36	22.88	100.00
	50.47	56.92	52.94	52.91
A/A	30	14	13	57
	52.63	24.56	22.81	100.00
	28.04	21.54	25.49	25.56
Total	107	65	51	223
	47.98	29.15	22.87	100.00
	100.00	100.00	100.00	100.00

Fisher's Exact = 0.915

```
->  tab drd2_n bmi_class, row col exact nolog
```

```
+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
| column percentage |
+-------------------+
```

drd2_n	bmi_class: Normal	Over_Wt	Class I	Class II	Class III	Total
G/G	60	38	17	6	7	128
	46.88	29.69	13.28	4.69	5.47	100.00
	56.07	58.46	62.96	46.15	63.64	57.40
A/G	41	22	9	6	4	82
	50.00	26.83	10.98	7.32	4.88	100.00
	38.32	33.85	33.33	46.15	36.36	36.77
A/A	6	5	1	1	0	13
	46.15	38.46	7.69	7.69	0.00	100.00
	5.61	7.69	3.70	7.69	0.00	5.83
Total	107	65	27	13	11	223
	47.98	29.15	12.11	5.83	4.93	100.00
	100.00	100.00	100.00	100.00	100.00	100.00

Fisher's Exact = 0.975

```
->  tab drd2_n bmi_pool, row col exact nolog
```

```
+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
| column percentage |
+-------------------+
```

drd2_n	bmi_pool: Normal	Over_Wt	Cl_I-III	Total
G/G	60	38	30	128
	46.88	29.69	23.44	100.00
	56.07	58.46	58.82	57.40
A/G	41	22	19	82
	50.00	26.83	23.17	100.00
	38.32	33.85	37.25	36.77
A/A	6	5	2	13
	46.15	38.46	15.38	100.00
	5.61	7.69	3.92	5.83
Total	107	65	51	223
	47.98	29.15	22.87	100.00
	100.00	100.00	100.00	100.00

Fisher's Exact = 0.923

```
->  tab drd4_n bmi_class, row col exact nolog

+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
| column percentage |
```

```
+-------------------+

           |                        bmi_class
    drd4_n |    Normal    Over_Wt    Class I   Class II  Class III |     Total
-----------+---------------------------------------------------------+----------
       T/T |        32         21          9          4          5 |        71
           |     45.07      29.58      12.68       5.63       7.04 |    100.00
           |     29.91      32.31      33.33      30.77      45.45 |     31.84
-----------+---------------------------------------------------------+----------
       C/T |        54         34         10          6          4 |       108
           |     50.00      31.48       9.26       5.56       3.70 |    100.00
           |     50.47      52.31      37.04      46.15      36.36 |     48.43
-----------+---------------------------------------------------------+----------
       C/C |        21         10          8          3          2 |        44
           |     47.73      22.73      18.18       6.82       4.55 |    100.00
           |     19.63      15.38      29.63      23.08      18.18 |     19.73
-----------+---------------------------------------------------------+----------
     Total |       107         65         27         13         11 |       223
           |     47.98      29.15      12.11       5.83       4.93 |    100.00
           |    100.00     100.00     100.00     100.00     100.00 |    100.00

           Fisher's Exact  =                 0.822

->  tab drd4_n bmi_pool, row col exact nolog

+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
| column percentage |
+-------------------+

           |             bmi_pool
    drd4_n |    Normal    Over_Wt   Cl_I-III |     Total
-----------+---------------------------------+----------
       T/T |        32         21         18 |        71
           |     45.07      29.58      25.35 |    100.00
           |     29.91      32.31      35.29 |     31.84
-----------+---------------------------------+----------
       C/T |        54         34         20 |       108
           |     50.00      31.48      18.52 |    100.00
           |     50.47      52.31      39.22 |     48.43
-----------+---------------------------------+----------
       C/C |        21         10         13 |        44
           |     47.73      22.73      29.55 |    100.00
           |     19.63      15.38      25.49 |     19.73
-----------+---------------------------------+----------
     Total |       107         65         51 |       223
           |     47.98      29.15      22.87 |    100.00
           |    100.00     100.00     100.00 |    100.00

           Fisher's Exact  =                 0.554

->  tab gabra6_n bmi_class, row col exact nolog

+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
| column percentage |



+-------------------+
```

gabra6_n	bmi_class					Total
	Normal	Over_Wt	Class I	Class II	Class III	
C/C	21	15	8	2	4	50
	42.00	30.00	16.00	4.00	8.00	100.00
	19.63	23.08	29.63	15.38	36.36	22.42
C/T	55	27	10	10	5	107
	51.40	25.23	9.35	9.35	4.67	100.00
	51.40	41.54	37.04	76.92	45.45	47.98
T/T	31	23	9	1	2	66
	46.97	34.85	13.64	1.52	3.03	100.00
	28.97	35.38	33.33	7.69	18.18	29.60
Total	107	65	27	13	11	223
	47.98	29.15	12.11	5.83	4.93	100.00
	100.00	100.00	100.00	100.00	100.00	100.00

