Association of Gene Polymorphisms in Thyroid Function-related Genes with Differentiated Thyroid Cancer Risk and Post Thyroidectomy L-Thyroxine Suppressive Dose Requirements

Maha Meshal Alrasheed

A thesis submitted in partial fulfilment of the requirements of the University of Brighton for the degree of Doctor of Philosophy

September 2013

University of Brighton

In loving memory of my beloved uncle, Talal Abdulaziz Alrasheed (1962-2003)

Introduction: Patients with differentiated thyroid cancer (DTC) are usually managed with total thyroidectomy and subsequent radioiodine ablation of the remnant thyroid tissue. Since these patients become athyrotic, L-throxine (L-T4) therapy is required for life, in order to replace the thyroid hormones and also to suppress the endogenous thyroid stimulating hormone (TSH) which may have a growth-promoting effect on any residual thyroid cancer cells. The approximate dose required to achieve this suppression is about 2 μg/kg. However, there is wide variation between patients in their L-T4 requirement. Although factors such as the timing of the dose, compliance, weight and age play important roles, genetic factors are also thought to be important in this dose variability. Therefore, the aims of this study are to identify and evaluate the association of polymorphisms in six genes [(iodothyronine deiodinases (DIO) 1, 2 and 3, paired box gene 8 (PAX8), thyroid stimulating hormone subunit β (TSHβ), and sodium iodide symporter (NIS)], involved in thyroid hormone metabolic and functional pathways with DTC risk and L-T4 dose requirement.

Patients and Methods: Initially, 96 healthy Saudi individuals were recruited from the Family Medicine Polyclinics at KFSHRC to establish the single nucleotide polymorphisms (SNPs) present in genes of interest in the Saudi general population. This was followed by association studies in 507 patients and 597 controls. Candidate patients had undergone total thyroidectomy and received radioiodine ablation. They were on L-thyroxine suppressive therapy (Euthyrox, Merck Pharmaceuticals, NJ, USA), aiming at obtaining either suppressed (TSH <0.1 mU/L) or near-suppressed (0.1≤TSH<0.5 mU/L) serum TSH levels with FT4 in the normal range (12-22 pmol/L). Identification of SNPs was accomplished by sequencing all exons and exon-intron boundaries using MegaBACE DNA analysis system, and the association of the gene variants with risk of thyroid cancer and/or thyroxin dose requirement in the patient population was accomplished using real-time PCR with Taqman chemistry.

Result: Several novel and familiar SNPs were described in the studied genes with a minor allele frequency of 0.05 in the general population, of which 39 were selected for association studies. Following the adjustment for age, gender and smoking, the rs1321108, rs1321109 of the TSHβ gene, rs11123172, rs67776659 of the PAX8 gene and rs945006 of the DIO3 were associated with DTC. Interestingly, several haplotypes constructed from the studied SNPs in PAX8, TSHβ and NIS were associated with disease. The rs12084242, rs12095080, rs17109582 of DIO1, rs1321109 of TSHβ, rs2241975, and rs3738913 of PAX8 were associated with thyroxine dose requirement in the suppressed group, rs12095080 of DIO1 as well as rs2241975 and rs4849186 of PAX8 were associated with the difference in thyroxine dose requirement in the combined group, and rs4849186 and rs1478 of PAX8 were associated in the near-suppressed group.

Discussion and Conclusion: The study has revealed that TSHβ, NIS PAX8 and DIO3 are associated with DTC risk, and changes in DIO1, TSHβ and PAX8 may be important with respect to the requirement of altering thyroxine dose in therapy of DTC.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to acknowledge my supervisors, Dr. Paul Gard, Andrew Overall, Angela Macadam and Nduna Dzimiri for the superb guidance throughout my studies. I would never have been able to accomplish this work without all your advices. In particular my special thanks go to Dr. Gard for the being there for me every time I called upon him to provide me with all the necessary academic support. Equally I owe my deep-felt gratitude to Dr. Dzimiri for being always there for me whenever I needed him for both academic and personal support. I also would like to thank Dr. Ali Alzahrani for identifying the huge volume study candidates through his clinic and his contribution to the aims of the study, without whom it would have been impossible to perform the study in its present form.

My great appreciation goes to my family, in particular my father for standing by my side as well as his encouragement and motivation throughout my studies, my husband, Khaled, my daughter Haifa and son Sultan for the sacrifices they had to make to facilitate the smooth running of my studies, and my mother, sisters and brothers, especially Noura, Aljohara and Haifa, for their emotional, moral and personal support.

I also feel greatly indebted to the School of Pharmacy and Biomolecular Science of the University of Brighton for providing me with the opportunity to carry out my studies with them.

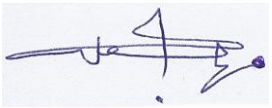
I also express my gratitude to King Saud University for funding my studies through the External Joint Supervision Program.

I would like to acknowledge Dr. Sultan Al-Sedairy, the Director of the Research Centre of King Faisal Specialist Hospital and Research Centre for providing technical support required in my study, for which I am deeply thankful.

Finally I would like to thank many friends whom I could not mention by name, for all their support, encouragement and motivation throughout my studies.

Declaration

I declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the original work of the author. The thesis has not been previously submitted to this or any other University for a degree, and does not incorporate any material already submitted for a degree.



Maha Meshal Alrasheed

Date: 30th September, 2013

1. Introduction

1.1. Foreword

Variations in the human genome are an important cause of human susceptibility to almost all diseases as well as alterations in the way different individuals respond to therapy with drugs and other xenobiotics. These changes play an important role in the inter-ethnic differences, particularly in response to therapy of complex disorders, such as cancer and cardiovascular diseases. The differences may manifest at the level of drug-receptor interaction or drug metabolic processes. In recent years, great strides have been made in understanding the role of gene polymorphisms in the pathogenesis and prognosis of various forms of malignancies such as breast, lung, gastric, leukaemic, lymphoid and thyroid cancers. In thyroid cancer, research has focused primarily on genetic variations in the thyroid stimulating hormone receptor (TSHR), transthyretin and the deiodinases, and their association with thyroid tumour formation. The production of the thyroid hormones is regulated by the hypothalamus–pituitary–thyroid axis, whereas their biological activity is regulated mainly at the tissue level by the iodothyronine deiodinases and thyroid hormone transporters. The genes that encode for all of these proteins are known to exist in different polymorphic forms, some of which are believed to be responsible for the differences observed in patient responses to drug therapy and in thyroid hormone bioactivity. However, many questions remain unanswered with regard to their role in thyroid cancer risk and effect on its management. In this study, key genes involved in thyroid hormone signalling and metabolic pathways were screened for polymorphisms in patients with differentiated thyroid cancer and matching controls. Subsequently, possible associations of these polymorphisms with the risk of development of thyroid cancer as well as variations in thyroid hormone requirement following thyroidectomy were analysed in the Saudi population.

1.2. Gene polymorphism, disease and response to drug therapy

Although less than 2% of the human genome is composed of genes that actually code for functional proteins, studies have shown that at least 60% of the genome is transcribed. However, most of the transcription is geared towards generating transfer and ribosomal RNAs as well as other RNAs involved in processes such as splicing of mRNA and the regulation of gene expression. Thus variations occurring in areas other than the coding loci, can also lead to adverse health effects.

Gene polymorphisms are variations in the DNA sequence between individuals, groups, or populations (Abecasis *et al.*, 2010; Watson JD *et al.*, 1996). This may occur as a result of, for example, a mutation either inherited from the parents or induced by various factors, such as viruses or radiation. These variations bring about different forms of the gene and often lead to a diverse phenotypic expression of the gene within a given population or species (Chakravarti, 1999). The majority of such mutations cause a change at a single base in the DNA sequence leading to genetic variations called single nucleotide polymorphisms (SNPs). The SNPs may be a result of two types of nucleotide (nt) base substitutions, which are defined as transition or transversion substitutions (Watson JD *et al.*, 1996). Accordingly, a transition substitution occurs between the two purine bases, adenine and guanine, or between the two pyrimidine bases, cytosine and thymine. This type of change constitutes two thirds of all SNPs. A transversion substitution, on the other hand, occurs as result of an exchange between a purine and a pyrimidine base. These changes in the DNA sequence are classified as missense, nonsense, insertion/deletion, duplication, frameshift or repeat expansion mutations(Kogai *et al.*, 2012). They are also called non-synonymous or synonymous, whereby the former describes a point mutation in which the change in the nucleotide results in an alteration in the resultant base, while a synonymous one does not lead to such changes. In a missense mutation, the new base change leads to an alteration of the

codon resulting in a different amino acid being incorporated into the protein sequence. Mutations that change the protein code into a stop codon are termed nonsense mutations. Such mutations lead to the synthesis of truncated proteins, rendering the native proteins non-functional, and are a primary cause of so-called monogenic disease. Silent mutations on the other hand cause no change in the final protein in product and can only be detected by sequencing the gene. Such changes are usually not detrimental. Extra base pairs may also be added or deleted from the DNA sequence, a process known as insertion/deletion mutation, which can involve one to several thousand base pairs. Deletions of a single or two bases or multiples of two bases lead to the shifting of the reading frame, known as frameshift mutation. This process may be detrimental, as the mRNA would be translated into new groups of codons and the resultant protein may be non-functional.

The sequencing of almost the entire human genome by the Human Genome Project about two decades ago revealed that SNPs make up the most common type of genetic variations in the human genome, and the majority of those within the coding sequences (cSNPs) often lead to a predictable change in the protein sequence and function (Altshuler *et al.*, 2010; Cargill *et al.*, 1999; Lander *et al.*, 2001). Thus, while the genomes of individuals are 99.9% identical, the 0.1% difference means there are millions of polymorphisms, the most common being the SNPs (Chakravarti, 1999; Sachidanandam *et al.*, 2001). Hence, SNPs are the simplest form and most common source of genetic differences in the human genome possibly constituting 90% of all human DNA polymorphisms. To date, more than 10 million common SNPs have been identified in human genes, with more being reported every day (Sachidanandam *et al.*, 2001). It is now well established that the number of SNPs varies within one gene and from one population to another, with some genes displaying more than 100 SNPs (Chakravarti, 1999; Lander *et al.*, 2001). This points to a high level of human protein

diversity, and may therefore have great impact on disease manifestation and ways in which patients respond to drug therapy.

Currently, SNPs are being used as markers in human genetic studies ranging from comparing variations among populations to disease linkage studies (Chakravarti, 1999; Collins *et al.*, 1997; Kruglyak, 1997; Lander, 1996; Risch *et al.*, 1996), but they can also provide important leads to mechanisms involved in common and genetically complex human diseases as well as pharmacogenetic traits (Risch *et al.*, 1996). Until recently, SNPs were believed to present the most important type of genetic variations associated with disease and differences in patient response to drug therapy. However, in the last decade it has become increasingly evident that individual SNPs have poor predictive power as pharmacogenetic loci. Instead, it is multiple SNPs within a region of a chromosome or gene that ultimately affect biological and therapeutic phenotypes (Nebert *et al.*, 2008; Vormfelde *et al.*, 2007). Thus, it is postulated that certain, variations in human genes occur together in three to five different patterns as combination(s) of alleles of closely linked loci that are found in a single chromosome and tend to be inherited together as haplotypes (Gabriel *et al.*, 2002; Michalatos-Beloin *et al.*, 1996; Stephens *et al.*, 2001). These haplotypes may be located on genome segments often containing one or more genes, several gene regulatory segments, as well as fragments with no known function. It is thought that for example, among Africans, each haplotype block constitutes about 11,000 nucleotides, whereas European and Asian haplotype blocks comprise about 22,000 (Gabriel *et al.*, 2002; Michalatos-Beloin *et al.*, 1996; Stephens *et al.*, 2001). These manifestations indicate that, for a particular disease, the same combination of SNPs either in a single gene or a cluster of genes may ultimately determine the pharmacogenomic traits worldwide. Accordingly, the currently available data overwhelmingly suggest that the identification of patterns of genetic variations, in the form of haplotypes rather than single or point variations per se,

may present a more promising approach to predicting a person's response to drug therapy, and allow for highly individualized treatment of patients with complex diseases, such as cancer or cardiac disorders and their associated risk factors, including hyperlipidaemia, obesity, diabetes or hypertension. To date, several haplotypes are already known to be the potential underlying causes for differences in patient responses to drug therapy of various diseases, including heart failure and cancer (Bai *et al.*, 2009; Lunceford *et al.*, 2008; Pacanowski *et al.*, 2008; Tzvetkov *et al.*, 2008).

1.2.1. Gene polymorphism in the general population

The fact that the number of SNPs varies from one gene to another, with some genes displaying more than 100 SNPs (Chakravarti, 1999) indicates a high level of genetic variation among ethnic groups, such as African, Asian, and Caucasian, in several genes that determine the extent of response to several types of drug therapy (Kurose *et al.*, 2012; McGraw *et al.*, 2012; Polimanti *et al.*, 2012). Furthermore, because ethnic groups reflect an absence of gene flow over evolutionary timescales, they can differ in a number of important ways. Thus, due to the nature of genetic inheritance, the more distant the common ancestry is between any two individuals, the less likely these individuals are to share ancestral mutations. Accordingly, new, recent mutations are likely to have occurred in the process of evolution. Besides, it is known that there is far more genetic variation within any ethnic group than there is between them, and certain SNPs may be more prevalent in particular ethnic groups simply as a result of their evolutionary history. Hence, while this paradigm may be valid with respect to the applicability of population-based genotyping in general, the ultimate impact of gene variants may vary between ethnic groups or even individuals within the same population or family. Specifically, the possible differences among individuals or ethnic groups may play an important role in non-Mendelian genetic disorders or complex diseases such as cancer or cardiovascular disease, where the actual manifestation of the disease

depends greatly on the environment. For this reason, it is essential to have adequate information on the actual prevalence of such mutations in order to characterize and ascertain their true impact for any given ethnic group of interest. To date, several studies have indeed demonstrated variability in many polymorphisms among different ethnic groups. Such differences in some important disease-causing or drug response-related alleles have been described for various gene families, with great impact on patient care strategies (Zuo *et al.*, 2012). Such information on population allelic frequency distribution is becoming increasingly helpful in explaining several therapeutic issues such as therapeutic failures, adverse drug reactions (ADRs), potential risk groups and the optimal doses for therapeutic efficacy. One important example is that of the cytochrome P450s family of genes, which shows variation among different ethnic groups such as African, Asians and Caucasians in several alleles that determine the extent of response to therapy with several types of drugs (Kudzi *et al.*, 2009; Kudzi *et al.*, 2010; Veiga *et al.*, 2009). These points are discussed further in detail below.

1.2.2. Gene polymorphism and disease

Genetic disorders can be grossly classified into monogenic (Mendelian) or complex polygenic inheritance in nature. Monogenic disorders occur in a definable fashion, as autosomal recessive, autosomal dominant, X-linked or mitochondrial defects (Kumar, 2008). Thereby, in autosomal recessive disease, two copies of the mutated gene are needed to cause the disease; autosomal dominant requires only one copy of the gene to be present, while in the X-linked disease, the gene resides on the X chromosome. Unlike the Mendelian disorders which are caused by a mutation in the DNA sequence of a single gene, most complex diseases (also sometimes referred to as complex traits, polygenic/polygenetic diseases) arise from intricate interactions of a combination of environmental factors and mutations in several genes (Kumar, 2008). Sequence variations leading to human gene polymorphisms are thought to predispose individuals

to a wide range of complex diseases including cancer (Gabriel *et al.*, 2002; Hemminki *et al.*, 2006; Lander, 1996; Shah, 2007), cardiovascular diseases (Chakravarti, 1999; Collins *et al.*, 1997; Kruglyak, 1997; Lander, 1996; Risch *et al.*, 1996; Shah, 2007), diabetes (Rigoli *et al.*, 1995; Ukkola *et al.*, 1994), hypertension (Lifton, 1996; Smithies *et al.*, 1995), hyperlipidaemia (Gudnason *et al.*, 1999; Kakko *et al.*, 2001; Kuivenhoven *et al.*, 1997) or obesity-related cardiovascular disorders (Ehrenborg *et al.*, 2000; Rajput-Williams *et al.*, 1988). However, although great progress has been accomplished in identifying genetic factors underlying complex disorders, much still remains to be learned in this regard. Moreover, in the short term this in itself does not offer much in the way of patient benefit. For this reason, much interest continues to be nurtured in identifying possible genetic targets for enhancing drug therapy, which is simpler and offers short- term benefits, if not a cure.

1.2.3. Gene polymorphism and patient response to drug therapy

It is now acknowledged that variation in the human genome is an important cause of differences in response to drugs and other xenobiotics, and that susceptibility to almost all diseases is determined to some extent by these genetic variations (Brockmoller *et al.*, 2008; Cargill *et al.*, 1999). The variability in drug responses can occur at the receptor-drug interaction level or in the drug biotransformation process, as a result of polymorphic changes in a receptor or metabolizing enzyme, respectively. Diet and other environmental factors, such as life style or co-medication, are additional sources of apparent ethnic differences in factors that influence the amount of a drug reaching its targets (Pharmacokinetics)(Chen, 2006). Furthermore, the majority of inter-ethnic differences in the pharmacokinetics of a drug are also attributable to the variability in polymorphic traits of metabolic pathways (Altshuler *et al.*, 2010; Brockmoller *et al.*, 2008; Wood *et al.*, 1991).

In recent years, there has been an explosion of studies targeted at understanding the role of genetic factors in drug metabolism. These studies have led to the birth of the concepts of Pharmacogenetics and Pharmacogenomics. Pharmacogenetics is a discipline that focuses on the interaction of single genes with drugs, leading to different phenotypes or variation in drug responses, by correlating gene expression or SNPs with the efficacy or toxicity of the drug. These phenotypes can range from life-threatening adverse drug reactions (ADRs) at one end of the spectrum to lack of therapeutic efficacy on the other. Pharmacogenomics, on the other hand, embraces the whole genome application of Pharmacogenetics on the influence of genetic variations on drug responses in patients. Thus, currently studies increasingly involve entire pathways encoding proteins that influence Pharmacokinetics (metabolic pathways), Pharmacodynamics and genome-wide approaches in understanding the influence of gene polymorphism in disease pathways and therapy thereof.

To date great strides have been made in understanding the influence of genetic changes in metabolic pathways for several families of drugs. Drug metabolizing enzymes (DMEs) are a diverse group of proteins that metabolize a vast array of compounds including drugs, environmental pollutants, and endogenous compounds such as steroids and prostaglandins. They are found mainly in the liver and are sometimes called mixed-function oxidases or monooxygenases. These enzymes are classified as either phase I or phase II enzymes. Phase I enzymes, such as the cytochrome P450s (CYP450s), are involved in phase I reactions, including oxidation, reduction, hydrolysis, cyclization and decyclization reactions, whereas phase II enzymes are engaged in conjugation reactions (e.g. involving glucuronic acid, sulfation, glutathione or amino acids), which are detoxification reactions and involve interactions of polar functional groups of phase I metabolites. Phase 2 enzymes facilitate the elimination of drugs and the inactivation of carcinogenic metabolites produced by the CYP450s. The balance between the phase 1

and 2 enzymes usually determines the metabolic fate of a particular chemical. The expression of these enzymes can markedly vary between individuals owing to differences in the extent of induction and by polymorphisms. Thus, inter-ethnic differences in the Pharmacokinetics of a drug may be attributed to structural variations in the gene involved in its metabolic pathways. These differences can manifest themselves in different ways, including changes in drug potency or metabolism.

Furthermore, since an individual receives one allele from each parent, different phenotypes can result depending on the effect of possible combinations of resultant variants on the metabolic rate of the enzyme. Such variations in metabolic activity also exist not only among individuals but also among ethnic groups. For example, about 7% of whites and 2 - 7% of blacks are poor metabolizers of certain drugs that are dependent on cytochrome P 2D6 (CYP2D6), the enzyme that metabolizes many β-adrenergic receptor blockers (β–blockers), antidepressants and opioids (Chong *et al.*, 2003). It is also thought that 20% of Asians are poor metabolizers of drugs, such as phenytoin, phenobarbital, omeprazole, and other agents that are dependent on CYP2C19 (Chong *et al.*, 2003). Accordingly, individuals may be classified as poor, intermediate, efficient or extensive (ultrarapid) metabolizers, depending on how slow or fast a drug is metabolized by the form of CYP450s harboured by the person. Thus, poor metabolizers carrying two copies of defunct variant alleles would lack functional enzyme and intermediate metabolizers would be heterozygous for a defective allele. On the other side of the spectrum are efficient metabolizers carrying two functional gene copies and ultrarapid metabolizers carrying more than two functional gene copies (i.e. inheriting multiple copies of wild-type alleles), which results in excess enzyme activity (Christoffolete *et al.*, 2006). Abolished enzyme activity is commonly seen where a deletion of a whole gene has occurred. It also has its origin in mutations causing altered splicing, stop codons, abolished transcriptional start sites and amino acid changes

leading to disease-related activities. Mutations in substrate recognition sites can also cause the synthesis of enzymes with altered substrate specificity (Dumitrescu *et al.*, 2005).

Apart from the CYP450 family, there are other enzyme pathways such as deiodination, deamination and glucuronidation that are also engaged in the biotransformation of several xenobiotics. Besides, defects in transporting proteins for these agents can also influence drug actions. These include, among others, organic anion and amino acid transporters, which are also important in the regulation of thyroid hormone function. This study will focus on these pathways that are particularly important in iodothyronine metabolism, and specifically on genes associated with the development of thyroid cancer, and their possible influence on response to drug therapy for this condition.

1.3. The thyroid hormone system

1.3.1. The thyroid gland

The thyroid gland (also simply referred to as the thyroid) comprises two cone-like lobes (wings) defined as lobus dexter (right lobe) and lobus sinister (left lobe) that are connected via the isthmus. The gland resides on the anterior side of the neck, against the larynx and trachea, reaching posteriorly to the esophagus and carotid sheath (Figure1). The thyroid gland controls the rate at which the body utilizes energy, makes proteins, and controls the sensitivity of the body to other hormones.

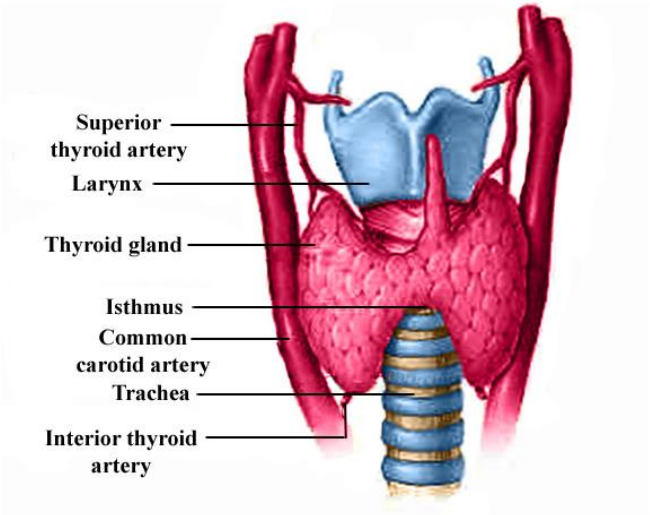


Figure 1. Schematic diagram of the thyroid gland showing the surrounding organs (modified) adopted from "The anatomy of the thyroid gland, Yale Medical group"

It does so by producing thyroid hormones, particularly triiodothyronine (T3) and thyroxine (tetraiodothyronine, T4) from the thyroid epithelial (follicular) cells, both of which regulate metabolic rate and influence the functional rate of multiple other systems in the body. The T3 and T4 are synthesized from iodine and tyrosine. Additionally, the thyroid gland also produces calcitonin in the parafollicular cells (also known as C-cells of the thyroid), which plays a role in calcium homeostasis (Figure 2). Output of hormones from the thyroid is regulated by the thyroid stimulating hormone (TSH) which is produced in the anterior pituitary gland, which in turn is regulated by thyrotropin-releasing hormone (TRH) produced by the hypothalamus. The thyroid selectively absorbs iodine (in form of iodine ions) from the blood in order to produce the thyroid hormones as well as storing iodine in the thyroglobulin. The most common problems of the thyroid gland are hyperthyroidism (overactivity) and hypothyroidism (underactivity) of the gland. However different forms of thyroid cancer are also prevalent.

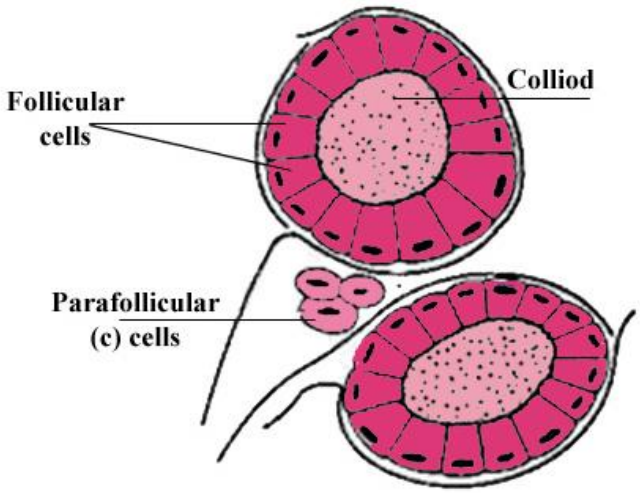


Figure 2. Functional unit of the thyroid (modified) adopted from www. Surgicalnotes.co.uk 2007-2011. (C) cell represents chief cells

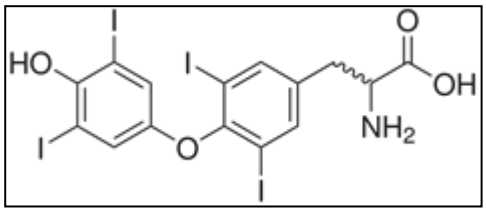
The parathyroid glands also small endocrine glands which are situated on the rear surface of the thyroid gland, between the two layers of the capsule and on the posterior side of the lobes. Normally humans have four of these organs, which control the amount of calcium in the blood and within the bones.

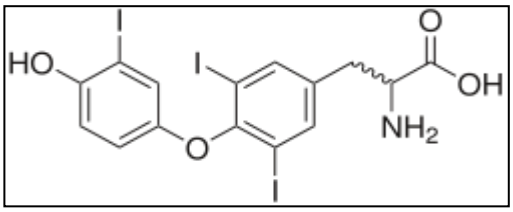
1.3.2. The synthesis and structural features of thyroid hormones

Thyroxine (T4) and triodothyronine (T3) are tyrosine-based thyroid hormones (THs) secreted by follicular cells of the thyroid gland. They are essential for the proper development and differentiation of all cells in the human body. They function by promoting the basal metabolic rate in body cells, which ultimately affects the rate of protein synthesis, neural maturation, and the regulation of growth in long bones. They also cause elevation of the body's response to catecholamines such as adrenaline.

The two THs are composed of two benzene rings: an inner tyrosyl ring, often also called the alpha ring, and an outer phenolic or beta ring (Figure 3). The first step in their physiological synthesis involves the organification of iodine (I_2) (Figure 4). Accordingly, I_2 is taken up and oxidized by hydrogen peroxide in the thyroid gland cells. It then binds to the 3 position of the tyrosyl ring in a reaction catalysed by thyroid

peroxidase enzyme, and then condensed onto tyrosine (Tyr) residues which reside along the polypeptide backbone of a protein molecule termed thyroglobulin.