Fisher's Exact = 0.309

-> tab gabra6_n bmi_pool, row col exact nolog

Key
frequency
row percentage
column percentage

gabra6_n	bmi_pool			Total
	Normal	Over_Wt	Cl_I-III	
C/C	21	15	14	50
	42.00	30.00	28.00	100.00
	19.63	23.08	27.45	22.42
C/T	55	27	25	107
	51.40	25.23	23.36	100.00
	51.40	41.54	49.02	47.98
T/T	31	23	12	66
	46.97	34.85	18.18	100.00
	28.97	35.38	23.53	29.60
Total	107	65	51	223
	47.98	29.15	22.87	100.00
	100.00	100.00	100.00	100.00

Fisher's Exact = 0.528

-> tab htr2c_n bmi_class, row col exact nolog

Key
frequency
row percentage
column percentage

htr2c_n	bmi_class					
	Normal	Over_Wt	Class I	Class II	Class III	Total
T/T	6	3	1	0	1	11
	54.55	27.27	9.09	0.00	9.09	100.00
	5.61	4.62	3.70	0.00	9.09	4.93
C/T	21	11	5	1	0	38
	55.26	28.95	13.16	2.63	0.00	100.00
	19.63	16.92	18.52	7.69	0.00	17.04
C/C	80	51	21	12	10	174
	45.98	29.31	12.07	6.90	5.75	100.00
	74.77	78.46	77.78	92.31	90.91	78.03
Total	107	65	27	13	11	223
	47.98	29.15	12.11	5.83	4.93	100.00
	100.00	100.00	100.00	100.00	100.00	100.00

Fisher's Exact = 0.804

```
->  tab htr2c_n bmi_pool, row col exact nolog
```

Key
frequency
row percentage
column percentage

htr2c_n	bmi_pool			
	Normal	Over_Wt	Cl_I-III	Total
T/T	6	3	2	11
	54.55	27.27	18.18	100.00
	5.61	4.62	3.92	4.93
C/T	21	11	6	38
	55.26	28.95	15.79	100.00
	19.63	16.92	11.76	17.04
C/C	80	51	43	174
	45.98	29.31	24.71	100.00
	74.77	78.46	84.31	78.03
Total	107	65	51	223
	47.98	29.15	22.87	100.00
	100.00	100.00	100.00	100.00

Fisher's Exact = 0.787

```
->  tab lepr_n bmi_class, row col exact nolog
```

Key
frequency
row percentage
column percentage

lepr_n	bmi_class					
	Normal	Over_Wt	Class I	Class II	Class III	Total
G/G	27	10	3	1	1	42
	64.29	23.81	7.14	2.38	2.38	100.00
	25.23	15.38	11.11	7.69	9.09	18.83
A/G	42	41	16	11	8	118
	35.59	34.75	13.56	9.32	6.78	100.00
	39.25	63.08	59.26	84.62	72.73	52.91
A/A	38	14	8	1	2	63
	60.32	22.22	12.70	1.59	3.17	100.00
	35.51	21.54	29.63	7.69	18.18	28.25
Total	107	65	27	13	11	223
	47.98	29.15	12.11	5.83	4.93	100.00
	100.00	100.00	100.00	100.00	100.00	100.00

Fisher's Exact = 0.018

```
->  tab lepr_n bmi_pool, row col exact nolog
```

Key
frequency
row percentage
column percentage

lepr_n	bmi_pool			
	Normal	Over_Wt	Cl_I-III	Total
G/G	27	10	5	42
	64.29	23.81	11.90	100.00
	25.23	15.38	9.80	18.83
A/G	42	41	35	118
	35.59	34.75	29.66	100.00
	39.25	63.08	68.63	52.91
A/A	38	14	11	63
	60.32	22.22	17.46	100.00
	35.51	21.54	21.57	28.25
Total	107	65	51	223
	47.98	29.15	22.87	100.00
	100.00	100.00	100.00	100.00

Fisher's Exact = 0.003

```
->  tab oprm1_n bmi_class, row col exact nolog
```

Key
frequency
row percentage
column percentage

oprm1_n	bmi_class				
	Normal	Over_Wt	Class I	Class II	Class III	Total
A/A	88	51	26	10	11	186
	47.31	27.42	13.98	5.38	5.91	100.00
	82.24	78.46	96.30	76.92	100.00	83.41
A/G	15	14	1	3	0	33
	45.45	42.42	3.03	9.09	0.00	100.00
	14.02	21.54	3.70	23.08	0.00	14.80
G/G	4	0	0	0	0	4
	100.00	0.00	0.00	0.00	0.00	100.00
	3.74	0.00	0.00	0.00	0.00	1.79
Total	107	65	27	13	11	223
	47.98	29.15	12.11	5.83	4.93	100.00
	100.00	100.00	100.00	100.00	100.00	100.00

Fisher's Exact = 0.172

```
->  tab oprm1_n bmi_pool, row col exact nolog
```

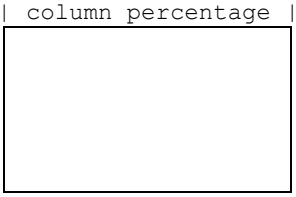
Key
frequency
row percentage
column percentage

oprm1_n	bmi_pool			
	Normal	Over_Wt	Cl_I-III	Total
A/A	88	51	47	186
	47.31	27.42	25.27	100.00
	82.24	78.46	92.16	83.41
A/G	15	14	4	33
	45.45	42.42	12.12	100.00
	14.02	21.54	7.84	14.80
G/G	4	0	0	4
	100.00	0.00	0.00	100.00
	3.74	0.00	0.00	1.79
Total	107	65	51	223
	47.98	29.15	22.87	100.00
	100.00	100.00	100.00	100.00

Fisher's Exact = 0.090

```
->  tab slc6a4_n bmi_class, row col exact nolog
```

Key
frequency
row percentage
column percentage



slc6a4_n	bmi_class					
	Normal	Over_Wt	Class I	Class II	Class III	Total
A/G	2	0	1	0	0	3
	66.67	0.00	33.33	0.00	0.00	100.00
	1.87	0.00	3.70	0.00	0.00	1.35
A/A	105	65	26	13	11	220
	47.73	29.55	11.82	5.91	5.00	100.00
	98.13	100.00	96.30	100.00	100.00	98.65
Total	107	65	27	13	11	223
	47.98	29.15	12.11	5.83	4.93	100.00
	100.00	100.00	100.00	100.00	100.00	100.00

Fisher's Exact = 0.464

-> tab slc6a4_n bmi_pool, row col exact nolog

Key
frequency
row percentage
column percentage

slc6a4_n	bmi_pool			
	Normal	Over_Wt	Cl_I-III	Total
A/G	2	0	1	3
	66.67	0.00	33.33	100.00
	1.87	0.00	1.96	1.35
A/A	105	65	50	220
	47.73	29.55	22.73	100.00
	98.13	100.00	98.04	98.65
Total	107	65	51	223
	47.98	29.15	22.87	100.00
	100.00	100.00	100.00	100.00