3,5,3',5'- Tetraiodothyronine(T4) 3,3',5- Triiodothyroine (T3)

Figure 3: The thyroid hormone structures (Wang *et al.*, 2010)

This reaction results in either a mono-iodinated tyrosine (MIT) or 3,5-di-iodinated tyrosine (DIT) being incorporated into thyroglobulin to form iodothyroglobulin, one of the most important constituents of the colloid material present in the follicle of the thyroid unit. The other synthetic reaction closely linked to organification is the coupling of the iodotyrosine molecules. Thereby, coupling of two di-iodotyrosine molecules by the deiodinases (5'-deiodinase) results in the formation of T4, while that of a di-iodotyrosine and a mono-iodotyrosine leads to the formation of tri-iodothyronine T3 or its isomer, the reverse T3 (rT3). These are further processed by decarboxylation and deiodination to produce iodothyronamine (T1a) and thyronamine (T0a). The di-iodotyrosine coupling to produce T4 is the major reaction with respect to the formation of THs.

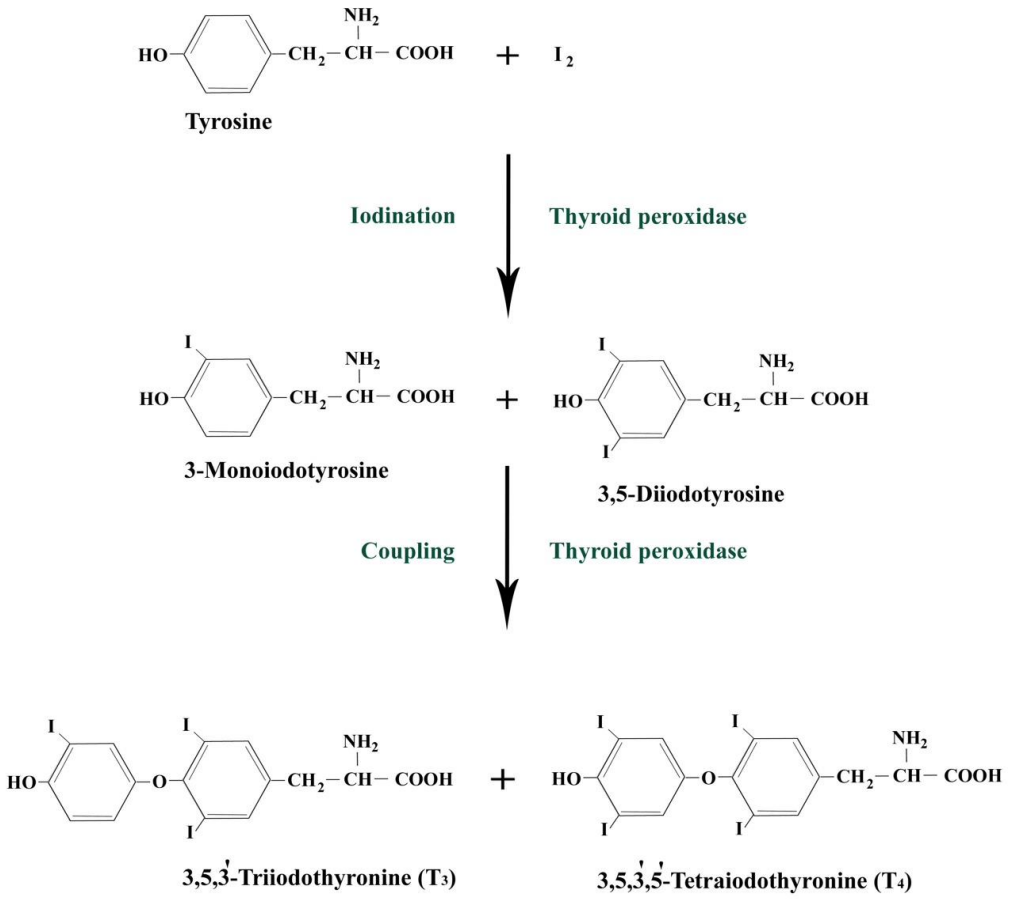


Figure 4: Physiological synthesis of thyroxine from Iodine and Tyrosine. I_2 – Iodine (modified according Miot et el) (Miot *et al.*, 2012)

The biological activity of TH i.e. the availability of the active hormone T3 for the nuclear TH receptors is mainly regulated at the tissue level by the iodothyronine deiodinases and TH transporters (Streckfuss *et al.*, 2005). Biological metabolism of thyroxine which involves deiodination, deamination, increasing the solubility of the phenol through oxidation as well as conjugation with glucuronic acid and sulphate occurs mostly in the liver, while decarboxylation occurs in the intestines. The metabolites are excreted in urine (Christoffolete *et al.*, 2006). Although T4, T3, and rT3 are synthesized within the thyroid gland, the T4 is the primary product that is secreted from the organ (Kohrle *et al.*, 1988; Robbins, 1981; Visser, 1996). Furthermore, while some T3 is produced in the thyroid and is biologically more active than T4, the major production of T3 actually occurs outside the thyroid gland, primarily by conversion of

T4 in the liver and kidneys through enzymatic removal of one I_2 from the outer ring of T4 (Chopra, 1976; Utiger, 1980). Accordingly, after the THs are formed, lysosomal proteases break up T4, as well as any T3 or rT3 formed, from the thyroglobulin molecule, and the hormones are released into circulation. The T3 and T4 released from the thyroid by proteolysis reach the bloodstream. There they will be bound to TH binding proteins and transported to their points of action (Robbins, 1981; Shen *et al.*, 2004; Visser, 1996).

1.3.3. Thyroid stimulating hormones

The thyroid stimulating hormones are a family of substances known as endogenous ligands. They are encoded by genes that share a number of characteristics. The hormone consists of two subunits known as the alpha (TSHα) and beta (TSHβ) subunits. When combined, the two subunits produce the active form, whereby the beta subunit (also called the buckle or seatbelt) wraps around the alpha subunit to form the functional hormone. The TSH is synthesized in the pituitary gland which is found at the base of the brain. It plays a major role in growth and in the function of the thyroid gland, and stimulates the production of thyroid hormones. This hormone plays an important role in regulating growth, brain development and metabolism. Its levels are regulated by the pituitary gland. Thus, when the thyroid hormone levels are too low, the pituitary gland releases the TSH into the bloodstream, which in turn signals the growth thyroid gland and production of the thyroid hormones.

1.3.4. Thyroid hormone metabolic pathways

The major processes in the regulation of the biological metabolism of thyroxine are deiodination, deamination and glucuronidation. Besides, the local availability of various iodothyronine metabolites is regulated through TH metabolism. Thus, during the development and maintenance of an adult organism, TH mediates their effects on gene expression, thermoregulation, energy metabolism, and many key reactions (Bianco

et al., 2006; Visser *et al.*, 1981). Cellular entry is required for the conversion of THs by the intracellular deiodinases and for binding of T3 to its nuclear receptors. These processes are described in greater detail below.

1.3.4.1 Iodothyronine deiodination

The family of three selenoproteins known as iodothyronine deiodinases (DIOs) types I, II and III (D1, D2 and D3, respectively, which are also sometimes referred as DIO1, DIO2, DIO3) (Braverman, 1994; Hesch *et al.*, 1975; Kohrle, 1999; Pitt-Rivers *et al.*, 1955) regulate the concentrations of TH in peripheral tissues through deiodination. This process involved the removal of I_2 moieties from compounds such as the THs, which is an important pathway of TH metabolism. These DIOs constitute a group of dimeric integral membrane protein that activate or inactivate TH (Bianco *et al.*, 2005; Callebaut *et al.*, 2003; Curcio-Morelli *et al.*, 2003; Leonard *et al.*, 2001), depending on whether they act on the phenolic or tyrosil rings of the iodothyronines, respectively (Kuiper *et al.*, 2005) (Figure 4). Since T4 is activated by outer ring deiodination (ORD) to T3, whereas both T4 and T3 are inactivated by inner ring deiodination (IRD) to 3,3',5-triiodothyronine and 3,3'-diiodothyronine, respectively (Figure 5) (Kuiper *et al.*, 2005), the DIOs inactivate the TH by removing specific I_2 moieties from the precursor molecule. Although these enzymes are homologous, they differ in a number of important features, including the regulation during foetal and neonatal development by thyroid state and during illness, their catalysis of outer ring deiodination (ORD) and/or IRD of sulfated iodothyronines and their inhibition by the thyrostatic drug propylthiouracil (Kuiper *et al.*, 2005). Besides, the enzymes also display different biochemical and regulatory properties, tissue distribution, developmental patterns of expression, as well as different responses to deiodinase inhibitors (Table 1).

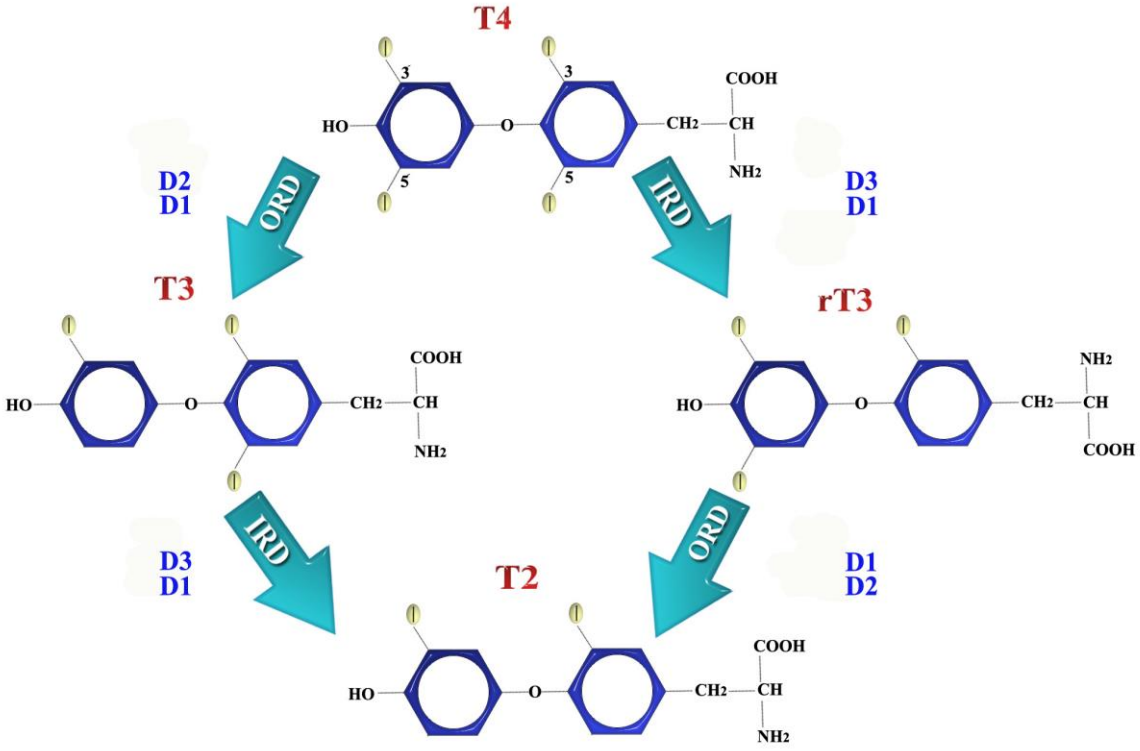


Figure 5: Major pathways of thyroid hormone deiodination. DIO1, DIO2, DIO3, deiodinase types 1, 2 and 3.The basic deiodination reactions are catalyzed by the deiodinases. These enzymes remove iodine moieties from the phenolic (outer rings) or tyrosil (inner rings) rings of the iodothyronines. These pathways can activate T4 by transforming it into T3 (via DIO1 or DIO2) or prevent it from being activated by converting it to the metabolically inactive form, reverse T3 (via DIO1 or DIO3). T2 is an inactive product common to both pathways that is rapidly metabolized by further deiodination (modified according Darras, V.M. and S.L. Van Herck (Darras et al., 2012).

The DIO1 is a kinetically inefficient enzyme that activates or inactivates T4 on an equimolar basis (Torlontano *et al.*, 2008). It is expressed mainly in the liver, kidney, and thyroid, and plays a key role in the production of the active hormone T3 from T4 and in the clearance of the metabolite rT3 (Bianco *et al.*, 2002). Its role in health remains unclear. The hepatic enzyme probably contributes to peripheral T3 production and may constitute the main site for the clearance of plasma rT3, mediated by the ORD activity of DIO1 (Table 1). However, this enzyme also possesses IRD activity, especially towards sulfated T4 and T3. Thus, in addition to the bioactivation of T4 to T3, DIO1 also catalyses hormone degradation. It is also distinguishable from the other

deiodinases by its sensitivity to inhibition by the anti-thyroid drug propylthiouracil (PTU) which is used in the treatment of hyperthyroidism.

The DIO2 is present in brain, pituitary, brown adipose tissue, thyroid, skeletal muscle, aortic smooth muscle cells, and osteoblasts, and has also been detected in the human heart (Bianco *et al.*, 2002). The enzyme catalyses the conversion of T4 to the more active form of the TH, T3. The latter is a critical regulator of thermogenesis and glucose metabolism. The DIO2 possesses only ORD activity, which changes in response to alterations in the thyroid state to maintain tissue T3 levels in the presence of varying plasma T4 and T3 levels. Hence it is thought to be important for the generation of local T3 in various tissues, but does not contribute much to the production of plasma T3. Accordingly, its regulation of local TH bioactivity is limited to certain tissues, such as the brain and skeletal muscle. In the brain, it is important for local production of T3, whereas in skeletal muscle it may also contribute to plasma T3 production (Bianco *et al.*, 2002; Maia *et al.*, 2011b; Peeters *et al.*, 2003c) (Table 1). It also plays a central role in the negative feedback regulation of TSH secretion in several human tissues, including hypothalamus and pituitary by converting T4 to T3 (Torlontano *et al.*, 2008).

Property	Type I 5'-deiodinase	Type II 5'-deiodinase	Type III 5'-deiodinase
Function	Systemic >local T3 production, degradation of rT3 and sulfated iodothyronines	local >systemic T3 production	Inactivation of T4 and T3
Expression	liver, kidney, thyroid, pituitary, heart, brown adipose tissue in sheep	(hypothyroid) pituitary, brain, brown adipose tissue, skin, placenta; thymus, pineal glands; glial cells	Placenta, brain; many tissues; not pituitary, thyroid, kidney, adult liver
Cosubstrate	DTT or DTE in vitro (K_M, 2–5 mM); not glutathione or thioredoxin in vivo; intermediate formation of an oxidized selenium residue.	DTT or DTE in vitro (K_M, 5–10 mM); higher concentrations than for DIO1; sequential two-substrate reaction with no intermediate formation of a selenium residue.	DTT or DTE in vitro (K_M, 10–20 mM); higher concentrations than for DIO1
Subcellular location	Endoplasmic reticulum in liver, inner plasma membrane in kidney and thyroid	Inner plasma membrane; p29 subunit associated with F-actin and perinuclear vesicles respectively	Endoplasmic reticulum
Cloned in species	human, rat, mouse, dog, chicken; not expressed in *Rana catesbeiana*, *Oreochromis niloticus* (tilapia), rainbow trout	human, rat, mouse, chicken, *Rana catesbeiana*, *Fundulus heteroclitus* (teleost), rainbow trout	human, rat, mouse, chicken, *Rana catesbeiana*, *Xenopus laevis*
Essential amino acid residues for catalytic activity	histidine, selenocysteine, cysteine, phenylalanine	Selenocysteine (?)	Selenocysteine
Enzyme Induction	T3, retinoids; TSH and cAMP in thyroid only; testosterone (liver), carbohydrate	cAMP; FGF; phorbol esters via PKC; ANP and CNP via cyclic GMP in glial cells	T3, FGF, EGF
Inhibition	PTU, iodoacetate, aurothioglucose, iopanoate	T4, rT3, iopanoate	Iopanoate

Table 1. Deiodinase isoenzymes, biochemical characteristics, regulation and function. In general, a given cell type will express only one type of deiodinase at a given time, though some tissues express none, and all 3 types of activity have been measured in the pituitary gland. FGF: fibroblast growth factor. PKC: protein kinase C. ANP: atrial natriuretic peptide. CNP: C-type natriuretic peptide. K_M, substrate concentration at which the reaction rate is half its maximum value. EGF, epidermal growth factor, adopted from Kohrle (Kohrle, 2000; Kohrle *et al.*, 2005).

DIO3 is present in brain, skin, placenta, pregnant uterus as well as various foetal tissues, and is often induced in critical illness (Bianco *et al.*, 2002; Peeters *et al.*, 2003a). Since it has only inner ring deiodination (IRD) activity, this 5'-deiodinase mediates the

degradation of TH. It is the major T3 and T4 inactivating enzyme, which contributes to TH homeostasis by protecting tissues from excess hormone (Table 1).

1.3.4.2. Other thyroid hormone metabolic pathways

Apart from deiodination, THs also undergo glucuronidation and sulfation. Thereby, T4 glucuronidation is mediated by uridine 5'-diphosphate (UDP)-glucuronosyltransferases (UGTs), such as UGT1A8 and UGT1A10 (Tong *et al.*, 2007). These enzymes are expressed in the gastrointestinal tract (but not the liver) and may be important for first-pass metabolism of the T4 metabolically active molecule T3 in a time-and tissue-specific fashion (Gereben *et al.*, 2008). Tyrosine sulfation is an important pathway in the metabolism of TH, in which its degradation by the type 1 deiodinase is accelerated (Yoder Graber *et al.*, 2007). This is a posttranslational modification process where a sulphate group is added to a tyrosine residue of a protein molecule, as happens with secreted proteins that pass through the Golgi apparatus.

1.3.5. Thyroid hormone transporters

The THs act intracellularly, which requires them to be transported by TH transporters across the cell membrane. A number of transporters including monocarboxylate transporters (MCTs), organic-anion transporting proteins (OATPs) and sodium iodide symporter (NIS) are involved in the transportation of THs. Several transporter families have been identified to date. However, only the monocarboxylate transporters 8, 10 (MCT8, MCT10), and organic-anion transporting protein 1C1 (OATP 1C1) show a high degree of specificity towards TH (Visser *et al.*, 2008). Among the OATPs, the Na^+ taurocholate co-transporting polypeptide (NTCP) and various other members of the OATP family mediate iodoothyronine transport. While the OATPs are known to be capable of transporting iodothyronine sulphates such as T4 sulphate (T4S), T3S, and rT3S, the major hepatic transporter for these conjugates has not yet been identified. Among these, OATP1C1 has a high affinity and specificity for T4 (van der Deure *et al.*,

2008b). Accordingly, it mediates the transport of T4, T4S, and rT3 and increases the access of these substrates to the intracellular active sites of the deiodinases. The expression of the T4 in capillaries throughout the brain suggests a critical role for its transport across the blood-brain barrier. Among the human MCTs, MCT8 is not only highly expressed in the liver and brain, but also widely distributed in other tissues, and shows preference for T3 as its ligand. Other transporters include the threonine (T) amino acid transporter 1 (TAT1), which does not however, transport iodothyronines (Friesema *et al.*, 2005).

Transthyretin (also called thyroxine-binding prealbumin) is a tetramer hormone, which is the thyroid hormone T4 carrier in serum and primary cerebrospinal fluid, thereby transporting up to 30% of serum thyroxine in normal individuals (Azad, 2011; Robbins, 2002). It also carries circulatory vitamin A by interacting with retinol-binding protein in a mutually enhancing fashion (Landers *et al.*, 2013). The sodium iodide symporter (NIS) (solute carrier family 5 SLC5A), is an integral plasma membrane protein which mediates the active transport of iodine anion (I^-) into the thyroid follicular cells as the critical first step for thyroid hormone biosynthesis (Smanik *et al.*, 1996). Although the mechanisms are not clear yet, it is thought to function through its translocation to the cell membrane to ensure correct insertion of iodine ion (Kogai *et al.*, 2012). It is also thought to be regulated by the p53 family proteins (Guerrieri *et al.*, 2013). It is expressed primarily in thyroid tissues, as well as in breast, colon and ovary (Smanik *et al.*, 1996). The thyroid hormones T3 and T4 are the only iodine-containing hormones among vertebrates, and the thyroid gland function depends on its adequate supply (i.e. sufficient dietary intake of I^- and proper NIS function) (Reinwein *et al.*, 1981). NIS-mediated radioiodine accumulation is reduced in most thyroid cancers due to decreased NIS expression/function, but it is enough for the radioiodide ablation to be effective in the majority of cases (Caillou *et al.*, 1998; Castro *et al.*, 2001; Patel *et al.*, 2002).

1.3.6. Transcription factors

The paired box gene 8 (PAX8) is a member of the thyroid-specific transcription factor (PAX) gene family which plays a critical role in the formation of tissues and organs during embryonic development and in maintaining the normal function of certain cells including thyroid follicular cells after birth (Ruiz-Llorente *et al.,* 2012). The protein is expressed in normal and neoplastic thyroid follicular epithelium and only a few other tissues. It has been demonstrated to be a critical regulator required for proper development and differentiation of thyroid follicular cells. Accordingly, the PAX genes provide instructions for making proteins that attach to specific areas of the DNA, and help in controlling their gene expression (Grasberger *et al.,* 2005). Thus, during embryonic development, as well as after birth, the PAX8 protein is believed to activate the genes that affect the formation of the kidney and the thyroid gland. This occurs through the regulation of several genes involved in the thyroid hormone production. PAX8 staining is thought to be useful in diagnosis of spindled or squamoid tumours of the neck (Bishop *et al.*, 2011).

Although the role of PAX8 in regulating genes responsible for thyroid differentiation has been well characterized, its involvement in cell survival and proliferation remains unclear. However, it is thought that that PAX8 is required for differentiated epithelial cell survival and can modulate their proliferation rate (Di Palma *et al.,* 2013). Furthermore, PAX8-specific silencing has been shown to induce apoptosis through a p53-dependent pathway that involves caspase-3 activation and cleavage of poly (ADP) ribose polymerase (Di Palma *et al.*, 2013). The p53 gene encodes a transcription factor that regulates cell growth, proliferation, cell cycle apoptosis and DNA repair (Farid *et al.*, 1994). Besides, although it was previously thought that p53 mutations are rarely found in thyroid cancer, recent evidence indicates that they may also be involved in early stages of the disease (Fagin *et al.*, 1996). The PAX8 gene may regulate cell

survival of differentiated epithelial cells through the transcriptional regulation of tumour protein p53-inducible nuclear protein 1 (tp53inp1), which is a positive regulator of p53-dependent cell cycle arrest and apoptosis. The tp53inp1 is a direct target of the PAX8 (Okamura *et al.*, 2001; Tomasini *et al.*, 2002) and has a crucial role in cell cycle control and in p53-mediated apoptosis (Di Palma *et al.*, 2013; Okamura *et al.*, 2001).

Besides, a PAX8-PPARγ fusion gene is common in FTC and some benign follicular adenomas (Dwight *et al.*, 2003; Eberhardt *et al.*, 2010; Kroll *et al.*, 2000). It is unclear how it exerts its oncogenic effect but it is believed that PAX8/PPARγ has a negative dominant effect on the wild type tumour suppressor effect of PPARγ. Although the mechanism here involves major gene rearrangements, it is thought that point SNPs/mutations might be another mechanism of loss or gain of function and contribution to thyroid cancer carcinogenesis (Vasko *et al.*, 2005; Wang *et al.*, 2007; Xing *et al.*, 2013; Xing *et al.*, 2005).

1.4. Thyroid cancers

1.4.1. Prevalence of thyroid cancer

The incidence of thyroid cancer has increased in many parts of the world over the last 3 decades, with a reported 2.4-fold increase ($P<0.001$) in the United States (Jemal *et al.*, 2010a; Jemal *et al.*, 2004; Jemal *et al.*, 2010b). However, the annual incidence varies considerably between different populations (Jemal *et al.*, 2010a). The reason for this significant increase is unclear, but it is at least partially due to improved diagnostic procedures providing clearer data on the prevalence of the disease (Davies *et al.*, 2006; Ward *et al.*, 2010).

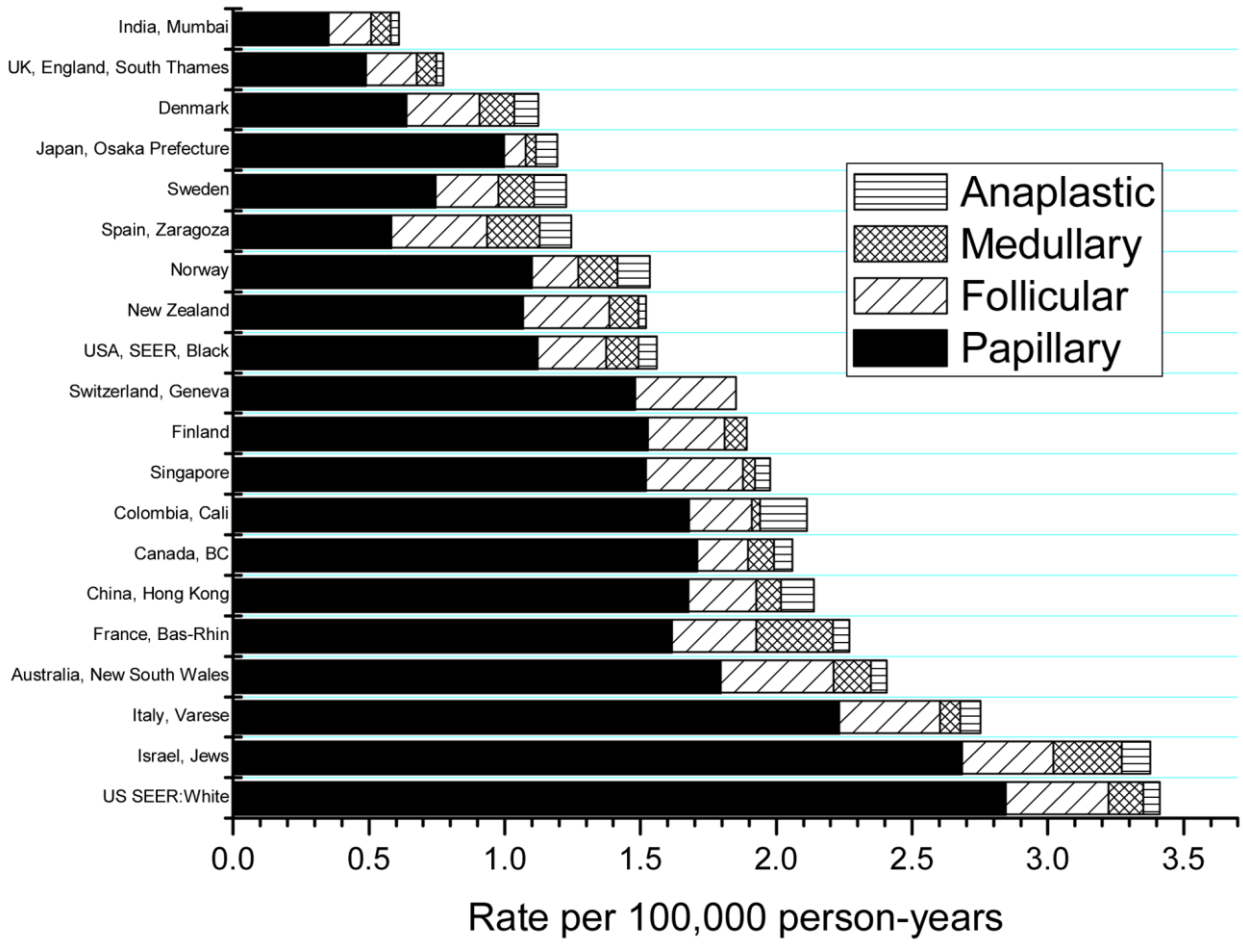


Figure 6: Rates for males from selected populations for the time period 1998–2002. Incidence rates of thyroid cancer for males (per 100,000 person-years) age standardized to the world population for selected populations for the time period 1998–2002 (figure adopted from Kilfoy, 2009) [78].