Fisher's Exact = 0.603

. for var comp_n - slc6a4_n: mlogit bmi_pool i.X if race < 2 , nolog rrr \ mlogit bmi_pool i.X i.race if race < 2 , nolog rrr

```
->  mlogit bmi_pool i.comp_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =       1.34
                                                Prob > chi2       =     0.8550
Log likelihood = -223.87636                     Pseudo R2         =     0.0030

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      comp_n |
        A/G  |   1.159664   .4671768     0.37   0.713     .5265274    2.554131
        A/A  |   .7666667   .3593277    -0.57   0.571     .3059556    1.921121
             |
       _cons |   .6086957   .2063351    -1.46   0.143     .3132281    1.182877
-------------+----------------------------------------------------------------
Cl_I_III     |
      comp_n |
        A/G  |   1.252723   .5797785     0.49   0.626     .5057161    3.103155
        A/A  |   1.107407    .569907     0.20   0.843     .4038798    3.036426
             |
       _cons |   .3913043   .1538525    -2.39   0.017     .1810675    .8456465
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.

->  mlogit bmi_pool i.comp_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(6)        =       5.49
                                                Prob > chi2       =     0.4825
Log likelihood =  -221.7994                     Pseudo R2         =     0.0122

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      comp_n |
        A/G  |   .9668534   .4036633    -0.08   0.936     .4265621    2.191488
        A/A  |   .6253156   .3025501    -0.97   0.332     .2422464     1.61414
             |
        race |
      Black  |   .3551674   .1934457    -1.90   0.057     .1221286    1.032878
       _cons |   .8024451   .2947988    -0.60   0.549     .3905752    1.648641
-------------+----------------------------------------------------------------
Cl_I_III     |
      comp_n |
        A/G  |   1.194703   .5680644     0.37   0.708     .4704631    3.033851
        A/A  |   1.049626   .5554211     0.09   0.927     .3720593    2.961128
             |
        race |
      Black  |   .8078728   .4023104    -0.43   0.668     .3044075    2.144029
       _cons |   .4216367   .1805585    -2.02   0.044     .1821488    .9760016
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
```

```
->  mlogit bmi_pool i.drd2_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =       1.50
                                                Prob > chi2       =     0.8266
Log likelihood =  -223.7949                     Pseudo R2         =     0.0033

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      drd2_n |
        A/G  |   .8556361   .2886691    -0.46   0.644     .4416886    1.657533
        A/A  |   1.993243    1.40092     0.98   0.326     .5026973    7.903401
             |
       _cons |   .6271186   .1315099    -2.23   0.026     .4157656    .9459124
-------------+----------------------------------------------------------------
Cl_I_III     |
      drd2_n |
        A/G  |   .9593496   .3510905    -0.11   0.910     .4682363    1.965571
        A/A  |   1.092593   .9796762     0.10   0.921      .188463    6.334181
             |
       _cons |   .4576271   .1063293    -3.36   0.001     .2902256    .7215855
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.

->  mlogit bmi_pool i.drd2_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(6)        =       5.96
                                                Prob > chi2       =     0.4276
Log likelihood = -221.56457                     Pseudo R2         =     0.0133

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      drd2_n |
        A/G  |   .9818917   .3395937    -0.05   0.958     .4985055    1.934003
        A/A  |   2.757189   2.039844     1.37   0.170     .6467316    11.75463
             |
        race |
      Black  |   .3359623   .1863274    -1.97   0.049     .1132941    .9962624
       _cons |   .6699165   .1423683    -1.89   0.059     .4416989     1.01605
-------------+----------------------------------------------------------------
Cl_I_III     |
      drd2_n |
        A/G  |   1.002007   .3757522     0.01   0.996     .4804751    2.089638
        A/A  |   1.202945   1.101525     0.20   0.840     .1998983     7.23907
             |
        race |
      Black  |   .7703315   .3851969    -0.52   0.602     .2890963     2.05264
       _cons |   .4679672   .1104376    -3.22   0.001     .2946705    .7431803
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
```

```
->  mlogit bmi_pool i.drd4_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =       3.97
                                                Prob > chi2       =     0.4105
Log likelihood = -222.56167                     Pseudo R2         =     0.0088

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      drd4_n |
        C/T  |    .916442   .3294505    -0.24   0.808     .4530099    1.853968
        C/C  |   .6122449   .2997391    -1.00   0.316     .2345304    1.598274
             |
       _cons |         .7   .1991649    -1.25   0.210     .4007865    1.222596
-------------+----------------------------------------------------------------
Cl_I_III     |
      drd4_n |
        C/T  |   .5993341   .2446137    -1.25   0.210     .2693131    1.333769
        C/C  |   1.008403   .4763137     0.02   0.986     .3995526     2.54504
             |
       _cons |   .5666667    .172025    -1.87   0.061     .3125543    1.027377
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.

->  mlogit bmi_pool i.drd4_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(6)        =       8.08
                                                Prob > chi2       =     0.2320
Log likelihood =  -220.5028                     Pseudo R2         =     0.0180

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      drd4_n |
        C/T  |   .8409619   .3076354    -0.47   0.636     .4105774    1.722493
        C/C  |   .5613107   .2780859    -1.17   0.244     .2125685    1.482203
             |
        race |
      Black  |   .3620838   .1937836    -1.90   0.058     .1268395    1.033626
       _cons |   .8431012    .253274    -0.57   0.570     .4679208    1.519103
-------------+----------------------------------------------------------------
Cl_I_III     |
      drd4_n |
        C/T  |   .5801017   .2389499    -1.32   0.186     .2587529    1.300538
        C/C  |   .9756998   .4641233    -0.05   0.959     .3840737    2.478665
             |
        race |
      Black  |   .7361219   .3600113    -0.63   0.531      .282265     1.91974
       _cons |   .6094571   .1974147    -1.53   0.126     .3230151    1.149909
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
```

```
->  mlogit bmi_pool i.gabra6_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =       2.49
                                                Prob > chi2       =     0.6469
Log likelihood = -223.30143                     Pseudo R2         =     0.0055

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
    gabra6_n |
        C/T  |   .6540881   .2728387    -1.02   0.309     .2887858    1.481483
        T/T  |   .9892473   .4339211    -0.02   0.980     .4187317    2.337082
             |
       _cons |        .75   .2561738    -0.84   0.400     .3839906     1.46488
-------------+----------------------------------------------------------------
Cl_I_III     |
    gabra6_n |
        C/T  |   .8233276   .3695298    -0.43   0.665     .3416114    1.984326
        T/T  |   .7038123   .3564519    -0.69   0.488     .2608305    1.899133
             |
       _cons |        .55   .2064582    -1.59   0.111     .2635359    1.147851
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.

->  mlogit bmi_pool i.gabra6_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(6)        =       6.60
                                                Prob > chi2       =     0.3595
Log likelihood = -221.24529                     Pseudo R2         =     0.0147

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
    gabra6_n |
        C/T  |   .6221277   .2625402    -1.12   0.261     .2720629    1.422623
        T/T  |   .9694618   .4300133    -0.07   0.944     .4064179    2.312536
             |
        race |
      Black  |   .3643126   .1942842    -1.89   0.058     .1280965    1.036123
       _cons |   .8759233   .3088586    -0.38   0.707     .4388585    1.748266
-------------+----------------------------------------------------------------
Cl_I_III     |
    gabra6_n |
        C/T  |   .8101081   .3647537    -0.47   0.640     .3351868    1.957939
        T/T  |   .6992791   .3545708    -0.71   0.481     .2588494    1.889095
             |
        race |
      Black  |   .7761155    .375504    -0.52   0.600     .3006726    2.003359
       _cons |   .5793325   .2246343    -1.41   0.159     .2709425    1.238736
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
```

```
->  mlogit bmi_pool i.htr2c_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =       1.94
                                                Prob > chi2       =     0.7467
Log likelihood =  -223.5747                     Pseudo R2         =     0.0043

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
     htr2c_n |
        C/T  |   .9166667   .7526958    -0.11   0.916     .1833456    4.583026
        C/C  |   1.054852   .7935908     0.07   0.943     .2414417    4.608622
             |
       _cons |         .6    .438178    -0.70   0.484     .1433909     2.51062
-------------+----------------------------------------------------------------
Cl_I_III     |
     htr2c_n |
        C/T  |       .625   .6091746    -0.48   0.630     .0925191    4.222101
        C/C  |   1.265823   1.087178     0.27   0.784     .2351263    6.814667
             |
       _cons |         .4    .334664    -1.10   0.273     .0776057    2.061704
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.

->  mlogit bmi_pool i.htr2c_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(6)        =       6.13
                                                Prob > chi2       =     0.4088
Log likelihood = -221.47992                     Pseudo R2         =     0.0137

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
     htr2c_n |
        C/T  |   .9385937   .7710654    -0.08   0.939     .1875884     4.69623
        C/C  |   1.239051   .9369126     0.28   0.777     .2814815    5.454164
             |
        race |
      Black  |   .3572929   .1920682    -1.91   0.056     .1245799    1.024709
       _cons |         .6    .438178    -0.70   0.484     .1433909     2.51062
-------------+----------------------------------------------------------------
Cl_I_III     |
     htr2c_n |
        C/T  |   .6321337   .6162262    -0.47   0.638     .0935465    4.271597
        C/C  |   1.363758   1.178227     0.36   0.720     .2508048    7.415466
             |
        race |
      Black  |   .6895325   .3381529    -0.76   0.448     .2637047    1.802983
       _cons |         .4    .334664    -1.10   0.273     .0776057    2.061704
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
```

```
->  mlogit bmi_pool i.lepr_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =      18.08
                                                Prob > chi2       =     0.0012
Log likelihood = -215.50334                     Pseudo R2         =     0.0403

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      lepr_n |
        A/G  |   2.634146   1.137351     2.24   0.025     1.130094    6.139954
        A/A  |       1.05   .5103512     0.10   0.920     .4050082    2.722167
             |
       _cons |   .3703704   .1371056    -2.68   0.007     .1792809    .7651356
-------------+----------------------------------------------------------------
Cl_I_III     |
      lepr_n |
        A/G  |   5.597561   3.267936     2.95   0.003     1.782606     17.5769
        A/A  |     1.6875   1.101313     0.80   0.423     .4695957    6.064059