It has been shown that the age-adjusted international thyroid cancer incidence rates from 1998–2002 varied 5-fold for males and nearly 10-fold for females by geographic region .This variation was observed for every continent except Africa, in which the incidence rates were generally low (Kilfoy *et al.*, 2009). Among the Europeans, Lithuania, Italy and Austria appear have the highest, while Greece Bosnia-Herzegovina and Albania display the lowest incidence rate (Figure 8).

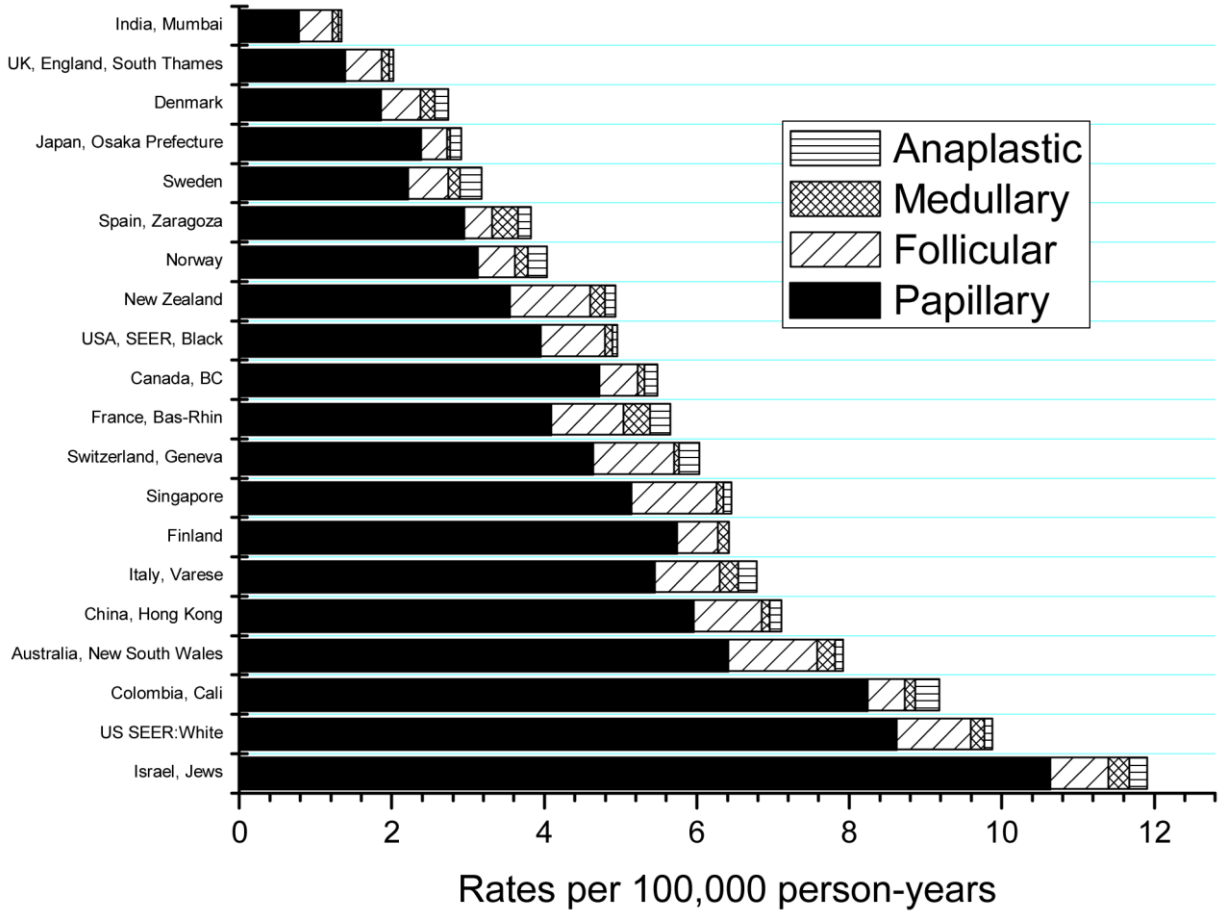


Figure 7: Rate for females from selected populations for the time period 1998–2002. Incidence rates of thyroid cancer for females (per 100,000 person-years) age standardized to the world population for selected populations for the time period 1998–2002 adopted from Kilfoy (Kilfoy *et al.*, 2009)

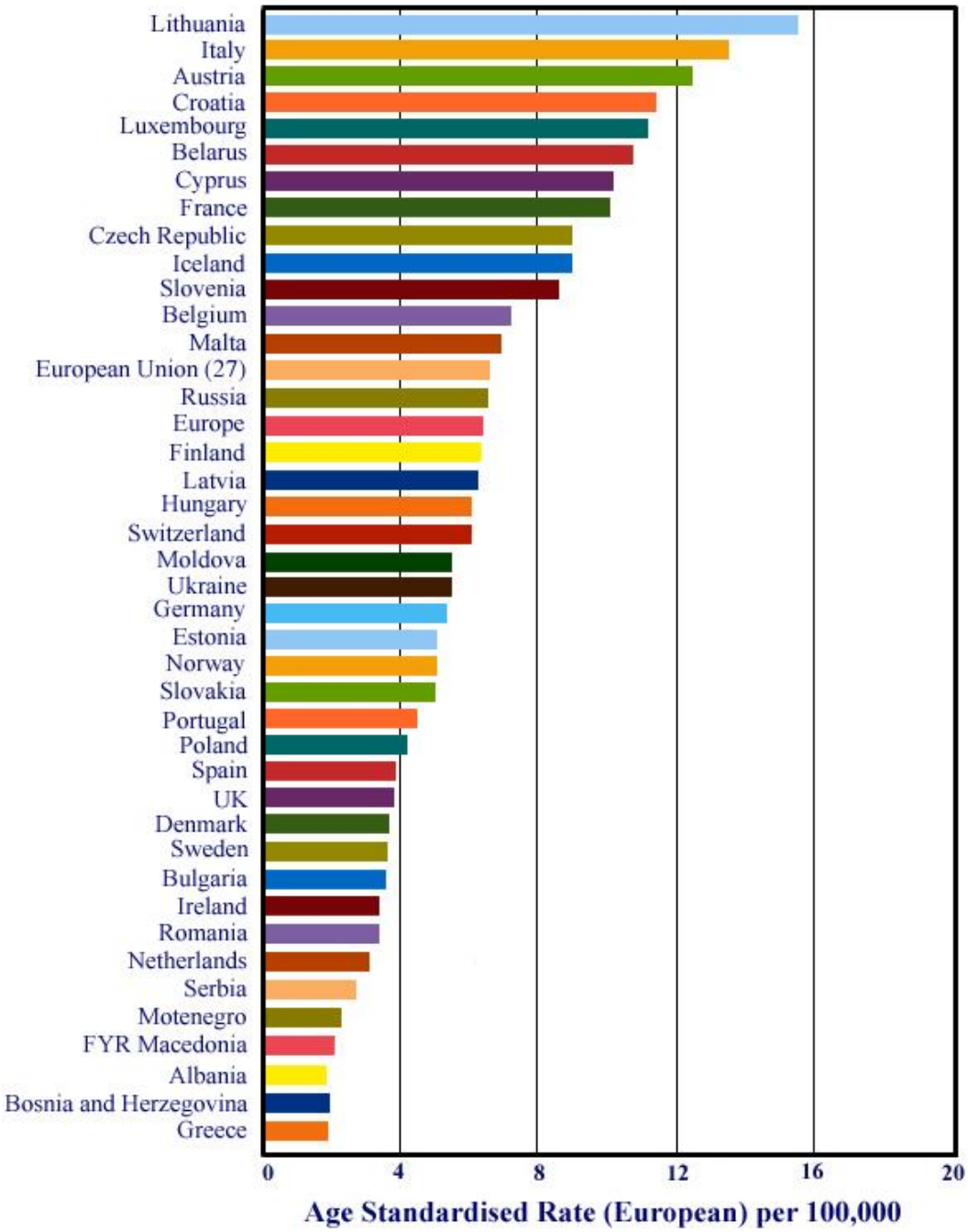


Figure 8. The figure displays the estimated incidence of cancer in Europe, adopted from Ferlay et al., 2013 (Ferlay *et al.*, 2013) and Bray et al., 2013 (Bray *et al.*, 2013).

In Saudi Arabia, thyroid cancer ranks number four among the ten most common cancers, preceded only by breast cancer, colorectal carcinoma and non-Hodgkin lymphoma (Saudi Cancer Registry, Figure 10). Among malignant tumours in females, it ranks number two, preceded only by breast cancer. Based on the 2006 National

Cancer Registry in Saudi Arabia, the age-standardized incidence rate (ASR) for thyroid cancer was 6.5 and 2.1 per 100,000 individuals in females and males, respectively (Figure 9). A comparison with selected countries worldwide showed a relatively high ASR in Saudi Arabia, albeit lower than some neighbouring countries and international rates (Figures 6 and 7). Although there might be differences in the methods and accuracy of these registries, they nonetheless indicate that incidence of thyroid cancer is undoubtedly on the rise in many countries around the world.

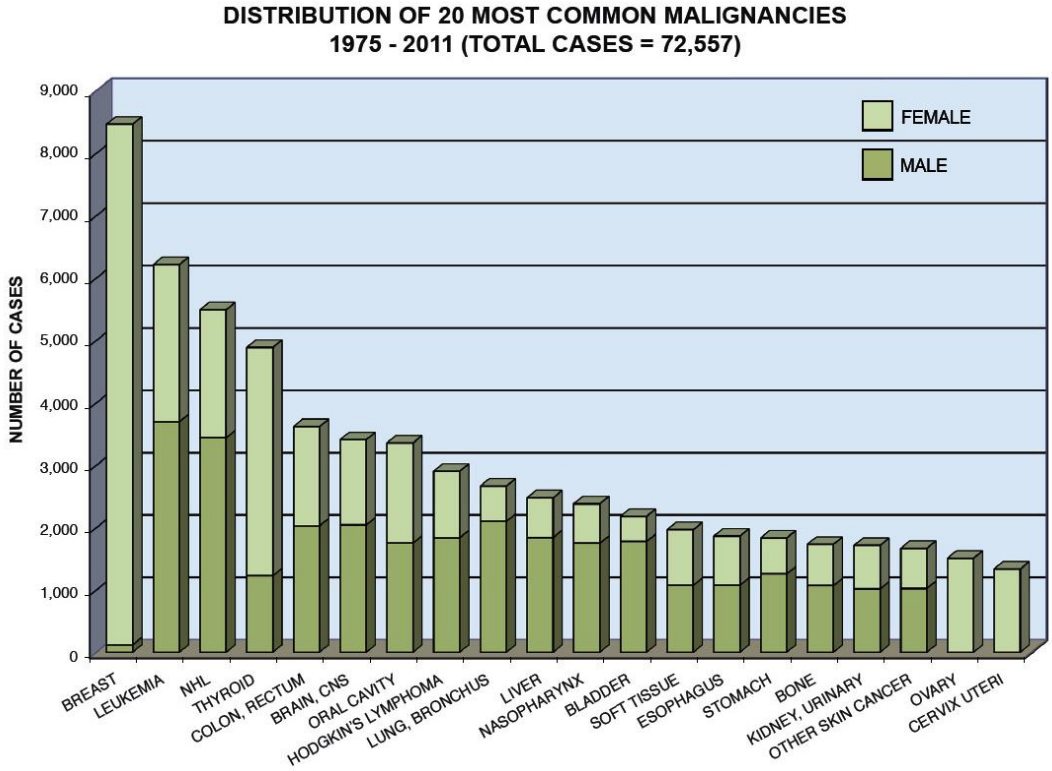


Figure 9: Distribution of the 20 most common malignancies in Saudi Arabia based on the KFSH and RC Tumour Registry, 2011

1.4.2. Types of thyroid cancers

Thyroid cancer is the most common endocrine malignancy. It accounts for about 1% of all new cases of cancer each year (Jemal *et al.*, 2010a; Jemal *et al.*, 2010b). Only about 5% of the growths in the thyroid gland (thyroid nodules) are cancerous (Welker *et al.*, 2003). It is classified into differentiated thyroid cancer (DTC) and undifferentiated or

anaplastic (ATC) thyroid cancer. The term "undifferentiated" is derived from the fact that the cancer cells do not look very much like normal thyroid tissue cells under the microscope. DTC represents more than 90% of cases and is further sub-classified into papillary (PTC) and follicular (FTC) subtypes. PTC and FTC account for approximately 80% and 15% of all cases, respectively (Greco *et al.*, 2009b). Apart from the follicular types, medullary thyroid cancer also exists as a rare subtype that originates from the parafollicular (C) cells. C cells are parafollicular cells in the thyroid gland located adjacent to the thyroid follicles and residing in the connective tissue that produce and secrete calcitonin.

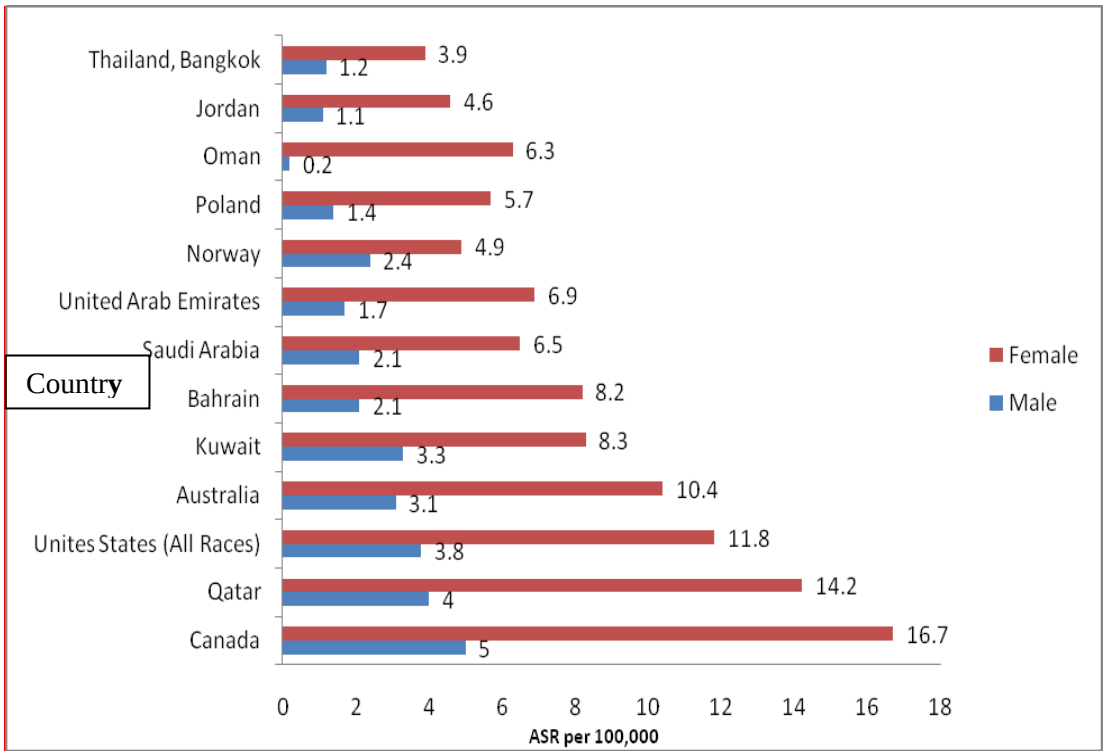


Figure 10. Comparison of age-standardized incidence rate for thyroid cancer among Saudis with that in selected countries. Data was plotted based on Saudi cancer registry 2006.

1.4.2.1. Papillary carcinoma

About 80% of the thyroid cancers are papillary carcinomas (also called papillary adenocarcinomas). These carcinomas typically grow very slowly, usually developing only in one lobe of the thyroid gland. Although they develop slowly, papillary carcinomas often spread to the lymph nodes in the neck. Several different subtypes of

papillary carcinoma can be recognized under the microscope. Other variants of papillary carcinoma (columnar, tall cell, diffuse sclerosis) are not as common and tend to grow and spread more quickly.

1.4.2.2. Follicular carcinoma

Follicular carcinoma (also sometimes called follicular cancer or follicular adenocarcinoma) is the next most common type of thyroid cancer, constituting approximately 10% of the thyroid cancers. Unlike papillary carcinoma, follicular carcinomas do not spread to lymph nodes and are usually retained in the thyroid gland. However, some can spread to other parts of the body, such as the lungs or bones (Ogawa *et al.*, 2006). It is thought that the prognosis for follicular carcinoma is not quite as good as that for papillary carcinoma. Nonetheless, it is still very good in most cases. Hürthle cell carcinoma, also known as oxyphil cell carcinoma, is a form of follicular carcinoma, which accounts for about 4% of the thyroid cancers (Kojic *et al.*, 2012). Because it is harder to find and treat (i.e. does not absorb radioactive iodine well), the prognosis may not be as good as that for typical follicular carcinoma, (Goffredo *et al.*, 2012).

1.4.2.3. Other types of carcinomas

The other types of carcinomas include the medullary thyroid carcinoma (MTC), anaplastic carcinoma (ATC), thyroid lymphoma and thyroid sarcoma. The MTC accounts for about 5% of thyroid cancers. It develops from the C cells of the thyroid gland, but may spread to lymph nodes, the lungs, or liver even before a thyroid nodule is discovered or a screening test is performed. It is an aggressive cancer that rapidly invades the neck, and is very hard to treat. There are 2 types of MTC: the sporadic and familial MTCs. The sporadic form, which occurs in about 80% of the MTC cases, is not hereditary and is found mostly in older adults and involves one thyroid lobe. On the other hand, the familial MTCs frequently develop during childhood or early adulthood,

and are often associated with an increased risk of other types of tumours. MTCs usually produce calcitonin and carcino-embryonic antigen, a protein that is produced in certain cancers, such as colorectal cancer (Heemstra *et al.*, 2008b; Namo Cury *et al.*, 2008).

Anaplastic carcinoma (also defined as undifferentiated carcinoma) is a rare form of thyroid cancer, which is believed to develop from an existing papillary or follicular cancer. This type of carcinoma makes up about 2% of all thyroid cancers.

Lymphomas are cancers that develop from lymphocytes, but very uncommon in the thyroid gland. Lymphocytes, the main cell type of the immune system, are pea-sized collections of immune cells scattered throughout the body, including the thyroid gland, but the cells are mostly found in lymph nodes. Additionally, thyroid sarcomas are rare cancers which originate in the supporting cells of the thyroid. They are often aggressive and hard to treat (Gereben *et al.*, 2008).

1.4.3. Diagnosis of thyroid cancers

Diagnosis of thyroid cancer is usually performed by fine needle biopsy (FNB), blood tests and ultrasound imaging. After a nodule is found during a physical examination, ultrasound is performed to confirm the presence of the nodule and to assess the status of the whole gland. Determination of the thyroid stimulating hormone (TSH) levels and presence of anti-thyroid antibodies will assist in deciding whether or not there is a functional thyroid disease, such as Hashimoto's thyroiditis, which is known to cause benign nodular goitre. One approach used to determine whether or not the nodule is malignant is use of the FNB, which has been described as the most cost-effective, sensitive and accurate test. This procedure or ultrasound-guided fine needle aspiration (FNA) usually yields sufficient thyroid cells to assess the risk of malignancy. However, in some cases, the suspected nodule may need to be removed surgically for pathological examination, or rarely, a biopsy is done with a large cutting needle, in order to procure a piece of nodule capsule.

Blood or imaging tests may be done prior to or in lieu of a biopsy. The possibility of a nodule secreting thyroid hormone and being less likely to be cancer or hypothyroidism is investigated by measuring the TSH level as well as the thyroid hormones T4 and T3. Tests for serum thyroid autoantibodies are sometimes done since these may indicate autoimmune thyroid disease. The blood assays may be accompanied by ultrasound imaging of the nodule to determine the position, size and texture and to determine whether or not the nodule is cystic. Hypoechoic, irregular borders, microcalcifications, or very high levels of blood flow within the nodule constitute suspicious findings in a nodule. On the other hand, hyperechoic, comet tail artefacts from colloid, no blood flow in the nodule and a halo, or smooth border make up less suspicious findings in benign nodules. Sometimes, technetium (Tc) or radioactive iodine (RAI) imaging of the thyroid is performed. A ^{123}I scan showing a "hot" nodule, accompanied by a lower than normal TSH strongly indicates that the nodule is not cancerous (Welker *et al.*, 2003).

1.4.4. Thyroid cancer therapy

The current strategies of cancer management include primarily surgery and radiation. Accordingly, surgery may be performed as a prophylactic, diagnostic, staging, curative, palliative or reconstitutive procedure. Once the cancer has been confirmed, the management will depend on the type of the disease. MTC is treated by total thyroidectomy and lymph node dissection (Kloos *et al.*, 2009). On the other hand, anaplastic thyroid cancer is a highly malignant tumour. However, it has very short survival (Alvarez *et al.*, 2000), and surgery represents the best cure for it (Cooper *et al.*, 2009). For DTC, the standard management consists of total thyroidectomy with or without lymph node dissection, followed by thyroid remnant ablation with radioactive iodine (Cooper *et al.*, 2009). The radiation is intended to remove any residual thyroid tissue or cancer following surgery. After surgery and radioactive iodine ablation, the patient usually becomes totally athyrotic (no remnant thyroid tissue) and will be

dependent totally on exogenous intake of thyroxine. Thyroid hormone therapy is given to replace the hormones (T4 and T3) and suppress tumour growth (Biondi *et al.*, 2010; Cooper *et al.*, 2009). Studies have described an association of decreased risk of recurrence with high dose of thyroid hormone without inducing significant hyperthyroidism (Cooper *et al.*, 1998). The basis of this beneficiary effect of TSH suppression is the fact that TSH promotes growth by its trophic action on thyroid tissue/tumours. While the level of TSH suppression depends on the risk category of the cancer, it is generally recommended that TSH is suppressed to <0.1 mU/l (normal range =0.5 - 4.5 mU/I, based on American Association of Clinical Endocrinologists) (Cobi *et al.*, 2001), at least for the first 5 years after the diagnosis of DTC (Cooper *et al.*, 2009; Pacini *et al.*, 2012) .

Although it is generally agreed that patients with persistent or recurrent disease ought to have their TSH suppressed to this undetectable level of <0.01 mU/l), it is still unclear to what level or for how long this should be done (Cooper *et al.*, 2009; Pacini *et al.*, 2012). On the other hand, patients who are apparently cured but have had history of more aggressive DTC are generally maintained on T4 therapy. This is done in order to keep TSH levels undetectable or close to the lower limit of the normal levels. Patients with smaller thyroid cancers who are likely to have lower risks of recurrence are often advised to have TSH close to the lower limit of the normal range, but not totally suppressed (Biondi *et al.*, 2010). For both bone and heart diseases, the levels of TSH suppression and optimal thyroid hormone dosing will also depend on the age of the patient, and associated risk factors (Biondi *et al.*, 2010). The thyroglobulin (Tg) level is a useful parameter in guiding therapy. If Tg is detected, this would indicate the likelihood of the patient having persistent, recurrent or metastatic disease. In this case, the patient needs to complete the TSH suppression. On the other hand, if the Tg is not detected, especially at the evaluation stage when the patient is taken off thyroid

hormones, this will point to the likelihood of the patient being cured and not requiring complete suppression (Biondi *et al.*, 2010; Cooper *et al.*, 2009). Other considerations for the level of TSH suppression include the presence or risk of osteoporosis and cardiac diseases such as atrial fibrillation. In the presence of these conditions, TSH suppression should be kept less stringent (Biondi *et al.*, 2010; Cooper *et al.*, 2009; Shargorodsky *et al.*, 2006).

Levothyroxine is a synthetic derivative of thyroxine, many generic versions of which have been approved by the U.S. Food and Drug Administration (FDA). According to the American Association of Clinical Endocrinologists, patients should be maintained on the same brand of levothyroxine, since small changes in its regimen can lead to significant changes in TSH serum concentrations (Michaud *et al.*, 1986). Changing the brand requires retesting of TSH in 6 weeks and adjustment of the dose accordingly, particularly since different formulations display different bioavailability.

1.4.5. Prognosis of thyroid cancer therapy

The prognosis of thyroid cancer is related to many factors, such as type of cancer, age of patient, tumour size and disease stage. PTC shows the best outcome, whereby 90% of patients remain alive after the diagnosis, followed by FTC (80%) and MTC, with ATC showing the poorest outcome (Gilliland *et al.*, 1997; Zarebczan *et al.*, 2010). The unusual forms of papillary carcinoma and follicular types have the same prognosis and follow the same treatment (Witt *et al.*, 2013). Because medullary cancer does not absorb radioactive I_2, the outlook is not quite as good as that for differentiated thyroid cancers. Furthermore, most of the time, papillary carcinomas can be successfully treated and are rarely fatal (Mallick, 2010).

1.5. Gene polymorphism, thyroid cancer manifestation and management

1.5.1. Gene polymorphism and thyroid gland disorders

Thyroid disorders constitute a heterogeneous group of abnormalities that are linked to mutations found in several genes that are responsible for several aspects of the thyroid function. These genes include those coding for protein that are involved in the development of the thyroid, such as thyroid transcription factor-1 and -2 (TTF-1 and TTF-2), paired box gene 8 (PAX8). The mutations could also occur in one of the genes encoding the proteins involved in TH biosynthesis, such as the thyroglobulin (TG), thyroperoxidase (TPO), hydrogen peroxide-generating system (DUOX2) (Smanik *et al.*, 1994), NIS (Tonacchera *et al.*, 2002b), TSH and TSH receptor (TSHR)(Rivolta *et al.*, 2005).

Unlike other types of cancers, information on the role of gene polymorphisms in thyroid cancer is still scanty. Probably the most critically studied gene in this regard is the TSHR. The TSH controls thyroid growth and hormone secretion through binding to this G protein-coupled receptor and the production of cyclic AMP (cAMP). It is thought that mutant TSHRs cause hyperactive thyroid nodules (Kohrle, 2000). Somatic gain-of-function mutations resulting in a phenotype of toxic adenoma, toxic multinodular goitre and congenital hyperthyroidism have also been described in the TSHR gene (Krohn *et al.*, 2002). In contrast, loss-of-function mutations have been associated with TSH resistance and congenital hypothyroidism (Sunthornthepvarakui *et al.*, 1995). Another mutation of the TSHR gene (Leu512Arg) found in the autonomously functioning papillary carcinoma has been associated with constitutive activation of the cAMP signal transduction pathway, causing thyrotoxicosis and a hot thyroid nodule in an autonomously functioning papillary carcinoma (Barbaro *et al.*, 2006). However, while somatic gain-of-function TSHR mutations are present in

hyperfunctional follicular adenomas and goitres, they may be absent in non-functioning thyroid lesions (Bianco *et al.*, 2006).

Besides, several microsatellite markers and intronic SNPs in the TSHR gene, such as, Asp727Glu have been associated with lower plasma TSH levels in healthy blood donors, without having any effect on free thyroxine (FT4) (Hansen *et al.*, 2007; Peeters *et al.*, 2003b). However, their contribution to the genetic variance is thought to be small. One study has suggested that the variant triggers an increased cAMP response of the receptor to TSH (Leonard *et al.*, 2001). Another study pointed to an association of TSHR mutations with iodine supply (Curcio-Morelli *et al.*, 2003), while in non-diabetic elderly men, variation in serum thyroid levels and the Asp727Glu have been associated with relative insulin resistance (Peeters *et al.*, 2007). Moreover, some investigators were unable to replicate these findings, possibly suggesting that the Asp727Glu is linked to other polymorphisms in a haplotype form (Dechairo *et al.*, 2005). Data on the response of the Pro52Thr variant of TSHR-β to TSH stimulation in healthy blood donors is also somewhat conflicting, with some studies failing to find any association with changes in serum TSH or iodothyronine levels. This has been attributed to subtle effects or low frequency of the Thr variant (Peeters *et al.*, 2003b).

Individuals carrying a heterozygous loss-of-function mutation seem to have a dominant transmission of partial TSH resistance, possibly due to intracellular entrapment and reduced maturation of the wild-type TSHR by inactive mutants (Calebiro *et al.*, 2005). Polymorphisms leading to a relative loss of function are likely to have an effect via this mechanism, and could possibly account for some of the so-called 'euthyroid outliers' with elevated TSH determinations (Calebiro *et al.*, 2005). However, only a few TSHR polymorphisms resulting in amino acid substitutions have been identified. Two of these, Asp36His and Pro52Thr, are located in the extracellular and another, Asp727Glu, in the intracellular domain of the receptor (Leonard *et al.*, 2001). In one study, the T

variant of the Pro52Thr was found to be significantly higher in hyperthyrotropinemic infants (p=0.03) compared to newborn controls (Teofoli *et al.*, 2007).

Another gene of interest is the thyroid hormone receptor-beta (THRβ). One of the THRβ mutations, the -G9A, has been significantly associated with higher serum TSH (p = 0.01) in a Danish twins study (Sorensen *et al.*, 2008). However, some studies concluded that the THRB mutations are not a relevant mechanism for human thyroid carcinogenesis (Rocha *et al.*, 2007).

Protein changes in the thyroglobulin, such as Glu2511Arg, are also thought to confer an increased risk for benign thyroid disorders (Hansen *et al.*, 2007; Pitt-Rivers *et al.*, 1955). It was argued that, although the risk of non-medullary thyroid cancer associated with the R-allele of this SNP is modest, its high prevalence in the general population suggests that it may nonetheless contribute significantly to the incidence of non-medullary thyroid cancer (Hansen *et al.*, 2007). Other SNPs include two cysteine substitutions, Cys1263Arg and Cys1995Ser, which have been associated with a defect in intracellular transport of thyroglobulin in congenital goitre, a variant of adenomatous goitre, in a familial study (Gabriel *et al.*, 1999).

Other genes discussed with respect to thyroid cancer manifestation include the phosphodiesterase 8B (PDE8B), the receptor tyrosine kinase (RET) (Arnaud-Lopez *et al.*, 2008; Krohn *et al.*, 2002; Pitt-Rivers *et al.*, 1955; Sykiotis *et al.*, 2003), cytochrome P450 1A1 (CYP1A1), glutathione S-transferase (GST) and glutathione S-transferase mu 1 (GSTMI) (Siraj *et al.*, 2008b). Specifically, Siraj et al described an association of the cytochrome P450 1A1, GST and GSTMI with PTC risk in one study (Siraj *et al.*, 2008b), and of two RAD52 variants also with the disease in another study in the Saudi population (Siraj *et al.*, 2008a). Another study found an association of the N-RAS gene with an early stage and lower incidence of extrathyroidal extension, and BRAF with metastasis and poor disease-free survival in PTCs in the Middle Eastern population

(Abubaker *et al.*, 2008). The PDE8B and the receptor tyrosine kinase (RET) (Arnaud-Lopez *et al.*, 2008; Krohn *et al.*, 2002; Sykiotis *et al.*, 2003), both of which are related to papillary thyroid cancer risk (Pitt-Rivers *et al.*, 1955). Thus, a G691S of the RET proto-oncogene was described in a mixed medullary and follicular cell carcinoma of the thyroid (Sykiotis *et al.*, 2003). The observation of a strong association of the PDE8B with circulating TSH levels in Sardinians has led to the suggestion of a primary effect on cAMP levels in the thyroid (Arnaud-Lopez *et al.*, 2008). Accordingly, this would affect the production of T4 and T3 as well as the feedback to alter TSH release by the pituitary.

The expression and activity of iodothyronine deiodinases are altered in different types of neoplasia (Piekielko-Witkowska *et al.*, 2011). However, to date no inactivating mutations have been described in the THSR or deiodinases. Recent studies show that increased DIO3 expression may lead to enhanced tumour proliferation in support of the idea of the deiodinases having the potential to affect progression of cancer, leading to the notion of their potential use as cancer markers and modulators of tumour progression (Piekielko-Witkowska *et al.*, 2011).

1.5.2. Effect of gene polymorphism on thyroid function and management of thyroid cancer

The importance of gene polymorphisms in disease is particularly evident in various forms of cancer. DNA damage is an important mechanism in carcinogenesis. Hence, genes related to maintaining genomic integrity such as the p53 gene (Albrechtsen *et al.*, 1999; Hu, 2009) and Upstream stimulatory factor (USF-1) (Baron *et al.*, 2012), may influence cancer risk (Neta *et al.*, 2011). Multiple associated gene polymorphisms have been described for malignancies such as breast (Qiu *et al.*, 2008; Rastelli *et al.*, 2008; Tan *et al.*, 2008), lung (Li *et al.*, 2008a; Risch *et al.*, 2008), gastric (Gao *et al.*, 2008),

leukemic (Ansari *et al.*, 2007), lymphoma (Davidsen *et al.*, 2008) and many other forms of the disease (Dong *et al.*, 2008; Gao *et al.*, 2008; Li *et al.*, 2008a; Zaridze, 2008).

Despite the high heritability shown in thyroid cancer, the causative genes for this disease remain largely unknown. Recently however, at least five SNPs (rs965513, rs944289, rs966423, rs2439302, and rs116909374) have been associated with papillary thyroid carcinoma (PTC) by genome-wide association studies (GWAS) (de la Chapelle, 2013). Thus a GWAS by Gudmundsson et al described variants on 9q21.33 (rs965513) and 14q13.3 (rs944289) that predisposed individuals to thyroid cancer in European populations (Gudmundsson *et al.*, 2009). The gene nearest to the 9q22.33 locus was forkhead box E1 gene (FOXE1; TTF2), while NK2 homeobox 1 gene (NKX2-1; TTF1) was among the genes located at the 14q13.3 locus, both of which were associated with low concentrations of TSH, and the 9q22.33 allele was further associated with low concentration of T4 and high concentration of T3(Gudmundsson *et al.*, 2009). A number of studies have in the meantime supported a role for the FOXE1 gene in thyroid and various forms of cancers (Carre *et al.*, 2014; Penna-Martinez *et al.*, 2014; Srichomthong *et al.*, 2013; Zhu *et al.*, 2014). The rs966423 (2q35) and rs2439302 (8p12) were later also linked to PTC in another GWAS study by Gudmundsson and colleagues (Gudmundsson *et al.*, 2012). The association of the rs2439302 was linked to the expression of neuregulin 1 gene (NRG1), which encodes the signalling protein neuregulin 1, in blood (Gudmundsson *et al.*, 2012). Association of the SNPs, rs944289, rs965513, rs966423 and rs2439302, with PTC was replicated in a Chinese cohort (Wang *et al.*, 2013a) and through the TCUKIN study (Jones *et al.*, 2012). However, a study by Liyanarachchi suggested that each of these variants showed highly significant, but moderate to low disease risk (Liyanarachchi *et al.*, 2013). These authors concluded that while the risk of having PTC increased with the genetic risk score, the predictive power of their cumulative effect was moderately high and clinical use needed further

verification. Furthermore, a linkage study by He suggested that SRGAP1 is a candidate gene for PTC susceptibility, with low-penetrance possibly as a modifier gene (He *et al.*, 2013a). Another familial study identified one rare mutation on chr4q32 in a pedigree displaying non-medullary thyroid carcinoma. This mutation is localized to a long-range enhancer element whose ability to bind the transcription factors POU2F and YY1 was found to be significantly impaired. This was linked to the downregulation of an enhancer RNA (eRNA) that was transcribed in thyroid tissue from this region (He *et al.*, 2013b). Besides, some population studies have also indicated that microRNA variants are important factors underlying the process of carcinogenesis, such as papillary thyroid carcinoma (de la Chapelle, 2013; Wojcicka *et al.*, 2013). Several miRNAs including miR-146b-5p, miR-221-3p, miR-7-3p, miR-551b-3p, miR-486-3p, and miR-144-3p, have been implicated in the disease, apparently exhibiting isoforms of variable length and potentially distinct function in thyroid tumorigenesis (Swierniak *et al.*, 2013).

1.5.3. Gene polymorphism and patient response to thyroid cancer management

Drug metabolizing enzymes are generally classified as belonging to either phase I or phase II types of reaction. Phase I reactions include oxidation, reduction, hydrolysis, cyclization and decyclization reactions, while phase II enzymes engage in conjugation reactions (e.g. involving glucuronic acid, sulfation, glutathione or amino acids, which are detoxification reactions and involve interactions of polar functional groups of phase I metabolites) may explain part of the observed inter-individual variations in pharmacokinetics and pharmacodynamics of anticancer drugs (van Schaik, 2005). Recently, it has also been recognized that several genetic variations in metabolizing enzymes influence circulating thyroid hormone levels. However, such variants have not been extensively studied (Panicker *et al.*, 2008).

Serum thyroid parameters show substantial inter-individual variability in healthy subjects, whereas the intra-individual differences lie within a narrow range (Andersen *et*

al., 2002). Since the window between drug toxicity and suboptimal therapy for cancer treatment is often narrow, it can be assumed that inter-individual variations in drug metabolism complicate therapy often as a result of genetic variations, in addition to environmental factors including food or iodine intake. For example, in a study of thyroid function tests involving Danish twins, genetic heritability accounted for about 65% of the variation in serum levels of TSH, FT4 and FT3 (Hansen *et al.*, 2004). In another study from a Mexican-American population, heritability accounted for 26 to 65% of the variation in serum thyroid stimulating hormone (TSH), free T4 (FT4) and free T3 levels (Samollow *et al.*, 2004). Several other studies have described polymorphisms in the thyroid hormone pathway genes that alter serum thyroid function. In this study, the primary focus is on thyroid cancer, in which the role of gene polymorphisms in the thyroxine dose requirements for TSH suppression will be investigated and their association with the risk of thyroid cancer will be evaluated. Details of these two main objectives will follow.

1.5.3.1. Polymorphisms of thyroxine signalling pathway genes and disease management

TSHβ transcription is crucial for the synthesis of functional TSH, which has a growth-stimulating effect on normal thyroid and thyroid cancer cells. This effect is mediated through the TSH receptor located on the cell surface which mediates its action via 2 pathways, the $G_{S\alpha}$-adenylyl cyclase-cyclic AMP pathway (Kondo *et al.*, 2006) and the G_q- or G_{11}-mediated phospholipase inositol triphosphate pathway (Boutin *et al.*, 2011). In a study by Franco et al, thyroid hormone receptor transgenic mice developed FTC spontaneously, but failed to do so when they were crossed with TSH receptor -/- mice, indicating the importance of TSH pathway for the development of thyroid cancer (Franco *et al.*, 2011; Lu *et al.*, 2010). In another study, a thyroid-specific BRAFV600E knock-in mouse developed an aggressive PTC. However, when this mouse model was

crossed with TSH receptor-/- mouse, it failed to develop PTC. It is now acknowledged that genetic variations in the TSHR and transthyretin play an important role in intra-individual variation in TH bioactivity (Bartalena *et al.*, 1992; Curtis *et al.*, 1994; Episkopou *et al.*, 1993; Haymart *et al.*, 2008; Moses *et al.*, 1990; Sorensen *et al.*, 2008) So far, research has focused mainly on genetic variation in the TSHR and the deiodinases, and their association with thyroid parameters and/or clinical endpoints, such as insulin resistance.

1.5.3.2. Polymorphisms of the metabolic pathway genes and thyroid disease management

Since the major processes regulating thyroid hormone metabolism are deiodination, deamination and glucuronidation, as well as transmembrane transportation, functional mutations in genes regulating these processes are likely to influence the thyroid hormone requirement of the patient. The genes encoding the iodothyronine deiodinases are good candidates in this regard, since they affect circulating thyroid hormones. Polymorphisms described in DIO1 include the SNPs, Cys785Thr which appear to show varying effects on serum T3/rT3 ratio and DIO1 activity in healthy individuals (Peeters *et al.*, 2003b). The Cys785Thr seems to affect the serum T3/rT3 ratio (de Jong *et al.*, 2007; Peeters *et al.*, 2003a; Peeters *et al.*, 2003c; van der Deure *et al.*, 2008a), partly leading to decreased activity of DIO1 (Peeters *et al.*, 2006; Peeters *et al.*, 2003a). Also, in a study involving 156 healthy blood donors, two DIO1 SNPs located in the 3'-untranslated (3'-UTR) region, called DIO1a-C/T, DIO1b-A/G, were associated with changes in plasma T4, free T4, T3, rT3, and TSH levels (Peeters *et al.*, 2003b). Accordingly, the DIO1a-T was associated with a higher plasma rT3, a higher plasma rT3/T4, and a lower T3/rT3 ratio in a dose-dependent fashion, while the DIO1b-G allele was associated with lower plasma rT3/T4 and higher T3/rT3 ratios (Peeters *et al.*,

2003b). Both SNPs are located in the 3'-untranslated (3'-UTR) domain of the DIO1 gene, possibly pointing to some regulatory function on the gene or mRNA expression/maturation. Furthermore, some investigators suggested that the rs2235544 in the DIO1 gene can change deiodinase function, leading to an alteration in the balance of circulating freeT3 to free T4 (Panicker *et al.*, 2008). It can therefore be inferred that changes in the DIO1 gene are associated with alterations in serum thyroid hormone levels. Reduced DIO1 expression has also been reported in the non-thyroid syndrome and in several human neoplasias (Arnaldi *et al.*, 2005; Maia *et al.*, 2011a). However, some investigators could not establish similar associations for these SNPs with serum rT3 levels (Bianco *et al.*, 2002).

Unlike the DIO1, DIO2 mutations have been associated with various disorders, including insulin resistance, mental retardation in thyroid-deficient children and neurological disorders. Perhaps the most well-studied variant is the Thr92Ala that has been associated with decreased DIO2 enzymatic activity and an increased risk for type 2 diabetes mellitus (T2DM), obesity and insulin resistance (Canani *et al.*, 2005a; Grarup *et al.*, 2007; Maia *et al.*, 2007). However, one study reported an association of the Thr92Ala with increased rather than decreased insulin sensitivity in the Amish population (Mentuccia *et al.*, 2005). It was also suggested that this SNP predicts the T4 dose needed to achieve target TSH (thyrotropin) levels in thyroidectomized patients (Torlontano *et al.*, 2008). Thus, while no association between the Thr92Ala polymorphism and serum thyroid hormone levels has been actually observed in humans, some studies in euthyroid patients have suggested that patients with the Ala92Ala phenotype needed higher thyroxine doses to achieve TSH suppression (Torlontano *et al.*, 2008). No association has been described thus far for the Thr92Ala variant with thyroid hormone levels or thyroxine dose requirement in patients treated for DTC or Hashimoto thyroiditis (Heemstra *et al.*, 2008a). Data on the influence of the Thr92Ala

polymorphism on serum T3 level is conflicting, but the SNP has been associated with insulin resistance in different populations (Canani *et al.*, 2005b; Mentuccia *et al.*, 2005). However, other investigators have found no association in various populations (Guo *et al.*, 2004). The basis for this association is unclear. In one study, both DIO1 and DIO2 were under-expressed in nearly all papillary thyroid carcinomas, leading to the notion that this could not only help to better define a tumour signature for thyroid tumours, but may also become useful as targets for thyroid tumour treatment (Arnaldi *et al.*, 2005). One other interesting observation is the possible involvement of ORFa-Gly3Asp of the DIO2 gene in insulin resistance in various populations (Guo *et al.*, 2004). This SNP has been associated with the variations in the serum T3/T4 ratio in young compared to the elderly subjects (Peeters *et al.*, 2005). Also, observations of reduced DIO2 in muscle and thyroid homogenates of the Ala92 allele carriers has led to the speculation that the ultimate effects are a product of linkage with other polymorphisms, possibly in form of a haplotype (Canani *et al.*, 2005a). Intronic mutations of the DIO2 gene may affect the amount of T3 available and partly determine the overall risk of mental retardation in an iodine-deficient environment (Guo *et al.*, 2004). However, the role of the gene on neurological actions of the thyroid hormone remains unclear.

In contrast to DIO1 and DIO2, currently there is little information on the possible role of DIO3 gene polymorphism in the management of thyroid disorders. A study involving 156 healthy blood donors previously described some SNPs located in the 3'-UTR region of the DIO3, albeit with no definable relationship with plasma T4, free T4, T3, rT3, and TSH levels (Peeters *et al.*, 2003b).

Polymorphisms in Transthyretin may potentially influence the availability of the thyroid hormones at their various functional sites. Notably, the substitution of Thr with Met in Thr119Thr in the iodothyronine protein was identified in individuals with transient euthyroid hyperthyroxinemia (Alves *et al.*, 1993). Furthermore, while healthy carriers

of Met119 have shown normal serum thyroid hormone concentrations, it remains uncertain whether or not the T4 binding to TTR is increased in such individuals (Alves *et al.*, 1993).

1.5.3.3. Polymorphism of TSH and TSH receptor and thyroid disease management

Thyroid stimulating hormone (TSH) is a glycoprotein hormone secreted from thyrotrops of the anterior pituitary which regulates secretion of the thyroid hormone and therefore thyroid function. It is a heterodimer containing alpha (α-) and beta (β-) subunits. The α-subunit is common to all members of the glycoprotein hormone family including the luteinizing hormone, follicle-stimulating hormone, TSH, and human chorionic gonadotropin, while the β-subunits are unique for each hormone, thereby providing the specific immunogenic and hormonal functions (Magner, 1990; Pierce *et al.*, 1981; Shupnik *et al.*, 1989). Therefore, transcription of TSH β-subunit gene is crucial to the synthesis of functional TSH. The TSH subunit genes are negatively regulated at the transcriptional level by T3 and T4, while thyrotropin releasing hormone (TRH) and the hypothalamic hormone positively enhance TSH gene expression.

As stated previously, variations in the TSHR and transthyretin genes play an important role in intra-individual variations in thyroid hormone bioactivity (Bartalena *et al.*, 1992; Curtis *et al.*, 1994; Episkopou *et al.*, 1993; Haymart *et al.*, 2008; Moses *et al.*, 1990; Sorensen *et al.*, 2008). A homozygous missense mutation in the TSHβ gene due to single base substitution G>A at codon 85 leading to a change from glycine to arginine was recently described in a 15-month-old male of consanguineous parents with congenital hypothyroidism (Muthukrishnan *et al.*, 2010) and in three cases with similar phenotype from Japan (Muthukrishnan *et al.*, 2010). Apart from being implicated in thyroid cancer, mutations in TSHβ may cause a rare disease defined as congenital central hypothyroidism (CCH) (Partsch *et al.*, 2006). One such mutation described in the German population is the homozygous deletion, delta 313T, in which the TSH

serum levels remained unaltered, leading to a negative result in the neonatal TSH screening (Partsch *et al.*, 2006). Severe isolated TSH deficiency has also been reported in two children from a consanguineous Turkish family who were homozygous for a C>T (Q49X) transition at the nt 654 of the TSHβ subunit gene, leading to a protein change from glutamine (CAG) to a premature TAG stop codon at position 49. The resulting nascent peptide lacked the TSHβ subunit region essential for the dimerization with the α-subunit, thereby blocking the correct secretion of the mature TSH heterodimer (Vuissoz *et al.*, 2001). Free T3, free T4, and basal TSH levels were extremely low in both affected individuals, and TRH stimulation failed to increase serum TSH, but not pituitary prolactin, confirming isolated TSH deficiency(Vuissoz *et al.*, 2001). Also, mutations in the TSHR have been described in a study involving 139 children of consanguineous families with a non-syndromic autosomal recessively inherited non-goitrous congenital hypothyroidism (CHNG) phenotype (Cangul *et al.*, 2010). Besides, homozygous germline TSHR mutations were discovered in six families with no clinical consequences.

1.5.3.4. Transcription factor polymorphisms and thyroid disease

The PAX8 gene has recently been investigated as a useful marker of thyroid epithelial neoplasms (Schmitt *et al.*, 2010). In this regard, several mutations have been described with variable effects. Thereby, some mutations are thought to cause congenital hypothyroidism (Congdon *et al.*, 2001), while others gently reduce thyroid hormone levels, possibly as a result of the unusual small thyroid gland size in people with the PAX8 mutation(Komatsu *et al.*, 2001). For example, sequencing of the PAX8 gene in a girl suffering from thyroid dysgenesis revealed a maternally inherited heterozygous mutation in the promoter region (-456C>T) (Hermanns *et al.*, 2011). However, despite being well-characterized with respect to its role in regulating genes involved in thyroid

differentiation, the biological pathways controlled by this transcription factor remain largely unknown, due to lack of identification of its additional targets.

1.5.3.5. Thyroid transporter polymorphisms and thyroid disease

Currently, only a few studies are available in the literature pertaining to the role of thyroid transporter in thyroid disease. A study of the NIS expression showed lack of NIS mRNA expression in 5 of 19 papillary thyroid cancers, while it was expressed in all 15 except 1 follicular thyroid adenoma (Arturi *et al.*, 1998). The authors concluded that early detection of the loss of NIS gene expression in the primary cancer may provide useful therapeutic information in DTC patients (Arturi *et al.*, 1998). Another study showed a 93% decrease in NIS expression in thyroid carcinomas and 83% cold adenomas consisting of toxic adenomas and Graves thyroid tissues (Lazar *et al.*, 1999). In addition, the authors found an inverse relationship between the tumour stage and the NIS expression levels (Lazar *et al.*, 1999), leading to the conclusion that reduction in NIS gene expression occurs in most hypofunctional, benign and malignant thyroid tumours. This is thought to be triggered by DNA methylation in critical regulatory regions (Venkataraman *et al.*, 1999b). A Gly543Glu substitution in the NIS gene has also been described in 2 Japanese siblings with I^- transport defect, which was supported by the observations of minimal I_2 uptake in NIS constructs carrying this mutation in COS-7 cells (Kosugi *et al.*, 1998). Another study observed that, while the homozygous mutant NIS-272X (C1163A substitution, resulting in the TGA stop codon at nt 272) causes congenital hypothyroidism, the C272X was enough to maintain active thyroidal iodine-uptake and function (Pohlenz *et al.*, 1997). A G1530A transition of the NIS gene, resulting in a gly395arg substitution, has also been described in a large Hutterite family in which 18 children carried an autosomal recessive form of congenital hypothyroidism due to an iodide transport defect, confirming that this mutation was the

direct cause of the iodide transport defect in these patients (Couch *et al.*, 1985; Kosugi *et al.*, 1999).

The low radioiodide uptake due to a defect in iodine transport observed in the majority of non-functioning thyroid nodules has been attributed to the defect in the expression and therefore targeting of human NIS (hNIS) protein to the cell membrane (Tonacchera *et al.*, 2002b). In this gene, a silent polymorphism C544G was described in one non-functioning nodule among patients who were originally subjected to surgery for a solitary non-functioning thyroid nodule originating in an otherwise normal gland after the oral administration of a tracer dose of iodine-131 (^{131}I) (Tonacchera *et al.*, 2002b). Also, the oncogene BRAF(V600E) has been suggested as the most frequent genetic event and prognostic factor in PTC, correlating with a high risk of recurrences and less differentiated tumours due to the loss of NIS-mediated ^{131}I uptake(Riesco-Eizaguirre *et al.*, 2006).

1.5.4. Gene-gene and gene-environment interactions in thyroid disease management

Gene-gene and gene-environment interactions are important features of complex disease manifestations. At present only very little information is available in the literature regarding the role of gene-gene or gene-environment interactions in thyroid cancer. However, an interesting feature in this regard is the recent observation in some follicular thyroid cancer cells that PAX8 gene which resides on chromosome 2 is fused with the peroxisome proliferator-activated receptor (*PPAR*) gene on chromosome 3 [t(2;3)(q13;p25)], leading to a translocation called PAX8/PPAR gamma (PPAR- γ) rearrangement, and expression of a PAX8- PPAR- γ fusion protein, PPFP observed in about 35% of FTCs and a small fraction of follicular adenomas (Kroll *et al.*, 2000). The most common molecular alterations associated with this phenomenon in thyroid cancer include the serine/threonine-protein kinase B-Raf (BRAF) and RAt sarcoma (RAS) point mutations and RET/PTC and PAX8/PPAR-γ rearrangements in more than 70% of

papillary and follicular thyroid carcinomas(Nikiforov, 2011). Thus, FTC is characterized by PAX8/PPAR-γ rearrangements and RAS mutations, whereas *BRAF* mutations RET/PTC and TRK oncogenes have been detected in PTC (Greco *et al.*, 2009b) (Diallo-Krou *et al.*, 2009). However, the mechanism by which PPFP contributes to follicular thyroid neoplasia is not fully understood. One study described the similar frequency of PAX8- PPAR- γ rearrangement in the follicular variant of papillary thyroid carcinoma (FVPTCs), follicular thyroid carcinomas (FTCs) (45.5%), and follicular thyroid adenomas (FTAs) (33.3%)(Castro *et al.*, 2006). In FVPTCs, the PAX8-PPAR-γ rearrangement was significantly associated with multifocality and vascular invasion (Castro *et al.*, 2006). FVPTC shows some of the molecular features of follicular tumours (Castro *et al.*, 2006). For example, a PAX8-PPAR-γ rearrangement was found among patients with encapsulated FVPTC (Rivera *et al.*, 2010). This has been shown to be a significant diagnostic utility especially in thyroid fine-needle aspiration samples with indeterminate cytology (Nikiforov, 2011). The use of these and other molecular markers is believed to improve the accuracy of cancer diagnosis, thereby permitting more individualized management of thyroid cancer patients (Nikiforov, 2011).