             |
       _cons |   .1481481   .0793716    -3.56   0.000     .0518392    .4233841
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.

->  mlogit bmi_pool i.lepr_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(6)        =      21.65
                                                Prob > chi2       =     0.0014
Log likelihood = -213.72111                     Pseudo R2         =     0.0482

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
      lepr_n |
        A/G  |   2.379771   1.043598     1.98   0.048     1.007532    5.620973
        A/A  |   .9055872   .4484436    -0.20   0.841     .3430981    2.390244
             |
        race |
       Black |   .3862421   .2099155    -1.75   0.080     .1331212    1.120655
       _cons |   .4577492    .177253    -2.02   0.044     .2142985    .9777684
-------------+----------------------------------------------------------------
Cl_I_III     |
      lepr_n |
        A/G  |   5.483833   3.221695     2.90   0.004     1.733852    17.34428
        A/A  |   1.636856   1.080643     0.75   0.455     .4487988    5.969929
             |
        race |
       Black |   .8592671   .4379452    -0.30   0.766     .3164417    2.333257
       _cons |   .1549294    .086005    -3.36   0.001     .0521929    .4598924
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
```

```
->  mlogit bmi_pool i.oprm1_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =       9.90
                                                Prob > chi2       =     0.0421
Log likelihood = -219.59303                     Pseudo R2         =     0.0221

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
     oprm1_n |
        A/G  |   1.604542   .6610234     1.15   0.251     .7156183    3.597665
        G/G  |   7.67e-07   .0006631    -0.02   0.987            0           .
             |
       _cons |   .5814038   .1033971    -3.05   0.002     .4102984    .8238646
-------------+----------------------------------------------------------------
Cl_I_III     |
     oprm1_n |
        A/G  |   .3906867   .2575162    -1.43   0.154     .1073421     1.42196
        G/G  |   6.40e-07   .0006457    -0.01   0.989            0           .
             |
       _cons |   .5115916   .0948259    -3.62   0.000     .3557536    .7356945
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
Note: 3 observations completely determined.  Standard errors questionable.

->  mlogit bmi_pool i.oprm1_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(6)        =      13.39
                                                Prob > chi2       =     0.0373
Log likelihood = -217.85068                     Pseudo R2         =     0.0298

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
     oprm1_n |
        A/G  |   1.389961   .5818687     0.79   0.432     .6118859    3.157436
        G/G  |   3.55e-07     .00042    -0.01   0.990            0           .
             |
        race |
      Black  |   .3918285   .2109207    -1.74   0.082     .1364255    1.125373
       _cons |   .6716509   .1294526    -2.07   0.039     .4603464    .9799467
-------------+----------------------------------------------------------------
Cl_I_III     |
     oprm1_n |
        A/G  |   .3622646   .2408165    -1.53   0.127     .0984435    1.333106
        G/G  |   3.17e-07   .0004377    -0.01   0.991            0           .
             |
        race |
      Black  |   .6670389   .3250417    -0.83   0.406     .2566662    1.733539
       _cons |   .5522675   .1131161    -2.90   0.004     .3696625    .8250753
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
Note: 3 observations completely determined.  Standard errors questionable.
```

```
->  mlogit bmi_pool i.slc6a4_n if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(2)        =       1.73
                                                Prob > chi2       =     0.4204
Log likelihood = -223.67842                     Pseudo R2         =     0.0039

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
    slc6a4_n |
        A/A  |   184109.6   1.00e+08     0.02   0.982            0           .
       _cons |   3.37e-06   .0018364    -0.02   0.982            0           .
-------------+----------------------------------------------------------------
Cl_I_III     |
    slc6a4_n |
        A/A  |   .4472933   .6375228    -0.56   0.572     .0273768    7.308053
       _cons |   .9984543   1.412029    -0.00   0.999      .062452    15.96285
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.

->  mlogit bmi_pool i.slc6a4_n i.race if race < 2 , nolog rrr

Multinomial logistic regression                 Number of obs     =        215
                                                LR chi2(4)        =       5.08
                                                Prob > chi2       =     0.2794
Log likelihood = -222.00637                     Pseudo R2         =     0.0113

------------------------------------------------------------------------------
    bmi_pool |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Normal       |  (base outcome)
-------------+----------------------------------------------------------------
Over_Wt      |
    slc6a4_n |
        A/A  |   82298.33   4.48e+07     0.02   0.983            0           .
             |
        race |
      Black  |   .4001523   .2133555    -1.72   0.086     .1407268    1.137821
       _cons |   8.43e-06   .0045892    -0.02   0.983            0           .
-------------+----------------------------------------------------------------
Cl_I_III     |
    slc6a4_n |
        A/A  |    .333869   .4977018    -0.74   0.462     .0179755    6.201128
             |
        race |
      Black  |   .7083739   .3604894    -0.68   0.498     .2612694    1.920598
       _cons |   1.409502   2.118468     0.23   0.819     .0740829    26.81721
------------------------------------------------------------------------------
Note: _cons estimates baseline relative risk for each outcome.
```