SNP ID	Amino Acid Change	Gene ID	SNP Effect	Reference
COSM26484	Leu512Arg	TSHR	Associated with constitutive activation of the cAMP signal transduction pathway, causing thyrotoxicosis and a hot thyroid.	Gozu H., *et al*., 2004. Thyroid; 14(11):975-80.
rs1991517 C>G	Asp727Glu	TSHR	May play a contributory role in the development of insulin resistance.	Ouchi N., *et al*., 2003. Curr Opin Lipidol; 14(6)561-6.
rs2234919 C>A	Pro52Thr	TSHR	Implicated in hyperthyrotropinemia	Lorenzo P.I., *et al*., 2011. Histochem Cell Biol; 136(5):595-607.
rs61747482 G>C	Asp36His	TSHR	Has been implicated in somatic gain-of-function mutations resulting in a phenotype of toxic adenoma, toxic multinodular goiter and congenital hyperthyroidism	Krohn, K. and R. Paschke, 2002. Mol Genet Metab; 75(3):202-8. Gabriel, E.M., et al., 1999. J Clin Endocronol Metab; 84(9):3328-35.
rs13063628 G>A	Intronic	TSHR	May contribute to increased TSHβ and elevated serum TSH levels	Sorensen H.G., *et al*., 2008. Thyroid; 18(10):1087-94.
rs190110651 G>A	Gly85Arg	TSHβ	Implicated in congenital hypothyroidism	Muthukrishnan J., *et al*., 2010. Indian J of Peadiatrics 77(1):94-6.
rs121918668 G>A	Q49X (Glu49Arg)	TSHβ	Leads to premature STOP codon resulting in blocked secretion of mature TSH heterodimer	Viussoz J.M., *et al*., 2001. J Clin Endocronol Metab; 86(9):4468-71.
rs2076738 T>C	Cys1263Arg	TG	Associated with a defective intracellular transport of thyroglobulin in congenital goiter	Hishinuma, A., et al., 1999. J Clin Endocronol Metab; 84(4):1438-44.
rs2076739 T>A	Cys1995Ser			
N.A.	Glu2511Arg	TG	Thought to confer increased risk for benign thyroid disorders	Lonn, S., et al., 2007. Cancer Epidemiol Biomarkers Prev; 16(1):174-7 Matakidou, A., et al., 2004. Carcinogenesis; 25(3):369-73.
rs1799939 G>A	Gly691Ser	RET proto-oncogene	Has been implicated in a mixed medullary and follicular cell carcinoma of the thyroid. This SNP may have a primary effect on cAMP levels in the thyroid which affects the production of T4 and T3 as well as the feedback to alter TSH release by the pituitary.	Van der Deure, W.M., et al., 2008. Endocrinology; 149(10):5307-14. Maruna, P., et al., 2008. Med Sci Monit; 14(4):CS31-6.

SNP ID	Amino Acid Change	Gene ID	SNP Effect	Reference
rs2235544	Intronic	DIO1	May alter deiodinase function leading to change in circulating T3 and T4.	Panicker, V., *et al*., 2008. J Clin Endocrinol Metab; 93(8):3075-81.
rs225014 A>G	Th92Ala	DIO2	Associated with decreased enzymatic activity of DIO2 and increased risk for type 2 diabetes mellitus	Canani L.H., *et al*., 2005. J Clin Endocronol Metab; 90(6):3472-8.
N.A.	ORFa-Gly3Asp	DIO2	Thought to be involved in the development of insulin resistance. Implicated in variations in serum T3:T4 ratios.	Guo T.W., *et al*., 2004. J Med Genet 41(8):585-90.
rs1803082	Thr119Thr	Transthyretin	Has been found in euthyroid hyperthyroxinema patients	Alves I.L., *et al*., 1993. J Clin Endocronol Metab 77(2):484-8.
rs76194198 C>T	Promoter region(nt-456)	PAX8	Implicated in thyroid dysgenesis	Chowdry I., *et al*., 2004. Life sciences; 75(24):2897-909.
rs1219091790	Gly543Glu	NIS	Detected in Japanese siblings with defective Iodine transport system.	Kosugi, S., *et al*., 1998. J Clin Endocronol Metab 83(9):3373-6.
rs121909175 C>A	C272X	NIS	Known to cause congenital hypothyroidism	Pohlenz, J., *et al*., 1997. Biochem Biophys Res Commun; 240(2):488-91.
COSM565389	Gly395Arg	NIS	Implicated in autosomal recessive form of congenital hypothyroidism.	Couch, R.M., et al., 1985. The Journal of pediatrics; 106(6): p. 950-3. Kosugi, S., *et al*., 1999. J Clin ndocronol Metab; 84(9):3248-53.
rs45531732	Ala544GAla	NIS	May contribute to mechanisms leading to defective iodine transport in non-functioning thyroid nodules.	Tonacchera, M., *et al*., 2002. J Clin Endocronol Metab; 87(1): p. 352-7.

Table 2. Summary of the discussed SNPs related to thyroid hormone functions

1.6. Objectives

1.6.1. Rationale for the study

As discussed above, many thyroid disorders result from a wide range of genetic changes possibly interacting with environmental factors. Also, the therapy of these disorders may be influenced by polymorphisms in genes involved in (a) its pharmacodynamics or pharmacokinetics, such as TSHR, and (b) metabolic pathways of the thyroid hormones, such as the three deiodinases.

The deiodinases DIO1, DIO2 and DIO3 were included in this study to evaluate specifically their association with thyroxine dose requirement. As mentioned previously (Section 1.3.4.1.), the three deiodinases are thyroxine-metabolizing enzymes which affect thyroxine level directly. Hence, if any mutation occurs in any of those genes, the level of thyroxine might be affected (Figure 5). Thus, for example, inactivating mutations in DIO1 and DIO2 may influence the conversion of T4 to the active T3, leading reduced formation of T3. On the other hand, since DIO3 converts the active T3 to its inactive isomer rT3, mutations in this gene could also affect this process. The fourth gene was TSHβ, which is crucial for the synthesis of functional TSH. This gene was included to evaluate its possible association with DTC risk. Genetic changes in the functional domains of this hormone may influence its growth-stimulating effect on thyroid cells, as described in Section 1.5.3.1.

The NIS gene was also included in this study for its role in the DTC pathway, since it is now known that its expression is reduced in most hypofunctional, benign and malignant thyroid tumours, possibly triggered by DNA methylation in its critical regulatory region (Venkataraman *et al.*, 1999a). Furthermore, as discussed above, BRAFV600E mutation occurs frequent in PTC and has been shown to repress NIS expression. The possibility of intragenic alterations in NIS cannot be ruled out as additional mechanism contributing to carcinogenesis.

The last gene selected for the study was the PAX8 gene. This gene may regulate cell survival of differentiated epithelial cells through the transcriptional regulation of tp53inp1, which regulates a p53-dependent cell cycle and apoptosis (Di Palma *et al.*, 2013; Okamura *et al.*, 2001; Tomasini *et al.*, 2002).

While some progress has already been recorded in identifying polymorphisms responsible for inter-individual differences in response to therapy, many issues remain unanswered with regard to the contribution of individual genes to the disease. In addition, inter-ethnic differences exist in the prevalence and the effect of these polymorphisms on thyroxine pharmacokinetics. The prevalence of important polymorphisms in genes related to thyroid cancer and therapy thereof in the Saudi population has not been well studied. In addition to their effect on thyroxine metabolism and turnover, some of these polymorphisms may be associated with increased risk of thyroid cancer itself. Therefore, the three genes, TSHβ, PAX8 and NIS were targeted to evaluate their potential role in the development of thyroid cancer, while the three deiodinases were selected to determine their possible association with thyroxine dose requirement and their potential use as pharmacogenetic markers in athyrotic thyroid cancer patients using the Saudi population as a study model.

1.6.2. Specific Objectives

The primary objectives of the study are to:

(a) Screen for the presence of SNPs in the following genes: metabolizers (DIO1, DIO2 and DIO3), TSHβ and transcription factors/symporters (PAX8 and NIS) in the Saudi population

(b) Assess the association of these SNPs and haplotypes with the risk of thyroid cancer.

(c) Evaluate the association of selected SNPs with thyroxine dose requirement

2. Materials and Methods

The study consists of two major components: (a) establishment of the mutations or polymorphic changes present in the genes of interest, and their distribution in the normal Saudi population using randomly selected healthy individuals from the general population and (b) association studies to evaluate the relevance of selected single nucleotide polymorphisms (SNPs) in the variability of thyroid cancer risk and thyroxine dose requirement in a larger cohort of Saudi thyroid cancer patients.

2.1. Study patients and blood sampling

Initially, (48-96) healthy Saudi individuals were recruited from the Family Medicine Polyclinics at KFSHRC to establish the SNPs present in genes of interest in the general population. This population size was statistically large enough, based on our power analysis calculations (section 2.8.1.) to provide adequate information on the presence of potentially informative SNPs in a population. Power analysis allows one to determine the number of samples required to confidently exclude false negatives (i.e. polymorphisms that appear to not be associated can confidently be considered not related). Other (48-96) DTC patients were recruited from thyroid cancer clinics to gather some indication of the potential relationships of these variants with the disease. This constituted the first stage of the association studies. The initial study was followed by evaluating the association of the gene variants with risk of thyroid cancer and/or thyroxine dose requirement in the patient population. In all, 507 patients and 597 controls of Saudi origin were recruited for this purpose. Candidate patients had undergone total thyroidectomy and received radioiodine ablation and were on L-thyroxine suppressive therapy (Euthyrox, Merck Pharmaceuticals, NJ, USA), aiming at obtaining either suppressed (TSH<0.1 mU/L) or near-suppressed (0.1≤TSH<0.5 mU/L) serum TSH levels with FT4 in the normal range (12-22 pmol/L).

The exclusion criteria for this group included patients on multiple drug regimes, or those who would have changed the thyroxine brand 3 months prior to the beginning of the study. Also excluded were patients on (a) drugs that could interfere with thyroxine treatment, such as, anti-epileptics and bile acid resins, (b) thyroid suppressors and other drugs that may alter thyroid hormone metabolism or production, such as lithium, furosemide, diphenyldantoin, and propanolol and (c) medicines that could affect the pituitary-thyroid axis (high dose of steroids, sex hormone such as testosterone, oestrogen replacement). Furthermore, pregnant women, patients with mental illness, those with other types of cancer, as well as patients with significant renal impairment (glomerular filtration rate<60 ml/min) and chronic liver disease were also excluded. Compliance was determined through a questionnaire aimed at obtaining information about each subject's medical history, medication use, smoking, and measurement of TSH and FT4 levels. All candidate patients signed a written consent before being entered into the study, and were interviewed by a clinical coordinator for completing the compliance questionnaire. The sample consent forms and questionnaires on patient demographics are included in Appendix B and C**.**

2.2. DNA extraction and quality assurance

2.2.1. DNA extraction

Five ml peripheral blood was sampled from the arm by venupuncture from each thyroid cancer patient and randomly selected blood donor controls visiting KFSHRC into 6 ml purple top vacutainer tubes (Beckton Dickenson, Franklin Lakes, NJ USA) containing K_2EDTA to a final concentration of 1.5- 2.0 mg/ml blood. The blood was immediately mixed with the EDTA by repeated gentle tilting of the tube, to prevent the coagulation of blood by binding calcium ions, and then delivered to the lab for DNA extraction.

The samples were then processed immediately or stored at 4^{o}C and DNA extracted within 72 hours.

Genomic DNA was isolated from whole blood using the Gentra Puregene DNA purification blood Core kit C (Qiagen Sciences, Germantown, Maryland, USA). Briefly, up-to 3 ml whole blood were added to 9 ml red blood cell lysis buffer containing 0.15 M ammonium chloride, and incubated for 5 minutes, with continuous mixing, to facilitate the complete lysis of erythrocytes. The mixture was centrifuged at 2000 g for 2 minutes and the supernatant discarded. The leucocyte pellet was re-suspended in 3 ml cell lysis buffer solution by vigorous vortexing to facilitate complete lysis of the white blood cells. Proteins in the lysate were removed by adding 1 ml protein precipitation solution (containing 3.9mM ammonium sulphate), vortexing for 20 seconds, and centrifuging at 2000 g for 5 minutes to precipitate the protein fraction. The DNA-containing supernatant was added to 3 ml isopropanol in a fresh 15 ml tube, and mixed gently to precipitate genomic DNA. The tube was centrifuged at 2000 g for 3 minutes, the supernatant discarded, and the DNA pellet washed twice in 3 ml of 75% ethanol. The DNA pellet was air-dried and dissolved in 250 µl hydration solution (containing Tris-EDTA buffer, pH8.0) at 65^{o}C for 1 hour.

2.2.2. DNA Quantification and purity

The concentration and fidelity of the genomic DNA was determined using the Nanodrop ND-1000 spectrophotometer (Wilmington, DE, USA) at 260 nm wavelength. Sample quality was confirmed by ensuring the 260:280 ratio was 1.8 ± 0.1, and the 260:230 ratio approximately 2.0± 0.2. Since nucleotides absorb UV light at 260 nm while proteins have a strong absorbance at 280 nm, a value of 260:280 ratio significantly below 1.8 may indicate protein contamination. On the other hand, a value of the 260:230 ratio significantly lower than 2.0 may indicate contaminants which absorb at 230 nm, such as EDTA, and phenolates. Stock samples were then dispensed in 50 µl

aliquots in 1.5 ml vials, labeled, and stored at -20^{o}C. Before their use in PCR, stock DNA samples were diluted in hydration solution to a working concentration of 50 ng/µl.

2.3. Generation of target gene amplicons by PCR

2.3.1. Primer design

Each of the target genes was portioned into segments of 300-600 bases, covering the coding and promoter regions, intron-exon junctions, and the 3' and 5'untranslated regions. Suitable forward and reverse primers for the Polymerase Chain Reaction (PCR) were designed using Primer3 input software (Whitehead Institute for Biomedical Research, Cambridge, MA 02142 USA), tagged with M13 universal forward and reverse oligonucleotide sequences and sourced from Metabion (Metabion, Martinsried, Germany). The primer oligonucleotide (oligo) lengths were restricted to between 18 and 27 bases, the GC content between 20 and 80 percent, and the melting temperatures between 55^{o}C and 63^{o}C. The nearest neighbor thermodynamic parameters for more accurate prediction of the melting temperature for the primer oligos were applied according to Santa Lucia 1998. These parameters are best suited for the correct prediction the melting temperatures of oligos with heterogeneous sequences which tend to melt with several stable intermediate states (SantaLucia, 1998).

2.3.2. PCR optimisation

The optimum annealing temperature for each primer set was determined empirically by gradient PCR method. For each set of primers (for example, forward and reverse primers for exon 14 on the NIS gene, approximately 454 base pairs), a 25 µl master-mix was prepared by mixing 2.5 µl of 10x PCR buffer containing 15 mM $MgCl_2$, 0.5 µl of 10 mM dNTP mix, 2.0 µl each of 5 µM forward and reverse primers, 0.33 µl HotStar Taq DNA Polymerase (5 Units/µl), 1.0 µl of 50 ng/ml genomic DNA, and 16.7 µl of

distilled water. A sufficient volume was prepared for each primer set to allow twelve duplicates of 25 µl reactions to be set up in one row of a 96-well plate which were subjected to PCR under a temperature gradient ranging from a minimum of 50°C to a maximum of 63°C (Figure 11). The gradient PCR program used was 95°C for 10 min, followed by 36 cycles of denaturation at 95°C for 15 sec, annealing at 50°C- 63°C for 30 sec, and extension at 72° C for 1 min, and a final extension step of 10 min at 72° C. The exact temperature applying to each column of the PCR plate was calibrated in the PTC200 Thermal Cycler (MJ Research, Ramsey, MN, USA). In one 96-well plate, eight primer sets were subjected to gradient PCR simultaneously.

	50°C	50.3°C	51.1°C	52.1°C	53.6°C	55.6°C	57.8°C	59.6°C	61.0°C	62.0°C	62.8°C	63.0°C
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	P1	P1	P1	P1	P1	P1	P1	P1	P1	P1	P1	P1
B	P2	P2	P2	P2	P2	P2	P2	P2	P2	P2	P2	P2
C	P3	P3	P3	P3	P3	P3	P3	P3	P3	P3	P3	P3
D	P4	P4	P4	P4	P4	P4	P4	P4	P4	P4	P4	P4
E	P5	P5	P5	P5	P5	P5	P5	P5	P5	P5	P5	P5
F	P6	P6	P6	P6	P6	P6	P6	P6	P6	P6	P6	P6
G	P7	P7	P7	P7	P7	P7	P7	P7	P7	P7	P7	P7
H	P8	P8	P8	P8	P8	P8	P8	P8	P8	P8	P8	P8

Figure 11. The figure represents a typical 96-well plate set up for the gradient PCR with temperature gradient ranging from 50°C in column 1 to 63°C in column 12. P1-P8 represent primer sets 1 to 8.

The generated amplicons were then run on agarose gel electrophoresis alongside a 100 bp DNA size marker to facilitate the selection of the annealing temperature that gives only one discrete band for each set of primers. An example is the set of primers for exon 14 of the NIS gene TGAGGCTCAGAGGCTTGC (forward) and

TTCAGTCCTTGTGAAGGGAGA which gave an annealing temperature of 61.0^{o}C and a PCR product 454 base pairs long (Figure 12).

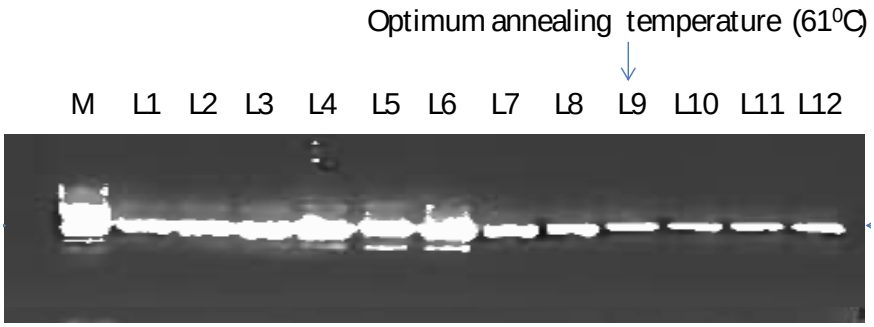


Figure 12. The figure represents a typical agarose gel of gradient PCR amplicons for the amplified NIS segment obtained within the temperature range from 50^{o}C (L1) to 63^{o}C (L12). M is DNA marker, Lanes 1-12 (L1-L12) show amplicons generated within the temperature gradient (50-63^{o}C) for primers, with Lanes 1-6 featuring multiple bands at lower temperatures.

For a number of primers, the PCR reactions resulted in amplicons with multiple bands throughout the temperature gradient, which implied non-specific amplification. In such cases, a variety of additives and enhancing agents were used in PCR to increase the specificity and consistency of the PCR reactions. Some of the more common additives include dimethyl sulphoxide (DMSO), formamide, betaine (N.N.N. trimethylglycine), tetramethyl ammonium chloride (TMAC), bovine serum albumin (BSA), glycerol, and non-ionic detergents, such as Triton X-100 and Tween 20. The Q solution provided with the Qiagen PCR kit (Qiagen GmbH, Hilden Germany) contains the mono-hydrate form of betaine, which improves specificity of PCR by suppressing the formation of secondary structures especially in the GC-rich regions, enhancing enzyme processivity, and stabilizing DNA polymerase.

In a few cases, the PCR reaction resulted in little or no product amplification. In such cases it was necessary to optimize the concentration of Mg^{2+} ions in the PCR buffer. Mg^{2+} ions bind to primers and nucleotides contained in the amplification reaction, and

are a critical DNA polymerase cofactor necessary for enzyme activity. The primer sets and optimum amplification conditions are summarized in Appendix D.

2.3.3. PCR for gene sequencing

PCR was then performed by the standard procedure on the PTC200 Thermal Cycler (MJ Research, Ramsey, MN, USA) using the Qiagen PCR kit (Qiagen GmbH, Hilden Germany). Briefly, a 25 µl master-mix was prepared for each reaction by mixing the PCR ingredients as described in section 2.3.2 above. The mixture was heated for 15 minutes at 95^{o}C to prime the HotStar Taq DNA polymerase, and thermally cycled 36 times at 95^{o}C for 30 sec, the optimum annealing temperature for 30 sec, and 72^{o}C for 1.30 min, followed by extension at 72^{o}C for 10 min and a final hold-step at 4^{o}C forever. The size and fidelity of the amplicons was monitored by agarose gel electrophoresis using a 100 base pair size marker (Invitrogen, Grand Island, NY, USA) as reference. For each PCR reaction plate, 5 µl of PCR product taken from each of 12 wells selected randomly across the plate, were run on an agarose gel. Figure 13 is an example of results obtained from the agarose gel electrophoresis of the PCR product of primers for exon 14 of the NIS gene.

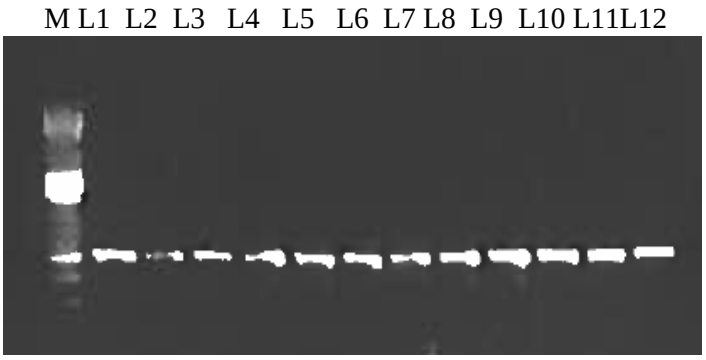


Figure 13. The figure is an agarose gel of amplicons obtained from PCR of exon 14 on the NIS gene. M is 100 bp marker, L1 to L12 are products obtained from test samples.

2.4. Screening for mutations by sequencing

Most common sequencing methods are based on Sanger's dideoxynucleotide sequencing principle. Sanger's method utilizes 2'3'dideoxynucleotides (ddNTPs) which differ from deoxynucleotides (dNTPs) by having an H atom tagged to the 3' carbon instead of an OH group. In Sanger sequencing, the DNA polymerase enzyme copies the DNA strand by adding the appropriate complementary nucleotides to the 3' end of the growing chain. Once a ddNTP is added, it terminates the PCR reaction selectively at a C, G, T, or A base, because it cannot form a phosphodiester bond with the next deoxynucleotide. The end product is a mixture of DNA fragments of different sizes which can then be separated by electrophoresis according to size. The dye terminator cycle sequencing method used here involves three steps: cycle sequencing, purification of sequencing products, and separation by capillary electrophoresis. During cycle sequencing PCR, each of the 4 ddNTPs labeled with one of the four different fluorescent dyes and all the reaction ingredients are mixed in one tube. The reaction is terminated each time a labeled ddNTP is added onto the growing DNA chain. Once thermal cycling has been completed, the reaction products are subjected to capillary electrophoresis for separation and identification. In capillary electrophoresis, the sample is passed through an electric field which causes the negatively charged DNA fragments to migrate through an array of capillaries containing a replaceable non-cross linked linear polyacrilamide matrix. Thus the fluorescently labelled extension products are separated by size based on their total charge. The fluorescent fragments then move across the path of a laser beam, which causes the dyes to fluoresce, and the signal detected and analyzed by an optical and electronic sensing device. Finally, the data capturing software converts the fluorescence signal to sequence data. Because each dye emits light at a different wavelength when excited by the laser, all four colors,

and thus all four bases, can be detected and distinguished in a single capillary. Analyzed sample data is depicted in the form of a chromatogram.

The prevalence of SNPs in the study population was determined by sequencing of amplicons, providing near-full coverage of the genes using the MegaBACE DNA analysis system (Amersham Biosciences, Sunnyvale, CA, USA) according to the manufacture's protocol. Briefly, 5 µl of PCR product were treated with 2 µl of ExoSAP-IT (USB Corporation, Cleveland, Ohio, USA) at 37^{o}C for 30 min to allow the removal of excess primers and dNTPs by hydrolytic enzymes, Exonuclease 1 and Shrimp Alkaline phosphatase. The enzymes were then deactivated at 80^{o}C for 15 min, and the cycle sequencing reaction initiated by mixing 100 ng DNA, 5umol primer, 8 µl of DYEnmic ET Dye Terminator (Amersham Biosciences, Buckinghamshire, UK) in a total reaction volume of 20 µl. The mixture was thermally cycled 40 times at 95^{o}C for 20 sec, 50^{o}C for 15 sec, and 60^{o}C for 1 min. Unincorporated dye-labelled terminators were removed by gel-filtration using the DyeEx on the 96 well-plate (Qiagen, GmbH, Hilden, Germany). The eluent was vacuum-dried and dissolved in 10 µl of loading solution (GE Healthcare UK Ltd, Little Chalfont, Buckinghamshire UK) for sequence detection by capillary electrophoresis in the MegaBACE sequencer (Amersham Biosciences, Piscataway, NJ, USA).

The sequence data were processed using several base callers currently available as part of the Megabace 1000 Capillary Sequencing System. Generally, a standard of accuracy reflecting a base calling error rate of <1000 was targeted. Data was analyzed for SNPs by the Lasergene Software (DNASTAR, Inc. Madison, WI, USA). The following is an example of a chromatogram obtained from sequencing of exon 14 on the NIS gene, showing the polymorphic change for the rs4808708 G>A at nucleotide position 18905 of the gene.

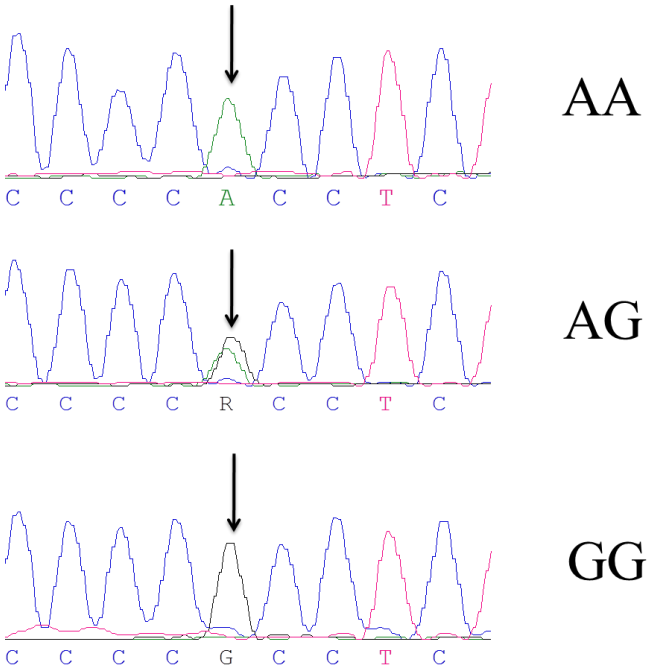


Figure 14. The figure shows chromatogram obtained during sequencing of exon 14 on the NIS gene. The variation from G>A in the rs4808708 loci on the three genotypes are depicted.