. for var comp n - slc6a4 n: logistic wt NnN i.X if race < 2 \ logistic wt NnN i.X i.race if race < 2

```
->  logistic wt_NnN i.comp_n if race < 2

Logistic regression                             Number of obs     =        215
                                                LR chi2(2)        =       0.82
                                                Prob > chi2       =     0.6623
Log likelihood = -148.50064                     Pseudo R2         =     0.0028

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
      comp_n |
        A/G  |   1.196078    .419409     0.51   0.610     .6015655    2.378134
        A/A  |         .9   .3569801    -0.27   0.791     .4136357    1.958245
             |
       _cons |          1   .2948839     0.00   1.000     .5610407    1.782402
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.comp_n i.race if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =       3.28
                                                  Prob > chi2       =     0.3500
Log likelihood = -147.27106                       Pseudo R2         =     0.0110

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      comp_n |
        A/G  |   1.057385   .3831063     0.15   0.878     .5197927    2.150979
        A/A  |   .7840873   .3210065    -0.59   0.552     .3514643    1.749233
             |
        race |
      Black  |   .5301688    .217184    -1.55   0.121     .2375302     1.18334
       _cons |   1.210903   .3895841     0.59   0.552     .6445433    2.274922
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.drd2_n if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(2)        =       0.82
                                                  Prob > chi2       =     0.6649
Log likelihood =  -148.5045                       Pseudo R2         =     0.0027

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      drd2_n |
        A/G  |   .8993902   .2574903    -0.37   0.711     .5131615    1.576312
        A/A  |   1.613281   1.052264     0.73   0.463     .4492759    5.793048
             |
       _cons |   1.084746   .1957783     0.45   0.652     .7615536    1.545096
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.drd2_n i.race if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =       3.58
                                                  Prob > chi2       =     0.3100
Log likelihood = -147.12054                       Pseudo R2         =     0.0120

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      drd2_n |
        A/G  |   .9930195   .2922212    -0.02   0.981     .5577893     1.76785
        A/A  |   1.992725   1.343872     1.02   0.307     .5313815    7.472886
             |
        race |
      Black  |   .5073555   .2100133    -1.64   0.101     .2254076    1.141974
       _cons |   1.138992   .2088617     0.71   0.478     .7951196    1.631583
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.drd4_n if race < 2
```

```
Logistic regression                               Number of obs     =        215
                                                  LR chi2(2)        =       0.72
                                                  Prob > chi2       =     0.6960
Log likelihood =  -148.5502                       Pseudo R2         =     0.0024

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      drd4_n |
        C/T  |    .774578   .2421691    -0.82   0.414      .419701     1.42952
        C/C  |   .7894737    .310703    -0.60   0.548     .3650407    1.707395
             |
       _cons |   1.266667   .3093601     0.97   0.333     .7848234    2.044338
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.drd4_n i.race if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =       3.54
                                                  Prob > chi2       =     0.3155
Log likelihood = -147.14234                       Pseudo R2         =     0.0119

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      drd4_n |
        C/T  |   .7258507   .2306319    -1.01   0.313     .3893914    1.353033
        C/C  |   .7422761   .2955963    -0.75   0.454     .3400884     1.62009
             |
        race |
      Black  |   .5140227    .206716    -1.65   0.098     .2337054    1.130566
       _cons |   1.455488   .3788536     1.44   0.149       .87387    2.424212
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.gabra6_n if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(2)        =       0.89
                                                  Prob > chi2       =     0.6418
Log likelihood = -148.46924                       Pseudo R2         =     0.0030

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    gabra6_n |
        C/T  |   .7256895   .2589496    -0.90   0.369     .3605906    1.460452
        T/T  |   .8684864   .3355675    -0.36   0.715     .4072622    1.852047
             |
       _cons |         1.3   .3866523     0.88   0.378     .7257296    2.328691
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.gabra6_n i.race if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =       3.50
                                                  Prob > chi2       =     0.3206
Log likelihood = -147.16211                       Pseudo R2         =     0.0118

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    gabra6_n |
        C/T  |   .7010223   .2523968    -0.99   0.324       .34615    1.419709
        T/T  |   .8543283   .3325543    -0.40   0.686     .3983709    1.832154
```

```
             |
        race |
       Black |   .5288172    .211116    -1.60   0.111     .2418167    1.156444
       _cons |    1.45498   .4478245     1.22   0.223     .7959173    2.659783
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.htr2c_n if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(2)        =       0.93