2.5. Association studies

Once the SNPs of interest were identified by sequencing, genotyping for the association study was performed in 507 patients and 597 controls of Saudi origin by real-time PCR using Taqman chemistry in the ABI Prism 7900HT Sequence Detection System (Applied Biosystems Inc. CA, USA). Primers and TaqMan probes were designed using the Primer Express software V2.0 (Applied Biosystems, Foster City, CA, USA) and procured from Applied Biosystems (Foster City, CA 94404, USA). The fluorogenic probes, bearing a suitable reporter dye on the 5'-end and a quencher dye on the 3'-end, hybridize to the specific complementary sequence bearing the SNP. One probe (for allele 1) was labelled with VIC dye, and the other (for allele 2) with FAM dye at the 5'-primer end. During the primer extension and synthesis of the nascent strand by Taq polymerase, exonuclease activity cleaves the annealed probe thereby releasing the reporter dye from its proximity to the quencher and allowing emission of fluorescence. Serial dilutions of the probes were run first to determine the optimal working concentration. For each reaction, a 25 µl master-mix was prepared by mixing 5 µl

containing 50 ng DNA, 12.5 µl of 2x Universal mix (Eurogentec, Liege Science Park, 4102 Seraing, Belgium), 1.25 µl of 20x probe assay mix (containing primers at 50 µM and probes at 5 µM stock concentrations), and 6.25 µl DNase-free distilled water. Three no-template controls were included in each 96-well plate for normalization of the emission signal. The thermal profile for amplification for the first cycle was set at 50^{o} C for 2 minutes (to optimize AmpErase UNG enzyme activity which prevents non-specific product carryover), 95^{o} C for 10 min, followed by 40 cycles of 94^{o} C for 15 sec, and 60^{o} C for 30 sec. The plates were then scanned for FRET signal using the 7900HT sequence detection system and data analysed using SDS 2.0 software (Applied Biosystems, Foster City, CA, USA).

The sample distribution obtained in the real time-PCR profile for the rs4808708 in NIS gene is given as an example in figure 15.

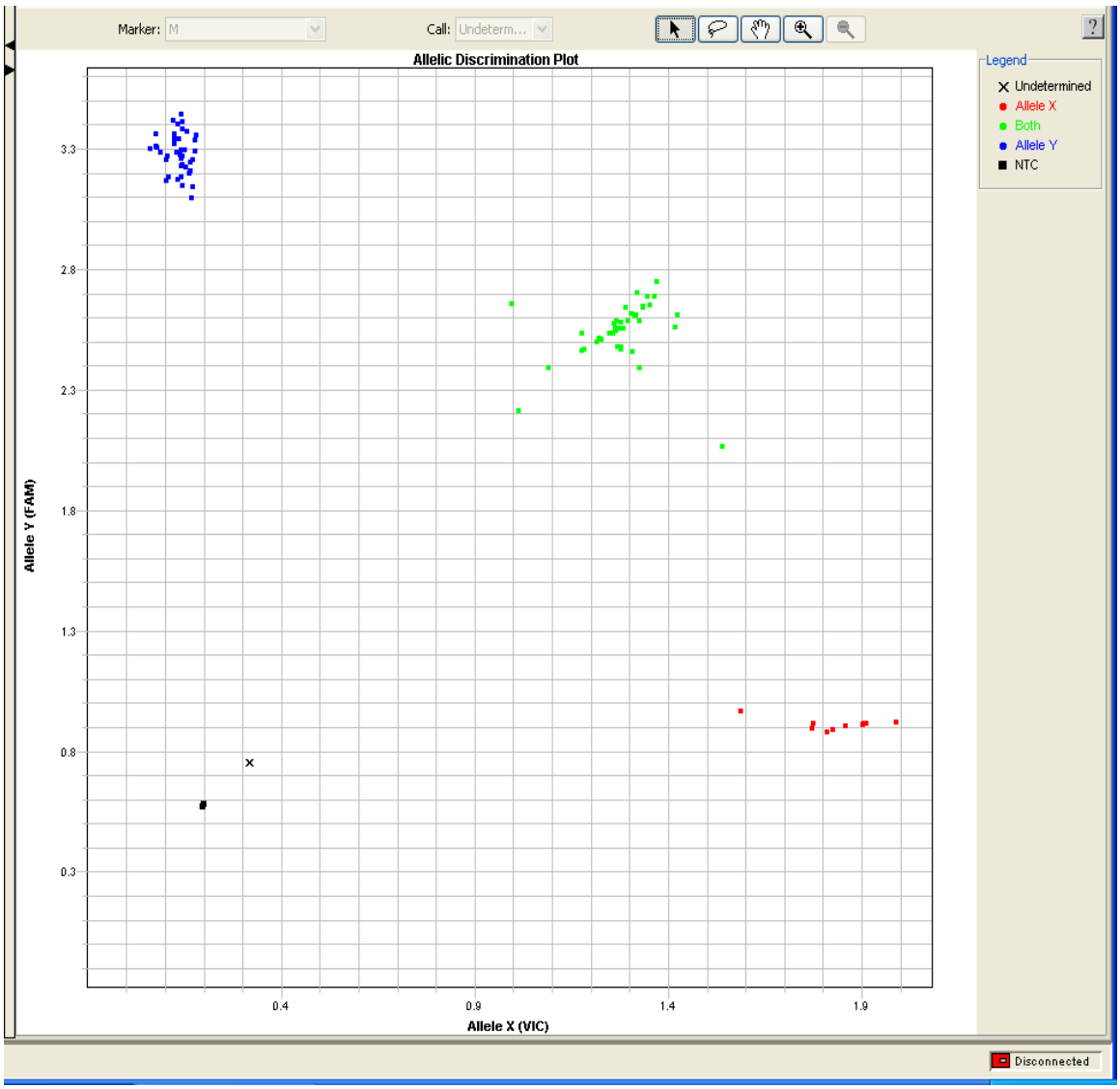


Figure 15. The figure is an allelic discrimination profile obtained for the NIS- rs4808708 genotypes on a 96 well-plate. The discrimination is indicated by the genotypes being in three different zones of the chart. For example the AA genotype (VIC, red) is found in the zone Y<1.0 and X>1.0, the GG (FAM, blue) in the zone Y>1.0 and X<1.0, and the heterozygote AG (both, green) in the zone Y>1.0 and X>1.0. The black symbols indicate three no template controls, while any other symbols outsides these zones represent undetermined genotypes.

2.6. Biochemical assays and other determinations

Values for TSH, FT4, thyroxine binding globulin (TBG) and thyroglobulin (Tg) antibodies were obtained from King Faisal Special Hospital and Research Centre Central Laboratory (American College of Pathology-accredited laboratory).

2.7. Fidelity test

Genotyping for randomly selected patients among the population sample was repeated under the same reaction conditions and comparable results were obtained in each case. This indicates that the ABI Prism 7900HT Sequence Detection System (Applied Biosystems Inc. CA, USA) machine used for real-time PCR is reliable. The actual presence/absence of the individual SNPs was verified by sequencing the reverse DNA strand. Re-sequencing of individual strands also ensured identification of SNPs with minor allele frequencies >5.0%. Furthermore, agarose gels were run for randomly selected samples to check the quality of amplification prior to subjecting them to sequencing.

2.8. Statistical Analysis

2.8.1. Power analysis and sample size calculation

The sample size was calculated based on the power of 80% using the assumption that SNPs are informative when they are present in 10% of the sample, against one sided alternative with the size of the critical region = 5%.

The null hypothesis is H_o : P_o

The alternative hypothesis being H_1:P_1

Appling the following formula by Fliess (Fliess, 1981)

$$n = \left\{ \frac{Z_\beta \sqrt{p_1 q_1} + Z_\alpha \sqrt{p_o q_o}}{p_o - p_1} \right\}^2 \qquad \text{(Equation 1)}$$

Whereby, P_o stands for the prevalence of the genotype or allele for the null hypothesis and equals to 0.10; P_1 is the prevalence of the genotype for the alternative hypothesis which is in this case >0.10; q=1-P, Z_α=1.64 is the cut-off point of the standardized

normal distribution at α level=5%; Z_β=0.84 represents the normal distribution cut- off for power of 80%.

$$n = \left\{ \frac{0.84\sqrt{(0.22)(0.78)} + 1.64\sqrt{(0.1)(0.9)}}{(0.1 - 0.22)} \right\}^2$$ (Equation 2)

After calculating using the above formula, the number of 48 individuals was determined as adequate sample size for the study.

2.8.2. Hardy-Weinberg's Equilibrium calculations

For a population in Hardy-Weinberg's equilibrium

$$p^2+2pq+q^2 = 1$$ (Equation 3)

Where p^2 = Frequency of MM genotype; 2pq = Frequency of MN genotype;

q^2 = Frequency of NN genotype, M is the common allele and N the rare allele

2.8.3. Association statistics

Comparison of genotypes and alleles distributions between patients and controls was tested by Chi- Square test. In a comprehensive logistic regression model we investigated the joint effects of potential risk factors. Associations between the outcome and a risk factor and/or a confounder were measured by the Mantel-Haenszel odds ratio. Ninety five (95%) confidence intervals were constructed on the population odds ratio. The conventional risk factors for thyroid cancer including (age, sex, smoking) were the risk factors of interest.

Comparison of genotypes and alleles between different groups for continuous dependent variables was done by Analysis of Variance (ANOVA) or Student's t-test as

appropriate. All statistical analyses were performed using the SPSS software version 20 (SPSS Inc., Chicago, USA).

2.8.4. Haplotype Association statistics

For the haplotype-based association analysis, the haplo.stats package (http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm) in the R Statistical Computing software (http://www.r-project.org/) was used to perform haplotype-based association analysis. Odd ratios for haplotypes were calculated using as reference the baseline haplotypes with the highest frequency, and the Haplotype Score statistic for the association of a haplotype with the binary trait was calculated as in Schaid *et al*. (2002)(Schaid *et al.*, 2002) and Lake *et al*., (2003)(Lake *et al.*, 2003). Significance of association was determined between haplotypes and the case-control status - a binomial trait denoting whether or not a patient had the disease.

2.8.5. Generalized Linear Models

In order to estimate the minimum adequate model for the association of the studied parameter with the DTC or thyroxine dose requirements, total deviance was calculated using the generalized linear models with fixed effects implemented using R (statistical programming language www.R-project.org/) according to (McCullagh *et al.*, 1989) Binomial error distributions were assumed for the thyroid cancer model and logistic regression used to select the best fitting model. Gaussian error distributions were assumed for L-thyroxine dose models. Initial maximal models included all main effects and parameters were estimated by maximum likelihood .The significance of each term was assessed by comparing the deviance value on removal of the term, sequentially in order to arrive at a minimal adequate model.

3. Results

3.1. Determination of the prevalence of single nucleotide polymorphisms in healthy individuals in the Saudi population

3.1.1. Foreword

It is now well-acknowledged that the distribution of genetic variations may vary between ethnical group. Hence the impact of gene polymorphisms on disease may also vary accordingly. Thus, in order to accurately evaluate the influence of alterations in a particular ethnic group, it is necessary to have the adequate knowledge about the distribution of such SNPs within the population of interest. The first part of the study was therefore directed at establishing the presence of informative SNPs in the six genes of interest, DIO1, DIO2, DIO3, PAX8, NIS and TSHβ in the Saudi population to enable us to evaluate the relevance of these genes for differentiated thyroid cancer in this population

3.1.2. Methodology

For establishing the SNPs present in genes of interest in the general population, we recruited 96 healthy Saudi individuals from the Family Medicine Polyclinics at KFSHRC. To estimate the required population size for the statistical power, the Equation 1 in Section 2.8.1 was used. Accordingly, it was determined that approximately 48 individuals were sufficient to provide informative data for the distribution of the SNPs in the general population.

DNA was extracted as described in the Methodology (section 2.2), and the prevalence of SNPs in the study population determined by direct sequencing of amplicons, providing near-full coverage of the genes using the MegaBACE DNA analysis system (Amersham Biosciences, Sunnyvale, CA, USA) according to the manufacture's protocol (section 2.3 and section 2.4). To obtain maximum number of potentially relevant SNPs

of interest, all exons, exon-intron junctions, promoter and untranslated regions were sequenced.

3.1.3. Results

3.1.3.1. Population Characteristics

Table 3 displays the demographic characteristics of the 96 studied healthy individuals. For technical reasons, sequencing could be completed for only part of these individuals in some of the genes.

Variable	Cases			Controls		
	All	**Male**	**Female**	**All**	**Male**	**Female**
N	96	24	72	96	37	59
Age(years)	43.1±14.0	44.2±14.7	42.7±13.8	45.2±16.8	52.0±17.7	40.9±14.8
Weight(kg)	75.3±16.6	76.0±11.6	75.1±18.1	81.0±20.3	86.5±19.3	77.2±20.4
BMI	30.1±6.9	27.5±4.8	30.9±7.2	31.5±7.8	30.3±5.4	33.2±9.0
T4 dose(µg/kg)	2.0±0.5	2.1±0.4	2.00±0.5	-	-	-
Smoking	13(14.0)	8(33.0)	5(7.0)	19(20.0)	13(35.0)	6(10.0)

Table 3: The demographics of the genotyped healthy individuals and case –control study group. The table shows the demographic characteristics of the 96 DTC case versus 96 controls. BMI, body mass index. Numbers in brackets are percentages of the total number in the group. Values for age, weight, BMI and T4 dose are given as mean± standard deviation.

3.1.3.2. Iodothyronine Deiodinase type I

Our first target genes were the three deiodinases 1, 2 and 3 (DIO1, DIO2 and DIO3). The sequencing results for the DIO1 gene are shown in Table 4. This gene consists of seven exons. Eight SNPs were identified in this gene, of which one was novel. One of the identified SNPs resides in exon 3, one other in intron 6 and the rest are found in the 3'-UTR in exon 7.

SNP ID	*Locus (change)*	*Genotype/allele*		*f(%)*	*HWE P-value*
ns3045 G>A	Exon 1 (G3045A= Val25Met)	Genotype	GG	46(98.0)	0.94123
			GA	1(2.0)	
			AA	0(0.0)	
		Allele	G	93(99.0)	
			A	1(1.0)	
rs2235544 C>A	Intron 3	Genotype	CC	27(56.0)	0.64941
			CA	16(33.0)	
			AA	5(11.0)	
		Allele	C	70(73.0)	
			A	26(27.0)	
rs11206244 C>T	Exon 4 3'-UTR	Genotype	CC	27(56.0)	1.0000
			CT	18(38.0)	
			TT	3(6.0)	
		Allele	C	72(75.0)	
			T	24(25.0)	
rs114860598 G>A	Exon 4 3'-UTR	Genotype	GG	44(92.0)	0.76324
			GA	4(8.0)	
			AA	0(0.0)	
		Allele	G	92(96.0)	
			A	4(4.0)	
rs113019354 C>A	Exon 4 3'-UTR	Genotype	CC	46(96.0)	0.88281
			CA	2(4.0)	
			AA	0(0.0)	
		Allele	C	94(98.0)	
			A	2(2.0)	
rs78848743 C>T	Exon 4 3'-UTR	Genotype	CC	46(96.0)	0.88281
			CT	2(4.0)	
			CC	0(0.0)	
		Allele	C	94(98.0)	
			T	2(2.0)	
rs17109585 G>C	Exon 4 3'-UTR	Genotype	GG	47(98.0)	0.94186
			GC	1(2.0)	
			GG	0(0.0)	
		Allele	G	95(99.0)	
			C	1(1.0)	
rs12095080 A>G	Exon 4 3'-UTR	Genotype	AA	35(73.0)	0.27786
			AG	13(27.0)	
			GG	0(0.0)	
		Allele	A	83(86.0)	
			G	13(14.0)	

Table 4. DIO1 genotyping in healthy individuals. The table presents the genotype and allele distributions in the DIO1 gene in the 48-96 healthy Saudi individuals. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide polymorphism identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel variants discovered in the study population. f(%) gives the frequencies of the genotypes of alleles as a percentage. HWE P-value is the statistical p-value for the Hardy-Weinberg's Equilibrium calculations.

Three of these SNPs, rs2235544 C>A, rs11206244 C>T and rs12095080 A>G, exhibited minor allele frequencies (MAFs) of greater than 0.1. Notably, the novel SNP ns3045 G>A leading to a substitution of valine with methionine at amino acid position 25 (Val25Met) in exon 1 was present in only one heterozygote. It also presented the only non-synonymous variant identified in this gene in our study population.

3.1.3.3. Iodothyronine Deiodinase type II

The iodothyronine Deiodinase type II (DIO2) gene is similarly short, constituting only four exons (Table 5). Fourteen SNPs were identified, distributed throughout the gene sequence. Notably, more than 50% of the variants exhibited a MAF rate > 0.1 (Table 5). Interestingly, all except three of the discovered SNPs reside in the 3'-UTR in exon 4 of the gene. Of the other three SNPs, two are intronic (one in intron 2 and another intron 3), while one presents a non-synonymous change (ACA>GCA) in exon four leading to Thr92Ala).

SNP ID	Locus (change)	Genotypes/alleles		*f*(%)	HWE P-value
rs56033314 T>C	Intron 3	Genotype	TT	90(97.0)	0.000001*
			TC	2(2.0)	
			CC	1(1.0)	
		Allele	T	182(98.0)	
			C	4(2.0)	
rs225011 A>G	Intron 3	Genotype	AA	35(40.2)	0.006*
			AG	30(34.5)	
			GG	22(25.3)	
		Allele	A	100(57.5)	
			G	74(42.5)	
rs225014 A>G	Exon 4 (ACA>GCA= Thr92Ala)	Genotype	AA	42(43.8)	0.55622
			AG	41(42.7)	
			GG	13(13.5)	
		Allele	A	125(65.1)	
			G	67(34.9)	
rs7140952 G>A	Exon 4 3'-UTR	Genotype	GG	81(88.0)	0.000001*
			GA	6(6.5)	
			AA	5(5.5)	
		Allele	G	168(91.0)	
			A	16(9.0)	
rs225016 A>T	Exon 4 3'-UTR	Genotype	AA	34(37.4)	0.00132*
			AT	30(33.0)	
			TT	27(29.6)	
		Allele	A	98(53.8)	
			T	84(46.2)	
rs225017 T>A	Exon 4 3'-UTR	Genotype	TT	53(57.0)	0.000001*
			TA	13(14.0)	
			AA	27(29.0)	
		Allele	T	119(64.0)	
			A	67(36.0)	

SNP ID	Locus (change)	Genotypes/alleles		*f*(%)	HWE P-value
ns31100	Exon 4	Genotype	CC	72(78.3)	0.24212
C>G	3'-UTR		CG	20(21.7)	
			GG	0(0.0)	
		Allele	C	164(89.1)	
			G	20(11.9)	
ns31116	Exon 4	Genotype	TT	59(64.1)	0.03606*
T>G	3'-UTR		TG	33(35.9)	
			GG	0 (0.0)	
		Allele	T	151(82.0)	
			G	33(18.0)	
rs181554842	Exon 4	Genotype	TT	67(72.8)	0.13152
T>G	3'-UTR		TG	25(27.2)	
			GG	0(0.0)	
		Allele	T	159(86.4)	
			G	25(13.6)	
ns31243	Exon 4	Genotype	TT	76(82.6)	0.36098
T>G	3'-UTR		TG	16(17.4)	
			GG	0(0.0)	
		Allele	T	168(91.3)	
			G	16(8.7)	
ns31278	Exon 4	Genotype	AA	82(89.1)	0.58147
A>G	3'-UTR		AG	10(10.9)	
			GG	0(0.0)	
		Allele	A	174(94.6)	
			G	10(5.4)	
ns31382	Exon 4	Genotype	AA	35(48.6)	0.00415*
A>G	3'-UTR		AG	22(30.6)	
			GG	15(20.8)	
		Allele	A	92(63.9)	
			G	52(36.1)	
rs75196191	Exon 4	Genotype	AA	93(98.9)	0.95865
A>T	3'-UTR		AT	1(1.1)	
			TT	0(0.0)	
		Allele	A	187(99.5)	
			T	1(0.5)	
rs17110436	Exon 4	Genotype	GG	88(93.6)	0.74925
G>A	3'-UTR		GA	6(6.4)	
			AA	0(0.0)	
		Allele	G	182(96.8)	
			A	6(3.2)	

Table 5: DIO2 genotyping in healthy individuals .The table gives the genotype and allele distributions in the DIO2 gene in the 96 healthy individuals. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide polymorphism identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles as a percentage. HWE P-value is the statistical p-value for the Hardy-Weinberg's Equilibrium calculations. A, C, G and T are the letter codes for the nucleotides adenine, cytosine, guanine and thymine, respectively. *Denotes significant deviation from HWE.

3.1.3.4. Iodothyronine Deiodinase type III

The iodothyronine deiodinase type III (DIO3) is the shortest of the three deiodinases, comprising only one exon. Only three changes were registered in the DIO3 gene, all of which are found in exon 1 (Table 6). One of these SNPs, C818G is a novel non-synonymous variant leading to the substitution of aspartic acid with glutamic acid at amino acid position 224 (Asp224Glu). However, the variant was present as a single heterozygous and a single homozygous mutant genotype, respectively.

SNP ID	Locus (change)	Genotypes /alleles		*f*(%)	HWE P-value
ns818 C>G	Exon 1 (C818G= Asp224Glu)	Genotype	CC	90(97.8)	0.000001*
			CG	1(1.1)	
			GG	1(1.1)	
		Allele	C	181(98.4)	
			G	3(1.6)	
rs113370931 G>T	Exon 1 3'-UTR	Genotype	GG	88(98.9)	0.95749
			GT	1(1.1)	
			TT	0(0)	
		Allele	G	177(99.5)	
			T	1(0.5)	
rs945006 T>G	Exon 1 3'-UTR	Genotype	TT	68(80.0)	0.95700
			TG	16(18.8)	
			GG	1(1.2)	
		Allele	T	152(89.4)	
			G	18(10.6)	

Table 6. DIO3 genotyping in healthy individuals. The table displays the genotype and allele distributions in the DIO3 gene in the 96 studied healthy individuals. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles as a percentage. HWE P-value is the statistical p-value for the Hardy-Weinberg's Equilibrium estimations. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. *Denotes significant deviation from HWE.

3.1.3.5. Paired box gene 8

Unlike the comparatively short deiodinase sequences, the paired box gene 8 (PAX8) gene is a relatively large gene consisting of 12 exons. The study identified 19 SNPs in this gene (Table 7), whereby more than 50% exhibited MAF values of greater than 0.1, similarly spreading across the whole gene sequence. Six of these (ns34596 G>A, ns37112 A>G, ns37508 A>C, ns43224 C>T, ns52857 insertion/deletion (Indel),

ns61053 T>C) were novel. Notably, the novel indel polymorphism introduces the TGA stop codon at nt 52857 in intron 10. With a prevalence of 10% in the general population, the presence of this indel in the study population is potentially of considerable significance.

SNP ID	Locus (change)	Genotypes/alleles		*f*(%)	HWE P-value
rs1867763 T>C	Intron 2	Genotype	TT	30(62.5)	0.02092*
			TC	12(25.0)	
			CC	6(12.5)	
		Allele	T	72(75.0)	
			C	24(25.0)	
rs13007173 T>C	Intron 4	Genotype	TT	37(78.7)	0.61498
			TC	9(19.1)	
			CC	1(2.2)	
		Allele	T	83(88.3)	
			C	11(11.7)	
ns34596 G>A	Intron 4	Genotype	GG	44(95.7)	0.88020
			GA	2(4.3)	
			AA	0(0.0)	
		Allele	G	90(97.8)	
			A	2(2.2)	
rs4849186 C>G	Intron 6	Genotype	CC	32(69.7)	0.22347
			CG	14(30.3)	
			GG	0(0.0)	
		Allele	C	78(84.8)	
			G	14(15.2)	
ns37112 A>G	Intron 6	Genotype	AA	44(91.7)	0.76324
			AG	4(8.3)	
			GG	0(0.0)	
		Allele	A	92(95.8)	
			G	4(4.2)	
ns 37508 A>C	Exon 7	Genotype	AA	22(45.8)	0.01007*
			AC	26(54.2)	
			CC	0(0.0)	
		Allele	A	70(72.9)	
			C	26(27.1)	
rs77365796 A>T	Intron 7	Genotype	AA	39(88.6)	0.68946
			AT	5(11.4)	
			TT	0(0.0)	
		Allele	A	83(94.3)	
			T	5(5.7)	
rs1110839 A>C	Exon 8	Genotype	AA	42(87.5)	0.000001*
			AC	0(0.0)	
			CC	6(12.5)	
		Allele	A	84(87.5)	
			C	12(12.5)	

SNP ID (PAX8)	Locus (change)	Genotypes/alleles		*f*(%)	HWE P-value
ns43224	Intron 8	Genotype	CC	92(96.8)	0.87574
C>T			CT	3(3.2)	
			TT	0(0.0)	
		Allele	C	187(98.4)	
			T	3(1.6)	
rs67776659	Intron10	Genotype	AA	83(90.2)	0.13959
A>G			AG	8(8.7)	
			GG	1(1.1)	
		Allele	A	174(94.6)	
			G	10(5.4)	
rs11123172	Intron10	Genotype	AA	43(45.7)	0.71706
A>G			AG	40(42.6)	
			GG	11(11.7)	
		Allele	A	126(67.0)	
			G	62(33.0)	
ns 52857(+/TGA)	Intron 10; INDEL	+/TGA	+/TGA	11(11.7)	
rs3748916	Intron10	Genotype	TT	25(26.6)	0.53601
T>C			CT	44(46.8)	
			CC	25(26.6)	
		Allele	T	94(50.0)	
			C	94(50.0)	
rs3748915	Intron10	Genotype	AA	89(94.7)	0.000001*
A>G			AG	0(0.0)	
			GG	5(5.3)	
		Allele	A	178(94.6)	
			G	10(5.4)	
rs4849176	Intron11	Genotype	AA	37(77.1)	0.000001*
A>G			AG	4(8.3)	
			GG	7(14.6)	
		Allele	A	78(81.3)	
			G	18(18.8)	
ns61053	Exon 12	Genotype	TT	43(89.6)	0.70345
T>C	3'-UTR		TC	5(10.4)	
			CC	0(0.0)	
		Allele	T	91(94.8)	
			C	5(5.2)	
rs1049137	Exon 12	Genotype	TT	36(75.0)	0.32230
T>C	3'-UTR		TC	12(25.0)	
			CC	0(0.0)	
		Allele	T	84(87.5)	
			C	12(12.5)	

SNP ID (PAX8)	Locus (change)	Genotypes/alleles		*f*(%)	HWE P-value
rs1479	Exon 12	Genotype	AA	37(77.1)	0.00001*
A>C	3'-UTR		AC	5(10.4)	
			CC	6(12.5)	
		Allele	A	79(82.3)	
			C	17(17.7)	
rs874898	Exon 12	Genotype	CC	30(62.5)	0.42353
C>G	3'-UTR		CG	17(35.4)	
			GG	1(2.1)	
		Allele	C	77(80.2)	
			G	19(19.8)	

Table 7. PAX8 genotyping in healthy individuals. The table shows the genotype and allele distributions in the PAX8 gene in the (48-96) healthy individuals. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. *f*(%) gives the frequencies of the genotypes / alleles as a percentage. HWE P-value is the statistical p-value for the Hardy-Weinberg's Equilibrium estimations. 3'-UTR, 3 prime untranslated region. A, C, G and T are the letter codes for the nucleotides adenine, cytosine, guanine and thymine respectively. *Denotes significant deviation from HWE.

3.1.3.6. Thyroid Stimulating Hormone- Beta

The thyroid stimulating hormone- beta (TSHβ) is a short gene consisting of 3 exons. In all, four SNPs were identified in our populations. Interestingly, two of these (rs72695872 C>T and rs1321108 A>G) reside in the promoter region, one rs10776792 G>A was non-synonymous in exon 2, and the fourth rs17477369 T>A was intronic (Table 8). The exonic variant rs10776792 G>A is non-synonymous, leading to the substitution of alanine to threonine at the amino acid position14 (Ala14Thr). However, it was found only as a single heterozygote in this study group.

SNP ID	Change (Locus)	Genotypes/alleles		*f*(%)	HWEP-value
rs72695872 C>T	Promoter region	Genotype	CC	20(34.5)	0.53860
			CT	30(51.7)	
			TT	8(13.8)	
		Allele	C	70(60.3)	
			T	46(39.7)	
rs1321108 A>G	Promoter region	Genotype	AA	21(43.8)	0.25692
			AG	24(50.0)	
			GG	3(6.2)	
		Allele	A	66(68.8)	
			G	30(31.2)	
rs10776792 G>A	Exon 2 (GCA > ACA =Ala14Thr)	Genotype	GG	47(97.9)	0.94186
			GA	1(2.1)	
			AA	0(0.0)	
		Allele	G	95(99.0)	
			A	1(1.0)	
rs17477369 T>A	Intron 2	Genotype	TT	42(87.5)	0.11178
			TA	5(10.4)	
			AA	1(2.1)	
		Allele	T	89(92.7)	
			A	7(7.3)	

Table 8. TSHβ genotyping in healthy individuals. The table gives the genotype and allele distributions in the DIO2 gene in the studied individuals. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles as a percentage. HWE P-value is the statistical p-value for the Hardy-Weinberg's Equilibrium estimations. A, C, G and T are the letter codes for the nucleotides adenine, cytosine, guanine and thymine respectively.

3.1.3.7. Sodium Iodide Symporter

The last gene investigated in the study was the NIS. Like the PAX8 gene, the Sodium Iodide Symporter (NIS) is a relatively large sequence containing 15 exons (Table 9). Twenty-one SNPs were identified in this gene. Similarly, more than 50% of the variants showed MAFs of greater than 0.1. Notably, all except two SNPs were either intronic or resided in the untranslated regions of the gene. Thereby, three of the SNPs were found in the 5'UTR and two in the 3'UTR of the gene. Eleven of these SNPs presented novel findings, among which (ns2734 /+T, ns8723 (+/CAAAA) and ns8781 +/A) are indels, also exhibiting a MAF of more than 0.1 in the healthy individuals.

SNP ID	Locus (change)	Genotype /allele		*f*(%)	HWE P-value
rs118133504 C>T	Exon 1 5'UTR	Genotype	CC	92(96.8)	0.875745
			CT	3(3.2)	
		Allele	TT	0(0.0)	
			C	187(98.4)	
			T	3(1.6)	
rs112077649 C>T	Exon 1 5'UTR	Genotype	CC	93(97.0)	
			CT	2(3.0)	0.917416
		Allele	TT	0(0.00)	
			C	188(98.9)	
			T	2(1.1)	
rs73518702 C>A	Exon 1 5'UTR	Genotype	CC	91(95.8)	0.833973
			CA	4(4.2)	
		Allele	AA	0(0.0)	
			C	186(97.9)	
			A	4(2.1)	
ns2650 G>A	Intron 3	Genotype	GG	47(97.9)	0.000001*
			GA	0(0.0)	
		Allele	AA	1(2.1)	
			G	94(97.9)	
			A	2(2.1)	
ns2732 A>C	Exon4 (GGA>GGC = Gly172Gly)	Genotype	AA	46(95.8)	0.88281
			AC	2(4.2)	
		Allele	CC	0(0.0)	
			A	94(97.9)	
			C	2(2.1)	
ns2734 /+T	Exon 4	/+T		8 (16.7)	
rs76616163 C>G	Intron 4	Genotype	CC	31(72.1)	0.287615
			CG	12(27.9)	
		Allele	GG	0(0.0)	
			C	74(86.0)	
			G	12(14.0)	
ns3857 A>G	Intron 4	Genotype	AA	85(92.4)	0.704441854
			AG	7(7.6)	
		Allele	GG	0(0.0)	
			A	177(96.2)	
			G	7(3.8)	
ns3944 G>A	intron 4	Genotype	GG	28(30.4)	0.000001*
			AG	64(69.6)	
			AA	0(0.0)	
		Allele	G	120(65.2)	
			A	64(34.8)	
rs7250061 C>T	Intron 5	Genotype	CC	43(93.5)	0.819166
		Allele	CT	3(6.5)	
			TT	0(0.0)	
			C	89(96.7)	
			T	3(3.3)	
ns8723 (+/CAAAA)	Intron 7 INDEL	+/CAAAA	+/CAAA	8(10.8)	
ns8753 G>C	Intron 7	Genotype	GG	61(82.4)	0.000001*
			GC	0(0.0)	
		Allele	CC	13(17.6)	
			G	122(82.4)	
			C	26(17.6)	

SNP ID	Locus (change)	Genotype /allele		*f*(%)	HWE P-value
ns8756 A>C	Intron 7	Genotype	AA	64(86.4)	0.000001*
		Allele	AC	0(0.0)	
			CC	10(13.6)	
			A	128(86.4)	
			C	20(13.6)	
ns8775 A>T	Intron 7	Genotype	AA	61(82.4)	0.000001*
			AT	0(0.0)	
		Allele	TT	13(17.6)	
			A	122(82.4)	
			T	26(17.6)	
ns8781 (+/AT)	Intron 7 INDEL	+/AT	+/AT	13(17.6)	
ns15897 G>A	Intron 12	Genotype	GG	89(96.7)	0.873686332
			GA	3(3.3)	
		Allele	AA	0(0.0)	
			G	181(98.4)	
			A	3(1.6)	
rs4808708 G>A	Intron 13	Genotype	GG	36(75.0)	0.045341
			GA	9(18.8)	
		Allele	AA	3(6.2)	
			G	81(84.4)	
			A	15(15.6)	
rs4808709 A>G	Intron 14	Genotype	AA	27(56.3)	0.576735
			AG	17(35.4)	
		Allele	GG	4(8.3)	
			A	71(74.0)	
			G	25(26.0)	
rs35036312 G>A	Intron 14	Genotype	GG	36(76.6)	0.024286
			GA	8(17.0)	
		Allele	AA	3(6.4)	
			G	80(85.1)	
			A	14(14.9)	
rs12327843 T>C	Exon 15 3'-UTR	Genotype	TT	26(54.2)	0.848324
			TC	19(39.6)	
		Allele	CC	3(6.2)	
			T	71(74.0)	
			C	25(26.0)	
rs76103356 G>C	Exon 15 3'-UTR	Genotype	GG	46(95.8)	0.88281
			GC	2(4.2)	
		Allele	CC	0(0.0)	
			G	94(97.9)	
			C	2(2.1)	

Table 9. NIS genotyping in healthy individuals. The table gives the genotype and allele distributions in the NIS gene in the studied individuals. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles as a percentage. HWE P-value is the statistical p-value for the Hardy-Weinberg's Equilibrium estimations. A, C, G and T are the letter codes for the nucleotides adenine, cytosine, guanine and thymine respectively. *Denotes significant deviation from HWE.

3.1.4. Conclusion

Put together the study identified 69 SNPs in the six studied genes. The data suggests that the genes particularly DIO2, PAX8 and NIS are highly polymorphic in our population. Twenty-four of the SNPs were novel findings. These included two non-synonymous variants ns3045G>A (Val25Met) in exon 1 of DIO1, ns818 C>G (Asp224Glu) exon 1 in DIO3, four indels comprising the ns52857 (+/TGA) in intron 10 of PAX8 and ns2734 (+/T) in exon 4, ns8723 (+/CAAAA) and ns8781 (+/AT) in intron 7 of the NIS gene. Among the familiar SNPs were also the non-synonymous variants, rs225014 A>G (Thr92Ala) in exon 4 of DIO2, rs10776792G>A (Ala14Thr) in exon 2 of the TSHβ gene. Thus, the study identified several variants that may lead to alterations in the encoded protein products of these genes, and therefore be of direct relevance with regard to their functional expression.

Notably, a great number of the discovered variants were non-coding in intronic regions, or constitute parts of the promoter or the 3'-UTR regions of the genes. This fact does not necessary render these findings irrelevant, since it is becoming increasingly clear that these regions can play a major role in disease. This subject is discussed further in the different Sections below. It should also be mentioned that the Hardy-Weinberg's equilibrium (HWE) tests indicated that, among the 69 SNPs described in the present study, about nineteen of them deviated significantly from the HW expectation. This fact may be due to societal nature of the population, which is also discussed to a more detailed extent below. Altogether, these results provide a good working basis in selecting suitable SNPs for further analysis with respect to their role in disease and therapy thereof, which is the primary subject of the study.

3.2. Association of the identified SNPs with differentiated thyroid cancer in the Saudi population by sequencing

3.2.1. Foreword

Following the discovery of several SNPs in the six genes in 96 healthy individuals, the association studies for these SNPs were performed in two stages. First, it was interesting to sequence a similar number of differentiated thyroid cancer (DTC) patients, in order to gather some indication of the potential relationships of these variants with disease. Accordingly, 96 DTC patients were sequenced and then the two data sets were compared against each other. The demographics of the DTC cases and healthy individuals employed in this set of experiments are given in Table 3. As indicated in this table, there were no significant differences between the DTC cases and healthy controls in the important confounding variables, such as age and body-mass index. This experiment would serve as the basis to select SNPs of interest for further association study in a larger cohort.

3.2.2. Methodology

For the association studies by sequencing, in addition to the 96 healthy individuals, 96 patients with differentiated thyroid cancer were also studied. DNA was extracted as described in the Methodology (section 2.2), and the prevalence of SNPs in the study population determined by direct sequencing of amplicons, providing near-full coverage of the genes using the MegaBACE DNA analysis system (Amersham Biosciences, Sunnyvale, CA, USA) according to the manufacture's protocol (Sections 2.3 and 2.4). To obtain the maximum number of possibly relevant SNPs in the gene, all exons, exon-intron junctions, promoter and untranslated regions were sequenced.

3.2.3. Results

3.2.3.1. Iodothyronine Deiodinase type I

We first evaluated the association for the deiodinases with disease. Table 10 compares the prevalence of all the SNPs in the DIO1 discovered in the cases versus those found in the healthy controls. To begin with, the sequencing experiments yielded several variants in the healthy individuals that were not present in the patient group and vice versa.

Among all 23 SNPs detected in the DIO1 gene, only three SNPs (rs2235544 C>A, rs11206244 C>T and rs12095080 A>G) were common between the healthy and the cases (Table 10). However, the statistical analysis for these three revealed no causative relationship for any of them, but rather negatively associated trend for the rs12095080 A>G [Odds ratio (95% Confidence Interval) = 0.134(0.04-0.43); p=0.0002]. Because of this scenario, the distribution of several other variants showed significant differences between DTC patient and controls. These included the five novel variants, ns (-334) A>C ($p<0.0001$), ns (-215)A ($p=0.0043$) in promoter region, and ns2950 G>A ($p=0.0007$) in the 5'UTR and ns19271A>C ($p=0.0001$) and ns19331A>C ($p=0.0022$) in the 3'-UTR in exon 4 of the gene. Several familiar intronic SNPs were also described exclusively in the DTC patients as positively associated, including rs1158818 C>G ($p<0.00001$), rs4926616 ($p<0.0001$), rs2294511A>T ($p<0.00005$) and rs12084242 ($p<0.00001$).

Notably also, the finding that the vast majority of the described SNPs were intronic or resided in the untranslated regions of the gene, pointed to the likelihood of their exerting an indirect functional impact on the protein product.

SNP I/D	Locus (change)	Allele	Controls *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns(-344) A>C	promoter	A C	96(100.0) 0(0.0)	127(66.16) 65(33.84)	0.00001	-
ns(-215) A>C	promoter	A C	96(100.0) 0(0.0)	176(91.70) 16(8.30)	0.0043	-
ns2,950 G>A	Exon 3 5'-UTR	G A	96(100.0) 0(0.0)	172(89.60) 20(10.40)	0.0007	-
rs2294510 C>T	Exon1 (CTC>CTT =Leu 13 Leu)	C T	96(100.0) 0(0.0)	189(98.44) 3(1.66)	0.5532	-
ns3045 G>A	exon 1 (GTG>ATG= Val25Met)	G A	93(98.94) 1(1.06)	192(100.0) 0(0.0)	0.3286	0
rs11588118 C>G	Intron 1	C G	96(100.0) 0(0.0)	133(83.13) 27(16.87)	0.00001	-
rs12033572 G>C	Intron 1	G C	96(100.0) 0(0.0)	184(95.83) 8(4.17)	0.0553	-
rs4926616 C>T	Intron 1	C T	96(100.0) 0(0.0)	129(67.19) 63(32.81)	0.00001	-
rs2294511 A>T	Intron 1	A T	96(100.0) 0(0.0)	80(54.80) 66(45.20)	0.00001	-
ns14,147 G>A	Intron 2	G A	96(100.0) 0(0.0)	136(95.80) 6(4.20)	0.0837	-
rs12084242 T>A	Intron 3	T A	96(100.0) 0(0.0)	147(76.60) 45(23.40)	0.00001	-
rs17109582 T>C	Intron 3	T C	96(10000) 0(0.0)	165(85.94) 27(14.06)	0.0002	-
rs2235544 C>A	Intron 3	C A	70(72.92) 26(27.08)	127(66.15) 65(33.85)	0.2827	1.38(0.80-2.37)
rs11206244 C>T	Exon 4 3'-UTR	C T	72(75.0) 24(25.0)	139(72.40) 53(27.60)	0.6739	1.14(0.65-2.00)
rs114860598 G>A	Exon 4 3'-UTR	G A	92(95.83) 4(4.17)	192(100.0) 0(0.0)	0.0118	0
ns19,253 A>C	Exon 4 3'-UTR	A C	96(100.0) 0(0.0)	180(95.74) 8(4.26)	0.0546	-
ns19,271 A>C	Exon 4 3'-UTR	A C	96(100.0) 0(0.0)	160(85.11) 28(14.89)	0.0001	-
ns19,331 A>C	Exon 4 3'-UTR	A C	96(100.0) 0(0.0)	171(90.96) 17(9.04)	0.0022	-
ns19,345 A>C	Exon 4 3'-UTR	A C	96(100.0) 0(0.0)	181(96.28) 7(3.72)	0.0995	-

SNP I/D	Locus (change)	Allele	Controls *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
rs113019354 C>A	exon 4 3'-UTR	C A	94(97.92) 2(2.08)	192(100.0) 0(0.0)	0.1103	0
rs78848743 C>T	exon 4 3'-UTR	C T	94(97.92) 2(2.08)	192(100.0) 0(0.0)	0.1103	0
rs17109585 G>T	exon 4 3'-UTR	G T	95(98.96) 1(1.04)	192(100.0) 0(0.0)	0.3333	0
rs12095080 A>G	exon 4 3'-UTR	A G	83(0.86) 13(0.13)	188(97.92) 4(2.08)	0.0002	0.14 (0.04-0.41)

Table 10. Association of DIO1 polymorphism with differentiated thyroid cancer. The table compares the DIO1 sequencing data for the 96 studied DTC patients versus 96 healthy controls. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles in absolute(percentage) values. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval.

3.2.3.2. Iodothyronine Deiodinase type II (DIO2)

A similar pattern was observed for sequencing analysis for the DIO2, whereby, among the 32 SNPs found in DIO2, only five (rs56033314 T>C, rs225011 A>G, rs7140952 G>A, rs225016 A>T and rs225017 T>A) were shared by the two groups, None of these variants showed any delineable trends either, except the rs225016 which was protective [0.6253 (0.4125-0.9479); p=0.0274]. Besides, the majority of the SNPs reside in the introns or 3'-UTR region. The majority of the SNPs were observed only in the patient group (Table 11).

SNP ID	Locus (change)	Allele	Control *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns19600 G>A	exon 3 (GAC21AAC =Asp7Asn)	G A	96 (100.0) 0 (0.0)	190(98.96) 2(1.04)	0.5540	-
ns20,045 T>G	intron 3	T G	96 (100.0) 0 (0.0)	180(97.83) 4(2.17)	0.3025	-
rs191477357 T>A	intron 3	T A	96 (100.0) 0 (0.0)	184(96.84) 6(3.16)	0.1010	-
rs1388378 C>A	intron 3	C A	96 (100.0) 0 (0.0)	175(93.09) 13(6.01)	0.0123	-

SNP ID	Locus (change)	Allele	Control *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
rs187037837 G>A	intron 3	G A	96 (100.0) 0 (0.0)	187(98.42) 3(1.58)	0.3286	-
rs2267873 T>C	intron 3	T C	96 (100.0) 0 (0.0)	174(90.63) 18(9.37)	0.0034	-
rs2267872 C>T	intron 3	C T	96 (1.00) 0 (0.0)	183(95.31) 9(4.69)	0.0643	-
rs12437279 G>A	intron 3	G A	96 (100.0) 0 (0.0)	146(77.66) 42(22.34)	0.00001	-
rs56033314 T>C	intron 3	T C	182(97.85) 4(21.15)	186(97.90) 4(2.10)	1	0.98(0.24-3.97)
rs225012 T>C	intron 3	T C	96 (100.0) 0 (0.0)	125(65.10) 67(34.90)	0.00001	-
rs225013 A>C	intron 3	A C	96 (100.0) 0 (0.0)	120(63.16) 70(36.84)	0.00001	-
rs225010 G>A	intron 3	G A	96 (100.0) 0 (0.0)	159(82.81) 33(17.19)	0.00001	-
rs225011 A>G	intron 3	A G	100(57.47) 74(42.53)	125(65.10) 67(34.90)	0.1620	0.72(0.47-1.10)
ns27682 A>G	intron 3	A G	96(100.0) 0 (0.0)	184(95.83) 6(4.17)	0.1010	-
rs201689423 C>T	intron 3	C T	96(100.0) 0 (0.0)	182(98.91) 2(1.09)	0.5477	-
rs225014 A>G	exon 4 Thr92Ala	A G	125(65.10) 67(34.90)	192(100.0) 0(0.0)	0.00001	0
rs7140952 G>A	exon 4 3’-UTR	G A	168(91.3) 16(8.7)	166(91.20) 16(8.80)	1	1.01(0.49-2.09)
ns29142 T>G	exon 4 3’-UTR	T G	96 (100.0) 0 (0.0)	184(95.83) 8(4.17)	0.0553	-
ns29156 A>G	exon 4 3’-UTR	A G	96 (100.0) 0 (0.0)	182(94.80) 8(5.20)	0.0550	-
rs225015 C>T	exon 4 3’-UTR	C T	96 (100.0) 0 (0.0)	168(87.50) 24(12.50)	0.0005	-
rs225016 A>T	exon 4 3’-UTR	A T	98(53.85) 84(46.15)	125(65.10) 67(34.90)	0.0274	0.63 (0.41-0.95)
rs225017 T>A	exon 4 3’-UTR	T A	119(63.98) 67(36.02)	127(66.84) 63(33.16)	0.5886	0.88(0.56-1.35)
ns31100 C>G	exon 4 3’-UTR	C G	164(0.89) 20(0.108)	192(1.00) 0(0.00)	0.00001	0(0)
ns31110/+G (Insertion)	exon 4 3’-UTR	/+G	44 (47.0)	0(0.0)	0.00001	0(0)
ns31116 T>G	exon 4 3’-UTR	T G	151(82.07) 33(17.03)	192(100.0) 0(0.0)	0.00001	0(0)
ns31156 C>G	exon 4 3’-UTR	C G	170(92.40) 14(7.60)	192(100.0) 0(0.0)	0.00001	0(0)
rs181554842 T>G	exon 4 3’-UTR	T G	159(86.41) 25(13.59)	192(100.0) 0(0.0)	0.00001	0(0)

SNP ID	Locus (change)	Allele	Control *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns31243 T>G	exon 4 3'-UTR	T G	168(91.30) 16(8.70)	192(100.0) 0(0.0)	0.00001	0(0)
ns31278 A>G	exon 4 3'-UTR	A G	174(94.60) 10(5.40)	192(100.0) 0(0.0)	0.0006	0(0)
ns31382 A>G	exon 4 3'-UTR	A G	92(63.89) 52(36.11)	192(100.0) 0(0.0)	0.00001	0(0)
rs75196191 A>T	exon 4 3'-UTR	A T	187(99.50) 1(0.50)	192(100.0) 0(0.0)	0.4947	0(0)
rs17110436 G>A	exon 4 3'-UTR	G A	182(96.81) 6(3.19)	192(100.0) 0(0.0)	0.0140	0(0)

Table 11. Association of DIO2 gene polymorphism with differentiated thyroid cancer. .The table compares the DIO2 sequencing data for the 96 studied DTC patients versus 96 healthy controls. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles in absolute(percentage) values. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval.

3.2.3.3. Iodothyronine Deiodinase type III

None of the SNPs described in DIO3 was shared between the two groups (Table 12). It is noteworthy however, that apart from the non-synonymous ns818 C>G found in healthy individuals only, one novel variant ns172 T>A identified in the cancer patients leads to a stop codon at nt 172 with a MAF of 0.05. Another novel variant ns148 T>G also found in cancer patients only with a MAF of 0.37 was non-synonymous, leading to a substitution of methionine with arginine at amino acid position 1.

SNP ID	Locus (change)	Allele	Controls *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns(-42) A>G	promoter	A G	96 (100.0) 0 (0.0)	188(97.92) 4(2.08)	0.3050	-
ns(-39) C>G	Promoter	C G	96 (100.0) 0 (0.0)	188(97.92) 4(2.08)	0.3050	-
ns134 T>C	exon 1 5'UTR	T C	96 (100.0) 0 (0.0)	163(94.80) 9(5.20)	0.0287	-
ns141 G>A	exon 1 5'UTR	G A	96 (100.0) 0 (0.0)	188(97.92) 4(2.08)	0.3050	-
ns148 T>G	Exon 1 (ATG>AGG= Met/Arg)	T G	96 (100.0) 0 (0.0)	103(62.05) 63(37.05)	0.00001	-

SNP ID	Locus (change)	Allele	Controls *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns170 G>A	exon 1 (AGG>CGA= Arg8Arg)	G A	96 (100.0) 0 (0.0)	188(97.92) 4(2.08)	0.3050	-
ns172 T>A	exon 1 (TTG>TAG= Leu9stop codon)	T A	96 (100.0) 0 (0.0)	170(94.44) 10(5.56)	0.0357	-
ns230 C>T	exon 1 (CTC>CTT= Leu28Leu)	C T	96 (100.0) 0 (0.0)	190(98.96) 2(1.04)	0.5540	-
ns818 C>G	exon1 (GAC>GAG= Asp224Glu)	C G	181(98.37) 3(1.63)	192(100.0) 0 (0.0)	0.1162	0(0)
rs113370931 G>T	exon 1 3'-UTR	G T	177(99.44) 1(0.54)	192(100.0) 0 (0.0)	0.4810	0(0)
rs945006 T>G	exon 1 3'-UTR	T G	152(89.41) 18(10.59)	192(100.0) 0 (0.0)	0.00001	0(0)

Table 12. Association of the DIO3 gene polymorphism with differentiated thyroid cancer. The table compares the DIO32 sequencing data for the 96 studied DTC patients versus 96 healthy controls. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles in absolute(percentage) values. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval

3.2.3.4. Thyroid Stimulating Hormone-Beta

In the Thyroid Stimulating Hormone-Beta (TSHβ) gene, 4 of the 8 SNPs described were common to both the cases and healthy controls. Two of these are in the promoter region, one in exon 2 and another resides in intron 2. Both SNPs in the promoter region rs72695872C>T [1.78(1.07-2.97); p = 0.029 and rs1321108A>G [2.15(1.28-3.62); p = 0.004] as well as the intronic rs17477369T>A [2.50(1.03-6.06); p = 0.048] were associated with disease in a potentially causative fashion.

SNP ID	Change (Locus)	Genotype /allele	Controls *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
rs1321109 A>G	Promoter region	A G	96 (100.0) 0 (0.0)	58(51.80) 54(48.20)	0.00001	-
ns (-466) A>G	Promoter region	A G	96 (100.0) 0 (0.0)	121(97.60) 3(2.40)	0.2587	-
rs72695872 C>T	Promoter region	C T	70(60.34) 46(39.66)	58(46.03) 68(53.07)	0.0288	1.78(1.07-2.97)

SNP ID	Change (Locus)	Genotype /allele	Controls *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
rs1321108 A>G	Promoter region	A G	66(68.75) 30(31.25)	94(50.54) 92(49.46)	0.0036	2.15(1.28-3.62)
rs10776792 G>A	exon 2 (GCA>ACA= Ala14Thr	G A	95(98.96) 1(1.04)	179(96.24) 7(3.76)	0.272	0.27(0.033-2.22)
rs17477369 T>A	intron 2	T A	89(92.71) 7(7.29)	122(83.60) 24(16.40)	0.0483	2.50(1.03-6.06)
rs28566771 C>G	intron 2	C G	96 (100.0) 0 (0.0)	123(78.85) 33(21.15)	0.0000 1	-
rs41312672 T>C	exon 3 3'-UTR	T C	96 (100.0) 0 (0.0)	186(98.94) 2(1.06)	0.5508	-

Table13. Association of TSHβ gene polymorphism with differentiated thyroid cancer. The table compares the TSHβ sequencing data for the 96 studied DTC patients versus 96 healthy controls. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles in absolute(percentage) values. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval

3.2.3.5. Paired Box gene 8

The Paired Box gene8 (PAX8) gene exhibited the maximum number of SNPs (42). Eleven of these were commonly shared between the two groups. Among the common variants, the rs3748915A>G [4.39(2.12-9.11), p = 0.00001] and rs4849176A>G [3.15 (1.75-5.67); p = 0.0001] were strongly implicated in disease. Interestingly, five variants ns34335G>C (p.Glu67Asp), ns34371 A>C (p.Lys80Gln), ns34421A>C (p.Lys96Gln), ns34428A>C (p.Gln98Pro), ns36839 C>A (p.Thr166Asn) that were found only in the cancer cases were non-synonymous Table14).