                                                  Prob > chi2       =     0.6269
Log likelihood = -148.44572                       Pseudo R2         =     0.0031

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     htr2c_n |
        C/T  |         .8   .5727129    -0.31   0.755     .1966629      3.2543
        C/C  |   1.139241   .7416179     0.20   0.841     .3180562    4.080627
             |
       _cons |          1   .6324555     0.00   1.000     .2895029    3.454197
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.htr2c_n i.race if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =       4.01
                                                  Prob > chi2       =     0.2609
Log likelihood = -146.91004                       Pseudo R2         =     0.0134

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     htr2c_n |
        C/T  |   .8145966   .5833939    -0.29   0.775     .2001396    3.315523
        C/C  |   1.289499    .844675     0.39   0.698     .3571487    4.655787
             |
        race |
       Black |     .49664   .2009729    -1.73   0.084     .2246933    1.097724
       _cons |          1   .6324555     0.00   1.000     .2895029    3.454197
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.lepr_n if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(2)        =      16.46
                                                  Prob > chi2       =     0.0003
Log likelihood = -140.68364                       Pseudo R2         =     0.0553

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lepr_n |
        A/G  |   3.480836   1.331707     3.26   0.001     1.644481     7.36781
        A/A  |   1.232143   .5223496     0.49   0.622     .5367922    2.828238
             |
       _cons |   .5185185   .1707695    -1.99   0.046     .2719121    .9887807
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.lepr_n i.race if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(3)        =      18.38
                                                  Prob > chi2       =     0.0004
```

```
Log likelihood = -139.72513                     Pseudo R2         =     0.0617

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lepr_n |
        A/G  |   3.266138   1.262491     3.06   0.002     1.531128    6.967187
        A/A  |   1.117466   .4817015     0.26   0.797     .4800764    2.601107
             |
        race |
      Black  |   .5634846   .2354016    -1.37   0.170     .2484756    1.277852
       _cons |     .59807   .2062802    -1.49   0.136     .3042035    1.175817
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.oprm1_n if race < 2
note: 2.oprm1_n != 0 predicts failure perfectly
      2.oprm1_n dropped and 3 obs not used


Logistic regression                             Number of obs     =        212
                                                LR chi2(1)        =       0.01
                                                Prob > chi2       =     0.9249
Log likelihood = -146.70683                     Pseudo R2         =     0.0000

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     oprm1_n |
        A/G  |   1.036879   .3985671     0.09   0.925     .4881284    2.202533
        G/G  |          1  (empty)
             |
       _cons |   1.093023   .1630995     0.60   0.551     .8158574    1.464349
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.oprm1_n i.race if race < 2
note: 2.oprm1_n != 0 predicts failure perfectly
      2.oprm1_n dropped and 3 obs not used


Logistic regression                             Number of obs     =        212
                                                LR chi2(2)        =       2.76
                                                Prob > chi2       =     0.2516
Log likelihood = -145.33117                     Pseudo R2         =     0.0094

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     oprm1_n |
        A/G  |   .9260163   .3617521    -0.20   0.844      .430619    1.991334
        G/G  |          1  (empty)
             |
        race |
      Black  |   .5160462   .2083992    -1.64   0.101     .2338518    1.138771
       _cons |   1.223881   .2015523     1.23   0.220     .8862589     1.69012
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.slc6a4_n if race < 2

Logistic regression                             Number of obs     =        215
                                                LR chi2(1)        =       0.00
                                                Prob > chi2       =     0.9631
Log likelihood =  -148.9116                     Pseudo R2         =     0.0000

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
    slc6a4_n |
        A/A  |   1.067961   1.517407     0.05   0.963     .0659372    17.29739
       _cons |          1   1.414214     0.00   1.000     .0625488    15.98751
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

->  logistic wt_NnN i.slc6a4_n i.race if race < 2

Logistic regression                               Number of obs     =        215
                                                  LR chi2(2)        =       2.54
                                                  Prob > chi2       =     0.2802
Log likelihood = -147.64029                       Pseudo R2         =     0.0085

------------------------------------------------------------------------------
      wt_NnN | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    slc6a4_n |
        A/A  |   .6111111   .8953278    -0.34   0.737     .0345971    10.79444
             |
        race |
       Black |   .5246914   .2152708    -1.57   0.116      .234786    1.172561
       _cons |   1.905882    2.80646     0.44   0.661      .106336    34.15952
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.
```