SNP ID	Locus (change)	Allele	Control *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns(-182) A>G	promoter	A G	96 (100.0) 0 (0.0)	177(98) 3(1)	0.3169	-
ns(-85) C>A	promoter	C A	96 (100.0) 0 (0.0)	178(98.90) 2(1.10)	0.5446	-
rs1867763 T>C	Intron 2	T C	72(75.0) 24(25.0)	117(64.29) 65(35.71)	0.0792	1.67(0.96-2.90)
ns34,335 G>C	exon 4 (GAG>CAG =Glu67Gln)	G C	96 (100.0) 0 (0.0)	182(98.91) 2(1.01)	0.5477	-

SNP ID	Locus (change)	Allele	Control *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns34,371 A>C	exon 4 (AAG>CAG= Lys80Gln)	A C	96 (100.0) 0 (0.0)	186(98.94) 2(1.06)	0.5508	-
ns34,421 A>C	exon 4 (AAA>CAA= Lys96Gln)	A C	96 (100.0) 0 (0.0)	182(98.91) 2(1.09)	0.5477	-
ns34,428 A>C	exon 4 (CAG>CCG= Gln98Pro)	A C	96 (100.0) 0 (0.0)	182(98.91) 2(1.09)	0.5477	-
rs13007173 T>C	Intron 4	T C	83(88.30) 11(11.70)	153(87.93) 21(12.07)	1	1.04(0.48-2.25)
ns34596 G>A	intron 4	G A	90(97.83) 2(2.17)	192(100.0) 0(0.0)	0.1041	0 (0)
rs74370449 G>A	intron 4	G A	96 (100.0) 0 (0.0)	162(96.43) 6(3.57)	0.0894	-
rs13015478 C>A	intron 5	C A	96 (100.0) 0 (0.0)	129(68.62) 59(31.38)	0.00001	-
ns36839 C>A	exon6 (ACT>AAT= Thr166Asn)	C A	96 (100.0) 0 (0.0)	96(55.17) 78(44.83)	0.00001	-
ns36994 G>T	intron6	G T	96 (100.0) 0 (0.0)	178(94.70) 10(5.30)	0.0343	-
ns37112 A>G	intron 6	A G	92(95.83) 4(4.17)	192(100.0) 0(0.0)	0.0118	0 (0-)
rs4849186 C>G	intron 6	C G	78 (84.80) 14 (15.20)	153(82.26) 33(17.74)	0.616	1.2(0.60-2.38)
rs4322837 A>G	exon 7	A G	96 (100.0) 0 (0.0)	187(98.42) 3(1.58)	0.3286	-
ns37508 A>C	exon 7	A C	70(72.92) 26(27.08)	192(100.0) 0(0.0)	0.00001	0 (0-)
rs77365796 A>T	intron 7	A T	83(94.32) 5(5.68)	192(100.0) 0(0.0)	0.0028	0 (0-)
rs78978662 A>G	exon 8	A G	96 (100.0) 0 (0.0)	186(97.90) 4(2.10)	0.3043	-
rs3738913 C>T	exon 8	C T	96 (100.0) 0 (0.0)	164(88.17) 22(11.83)	0.0007	-
rs1110839 A>C	exon 8	A C	84(87.5) 12(12.5)	160(83.33) 32(16.67)	0.3900	1.4(0.69-2.86)
ns43224 C>T	intron 8	C T	187(98.42) 3(1.58)	192(100.0) 0(0.0)	0.1220	0 (0-)
ns43,225 C>T	intron 8	C T	96 (100.0) 0 (0.0)	185(97.37) 5(2.63)	0.1720	-
ns43,288 C>T	intron 8	C T	96 (100.0) 0 (0.0)	186(98.94) 2(1.06)	0.5508	-
rs67776659 A>G	exon 10	A G	174(94.57) 10(5.43)	178(97.80) 4(2.20)	0.1712	0.39(0.12-1.27)
rs2241975 G>A	exon 10	G A	96 (100.0) 0 (0.0)	168(87.5) 24(12.5)	0.0005	-

SNP ID	Locus (change)	Allele	Control *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
rs11123172 A>G	exon 10	A G	126(67.02) 62(32.98)	113(58.85) 79(41.15)	0.1114	1.42(0.94-2.16)
ns52,229 /+TGA	exon 10	/+TGA	11(11.0)	15(15.0)	0.5272	1.41(0.61-3.27)
ns52,229 /-TGA	exon 10	/-TGA	0(0.00)	80(84.0)	0.0000	-
rs3748916 T>C	exon 10	T C	94(50.0) 94(50.0)	108(56.25) 84(43.75)	0.2581	1.29(0.86-1.93)
rs3748915 A>G	exon 10	A G	178(94.68) 10(5.32)	154(80.21) 38(19.79)	0.00001	4.39(2.12-9.11)
rs11123171 C>T	exon 10	C T	96 (100.0) 0 (0.0)	179(97.28) 5(2.72)	0.1688	-
ns53,236 G>T	exon 10	G T	96 (100.0) 0 (0.0)	182(98.91) 2(1.09)	0.5477	-
rs4849176 A>G	exon 11	A G	78(81.25) 18(18.75)	110(57.90) 80(42.10)	0.0001	3.15 (1.75-5.67)
ns60,253 A>T	intron 11	A T	96 (100.0) 0 (0.0)	179(98.35) 3(1.65)	0.3193	-
rs1049137 T>C	exon 12 3'-UTR	T C	84(87.5) 12(12.5)	155(80.73) 37(19.27)	0.1837	1.67(0.83-3.38)
rs1479 A>C	exon 12 3'-UTR	A C	79(82.30) 17(17.70)	155(80.73) 37(19.27)	0.8729	1.11(0.59-2.09)
rs1478 A>C	exon 12 3'-UTR	A C	96 (100.0) 0 (0.0)	168(88.42) 24(11.58)	0.0005	-
ns62,088 G>A	exon 12 3'-UTR	G A	96 (100.0) 0 (0.0)	140(93.33) 10(6.67)	0.0153	-
ns62140 G>A	exon 12 3'-UTR	G A	96 (100.0) 0 (0.0)	96(80.0) 24(20.0)	0.00001	-
ns61053 T>C	exon 12 3'-UTR	T C	91(94.80) 5(5.20)	192(100.0) 0(0.0)	0.0038	0 (0)
rs874898 C>G	exon 12 3'-UTR	C G	77(80.21) 19(19.79)	192(100.0) 0(0.0)	0.00001	0 (0)

Table 14. Association of PAX8 gene polymorphism with DTC. The table compares the PAX8 gene sequencing data for the 96 studied DTC patients versus 96 healthy controls. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles in absolute(percentage) values. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval

3.2.3.6. Sodium Iodide Symporter

Like the PAX8 gene, 33 SNPs were found in the Sodium Iodide Symporter (NIS) gene, but the great majority of them were uniquely identified with either healthy individuals or DTC patients.

SNP ID	Locus (change)	Allele	Control (%)	Cancer *f*(%)	P-value	OR(95%CI)
ns22 C>T	exon 1 5'UTR	C T	96(100.0) 0(0.0)	173(92.02) 15(7.98)	0.0032	-
ns44 C>T	exon 1 5'UTR	C T	96(100.0) 0(0.0)	182(96.81) 6(3.19)	0.0998	-
ns87 A>T	exon 1 5'UTR	A T	96(100.0) 0(0.0)	170(94.44) 10(5.56)	0.0357	-
rs118133504 C>T	exon 1 5'UTR	C T	187(98.42) 3(1.58)	182(98.91) 2(1.09)	1	0.69(0.11-4.15)
rs112077649 C>T	exon 1 5'UTR	C T	188(98.95) 2(1.05)	181(99.5) 1(0.5)	1	0.52(0.47-5.78)
rs73518702 C>A	exon 1 5'UTR	C A	186(97.90) 4(2.10)	192(100.0) 0(0.0)	0.0602	0 (0)
rs7255852 C>A	intron 1	C A	96(100.0) 0(0.0)	160(90.91) 16(9.09)	0.0017	-
rs115249627 A>G	intron 1	A G	96(100.0) 0(0.0)	174(91.60) 16(8.40)	0.0045	-
ns2650 G>A	intron 3	G A	94(97.92) 2(2.08)	192(100.0) 0(0.0)	0.1103	0 (0)
ns2732 A>C	Exon4 (GGA>GGC =Gly172Gly)	A C	94(97.92) 2(2.08)	192(100.0) 0(0.0)	0.1103	0 (0)
ns2734 /+T	exon 4	/+T	8 (16.70)	0(0.0)	0.0001	0 (0)
rs76616163 C>G	intron 4	C G	74(86.05) 12(13.05)	192(100.0) 0(0.0)	0.00001	0 (0)
ns3857 A>G	intron 4	A G	177(96.20) 7(3.80)	192(100.0) 0(0.0)	0.0063	0 (0)
ns3944 G>A	intron 4	G A	120(65.22) 64(34.78)	192(100.0) 0(0.0)	0.00001	0 (0)
rs7250061 C>T	intron 5	C T	89(96.74) 3(3.26)	192(100.0) 0(0.0)	0.0332	0 (0)
ns8723 (+/CAAAA)	intron 7	+/CAAA A	8(10.0)	7(7.0)	0.5869	0.66(0.23-1.90)
ns8749 /+AAAA	intron 7	/+AAAA	0(0.0)	1(1.05)	1	-

SNP ID (NIS)	Locus (change)	Allele	Control *f*(%)	Patients *f*(%)	P-value	OR(95%CI)
ns8753 G>C	intron 7	G C	122(82.43) 26(17.57)	192(100.0) 0(0.0)	0.0000	0 (0-)
ns8756 A>C	intron 7	A C	128(86.49) 20(13.51)	97(79.50) 25(20.50)	0.1414	1.65(0.87-3.14)
ns8775 A>T	intron 7	A T	122(82.43) 26(17.57)	119(99.16) 1(0. 84)	0.0000	0.039(0.01-0.3)
ns8781 Insertion (+/AT)	intron 7	(+/AT)	13(17.57)	0(0.00)	0.0000	0 (0-)
ns8989 C>T	intron 8	C T	96(100.0) 0(0.0)	114(96.61) 4(3.39)	0.1294	-
rs45602038 C>T	exon 13 TGC >TGT= (p.cys542cys)	C T	96(100.0) 0(0.0)	180(95.74) 8(4.26)	0.0546	-
ns15897 G>A	intron 13	G A	181(98.40) 3(1.60)	192(100.0) 0(0.0)	0.1162	0 (0-)
rs57670994 /-TAT	intron 13	/-TAT INDEL	0(0.0)	14(14.90)	0.0050	-
rs6748740 /- TAT	intron 13	/- TAT INDEL	0(0.0)	2(2.12)	0.5492	-
rs4808708 G>A	intron 13	G A	81(84.38) 15(15.62)	136(72.34) 52(27.66)	0.0266	2.06(1.09-3.90)
rs4808709 A>G	intron 14	A G	71(73.96) 25(26.04)	96(51.06) 92(48.04)	0.0002	2.72(1.59-4.66)
rs7250346 G>C	intron 14	G C	96(100.0) 0(0.0)	104(54.74) 86(45.26)	0.0000	-
rs35036312 G>A	intron 14	G A	80(85.11) 14(14.89)	147(77.37) 43(22.63)	0.1564	1.67(0.86-3.24)
rs12327843 T>C	exon 15 3'-UTR	T C	71(73.96) 25(26.04)	135(72.60) 51(27.40)	0.8876	1.07(0.61-1.87)
ns22,440 G>A	exon 15 3'-UTR	G A	96(100.0) 0(0.0)	159(98.15) 3(1.85)	0.2963	-
rs76103356 G>C	exon 15 3'-UTR	G C	94(97.92) 2(2.08)	192(100.0) 0(0.0)	0.1103	0 (0)

Table 15. Association of the NIS gene polymorphism with differentiated thyroid cancer. The table compares the NIS sequencing data for the 96 studied DTC patients versus 96 healthy controls. 3'-UTR, 3 prime untranslated region. SNP ID gives the single nucleotide identification number denoted with the international "rs" nomenclature for known SNP and referred to as "ns" for the novel SNPs discovered in the study population. f(%) gives the frequencies of the genotypes of alleles in absolute(percentage) values. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval

However, one novel indel ns8723 (+/CAAAA) was shared by the two groups, and another ns8749 /+AAAA was present only in patients. In addition two familiar indels rs57670994 /-TAT rs6748740 /- TAT were also found only in the cases.

3.2.4. Conclusions

As noted in the previous Section evaluating the prevalence of the SNPs in the general population, the distribution of the variants in the two groups (patients and controls) appeared to differ in a number of cases. In particular, generally more SNPs were reported in the disease cases than in healthy individuals. These variations might however be explained by fact that some of the genes were sequenced in only part of the healthy groups (i.e. the number of sequenced healthy individuals was not always the same as the cases), rather than any particular genetic phenomenon. Nonetheless, some SNPs appeared to be implicated in DTC in a potentially causative fashion, which lays the basis for further investigation in larger population sizes. Besides, combining the data of the two groups revealed that several of these SNPs would still occur with MAFs greater than 0.1, rendering their presence as potential susceptibility factor for the disease. In addition, the discovery of mutations involving indels in the population, some of which were novel and only present in the patients can be conceived as potentially very important for DTC manifestation. It is also noticeable that a number of SNPs found in the present study have recently been discussed with respect to their possible role in regulating thyroid hormone levels and thyroxin dose requirement (Hershman *et al.*, 1983; Panicker *et al.*, 2008). Besides, other variants in these genes have also been studied recently by various investigators with respect to their role in cancer and thyroid dyshormonogenesis, with some being implicated in changes in T3 levels, causing thyroid cancer or congenital central hypothyroidism, all of which are discussed further below.

3.3. Association of the identified SNPs with differentiated thyroid cancer by real-time polymerase chain reaction

3.3.1. Foreword

Following the initial association study for the discovered SNPs by sequencing (section 3.2.) comparing 96 each of the healthy group and differentiated thyroid cancer (DTC) patients, in this part of the study, 39 SNPs were selected from the six studied genes based on their frequencies and/or locations on the gene for further study by real-time PCR. These SNPs were common to both groups, and were considered informative for purpose of reproducibility of the original findings. The genotyping for the association study to evaluate their potential role as risk for (DTC) in the general population was performed in 507 patients and 597 controls of Saudi origin.

3.3.2. Methodology

For case-control study, the Taqman chemistry was used on the ABI Prism 7900HT Sequence Detection System (Applied Biosystems Inc. CA, USA) as described in materials and methods (section 2.5.). Briefly, primers and TaqMan probes were designed using the Primer Express software V2.0 (Applied Biosystems, Foster City, CA, USA) and procured from Applied Biosystems (Foster City, CA 94404, USA). Data was analyzed using the SDS 2.0 software (Applied Biosystems, Foster City, CA, USA). Figure 16 demonstrates typical results of a Taqman assay.

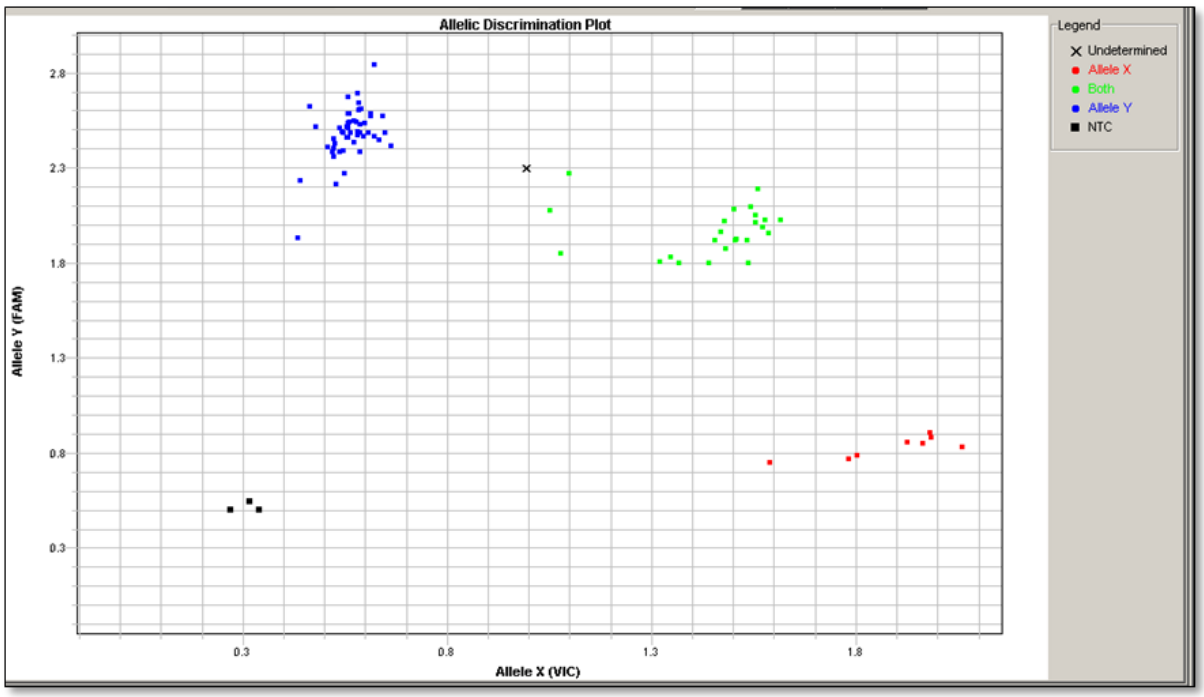


Figure 16. The figure is an allelic discrimination profile obtained for the rs1478 genotypes on a 96 well-plate. The discrimination is indicated by the genotypes being in three different zones of the chart. For example the GG genotype (VIC, red) is found in the zone Y<1.0 and X>1.0, the TT (FAM, blue) in the zone Y>1.0 and X<1.0, and the heterozygote GT (both, green) in the zone Y>1.0 and X>1.0. The black symbols indicate three no template controls, while any other symbols outsides these zones represent undetermined genotypes.

3.3.3. Results

3.3.3.1. Population Characteristics

The demographics of the DTC cases and healthy individuals employed in the population set of experiments are given in Table 16.

	Patients			**Controls**		
	All (507)	Male(92)	Female(415)	All(597)	Male(167)	Female(430)
Age(years)	45.6±15.0	47.2±15.2	44.6±12.7	45.0±13.0	52.3±15.6	43.1±14.1
BMI	30.3±6.6	28.1±5.7	30.8±6.6	29.4±6.7	28.45±5.9	29.8±6.9
Smoking	36 (7.1%)	33 (35.9%)	3 (0.7%)	90 (15.1%)	66 (39.5%)	24 (5.6%)

Table 16: Demographic data for the individuals in the population-based association study for cancer risk. There was no significant different between the two genders in both age and body mass index. However, there significantly more smokers among the male compared to the females in both the case and control groups. A total of 39 SNP in 6 genes were evaluated. Age and body mass index (BMI) are given as mean ±standard deviation.

As indicated in this table, there were no significant differences between the cases and healthy controls in the important confounding variables, such as age and body-mass index. However, there were significantly more smokers among controls than among patients ($p<0.0001$). Among the 507 DTC cases, 97.8% of the cases had papillary thyroid cancer (PTC) and 2.2 % had follicular thyroid cancer (FTC) (figure 17).

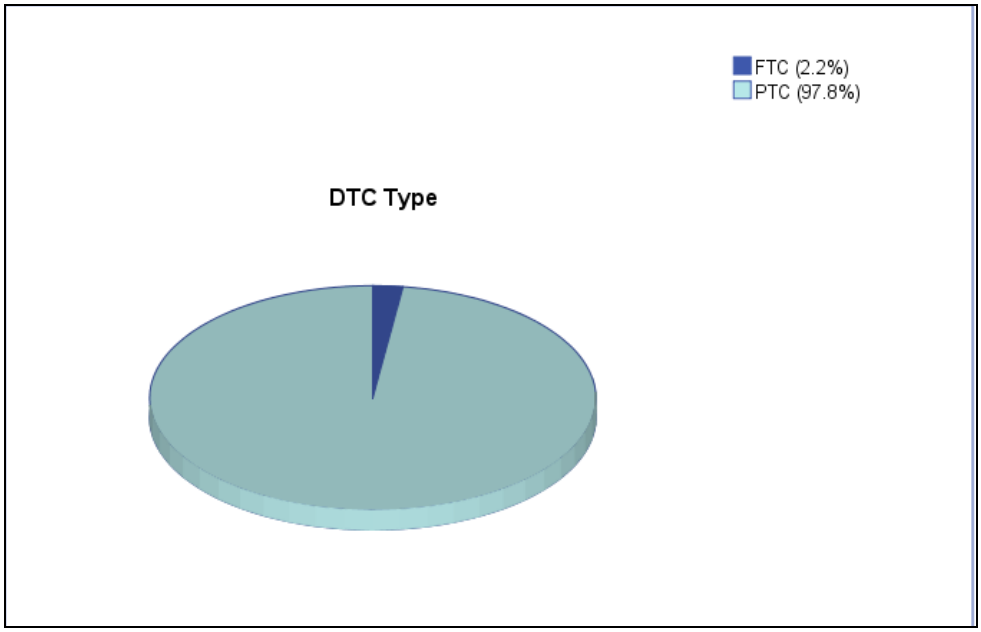


Figure 17. Distribution of the DTC subtypes in the 507 patients involved in the population-based association study for cancer risk.

3.3.3.2. Iodothyronine Deiodinase type I

The fact that the vast majority of the discovered SNPs were found only in the DTC patients confirmed the interest to find out further whether they were indeed associated with disease in large population sizes. Among the 39 SNPs selected for further study in the general population, 8 belonged to the DIO1 gene. These SNPs are given in Table 17. As indicated in this table, none of the SNPs was associated with disease. Only one variant rs2294511_AA [1.27(0.99-1.62); p=0.065] appeared to be weakly related in the univariate analysis, but could not retain this association following the correction for confounders. These observations seem to suggest that the original observations of

biased prevalence in the patient group might have been erroneous due to the small numbers engaged in the initial study.

SNP ID	Geno-types	Patients f(%)	Control f(%)	Univariate Analysis		Multivariate Analysis	
				P	OR(95%CI)	¶P	OR (95%CI)
rs12033572_G>C	G C	995(98.1) 19(1.9)	1168(97.8) 26(2.2)	0.72	0.85(0.47-1.55)	NS	NS
	GG GC+CC	488(96.3) 19(3.7)	572(95.8) 25(4.2)	0.75	0.89(0.48-1.63)	NS	NS
	GG+GC CC	507(100.0) 0(0.0)	(99.8) 1(0.2)	1.00	0.00(0.00)	NS	NS
rs12084242_T>A	T A	866(85.6) 146(14.4)	1020(85.7) 170(14.3)	0.97	1.01(0.79-1.28)	NS	NS
	TT AA+TA	371(73.3) 135(26.7)	437(73.4) 158(26.6)	1.00	1.00(0.77-1.31)	NS	NS
	TA+TT AA	495(97.8) 11(2.2)	583(98.0) 12(2.0)	1.00	1.08(0.47-2.46)	NS	NS
rs2294510_C>T	C T	983(96.9) 31(3.1)	1158(97.0) 36(3.0)	1.00	0.98(0.60-1.60)	NS	NS
	CC CT+TT	477(94.1) 30(5.9)	562(94.1) 35(5.9)	1.00	0.99(0.59-1.63)	NS	NS
	CC+CT TT	506(99.8) 1(0.2)	(99.8) 1(0.2)	1.00	0.84(0.53-13.60)	NS	NS
rs2294511_A>T	A T	632(62.5) 380(37.5)	697(59.1) 483(40.9)	0.11	1.15(0.97-1.36)	NS	NS
	AA AT+TT	197(38.9) 309(61.1)	197(33.4) 393(66.6)	0.065	1.27(0.99-1.62)	NS	NS
	AA+AT TT	435(86.0) 71(14.0)	500(84.7) 90 (15.3)	0.62	1.10(0.78-1.54)	NS	NS
rs12095080_A>G	A G	851(85.4) 145(14.6)	1000 (84.9) 178 (15.1)	0.76	1.04(0.82-1.32)	NS	NS
	AA AG+GG	365 (73.3) 133 (26.7)	424(72.0) 165(28.0)	0.63	1.06(0.81-1.39)	NS	NS
	AA+AG GG	486 (97.6) 12 (2.4)	576 (97.8) 13 (2.2)	0.98	0.91(0.41-2.02)	NS	NS

SNP ID	Geno-types	Patients f(%)	Control f(%)	Univariate Analysis		Multivariate Analysis	
				P	OR(95%CI)	¶P	OR(95%CI)
rs11206244_C>T	C T	720 (71.4) 288 (28.6)	842(70.9) 346(29.1)	0.81	1.02(0.85-1.23)	NS	NS
	CC CT+TT	259(51.4) 245(48.6)	307(51.7) 287(48.3)	0.97	0.98(0.77-1.25)	NS	NS
	CC+CT TT	461(91.5) 43(8.5)	535(90.1) 59 (9.9)	0.48	1.18(0.78-1.78)	NS	NS
rs4926616_C>T	C T	574(56.8) 436(43.2)	669(56.2) 521(43.8)	0.80	1.02(0.86-1.21)	NS	NS
	CC CT+TT	161(31.9) 344(68.1)	184(30.9) 411(69.1)	0.78	1.04(0.81-1.35)	NS	NS
	CC+CT TT	413(81.8) 92(18.2)	485(81.5) 110(18.5)	0.97	1.01(0.74-1.38)	NS	NS
rs17109582_T>C	T C	857 (85.0) 151 (15.0)	1007(84.8) 181(15.2)	0.91	0.98(0.77-1.23)	NS	NS
	TT TC+CC	368(73.0) 136(27.0)	428(72.1) 166(27.9)	0.77	0.95(0.73-1.24)	NS	NS
	TT+TC CC	489(97.0) 15(3.0)	579(97.5) 15(2.5)	0.78	1.18(0.57-2.44)	NS	NS

Table 17: Association of DIO1 variants with differentiated thyroid cancer risk. The table compares the DIO1 real-time PCR data for the 507 studied DTC patients versus 597 healthy controls. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature. f(%) gives the frequencies of the genotypes of alleles as a percentage. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval; NS, non-significant; P, P value. ¶Each SNP was entered in a multivariate analysis including age, sex, and smoking.

3.3.3.3. Iodothyronine Deiodinase type II

Seven DIO2 variants were selected from the 39 SNPs for further analysis (Table 18). However, none of these variants showed any delineable trends either, pointing to lack of association with disease at least for the studied SNPs.

SNP ID	Geno-types	Patients f(%)	Control f(%)	Univariate Analysis		Multivariate Analysis	
				P	OR(95%CI)	¶P	OR(95%CI)
rs225010 C>T	C T	628(61.9) 386(38.1)	735(61.7) 457(38.3)	0.93	1.01(0.85-1.20)	NS	NS
	CC CT+TT	197(38.9) 310(61.1)	239(40.1) 367(59.9)	0.71	0.94(0.74-1.21)	NS	NS
	CC+CT TT	431(85.0) 76(15)	496(83.2) 100(16.8)	0.46	1.14(0.82-1.58)	NS	NS

SNP ID	Geno-types	Patients f(%)	Control f(%)	Univariate Analysis		Multivariate Analysis	
				P	OR(95%CI)	¶P	OR(95%CI)
rs225011 T>C	T	626(61.9)	731(61.3)	0.83	0.97(0.82-1.16)	NS	NS
	C	386(38.1)	461(38.7)				
	TT	197(38.9)	236(39.6)	0.87	1.02(0.80-1.31)	NS	NS
	TC+CC	309(61.1)	360(60.4)				
	TT+TC	429(84.8)	495(83.1)				
	CC	77(15.2)	101(16.9)	0.43	0.88(0.63-1.21)	NS	NS
rs225014 T>C	T	711(70.3)	818(68.6)	0.43	0.92(0.77-1.11)	NS	NS
	C	301(29.7)	374(31.4)				
	TT	256(50.6)	291(48.8)	0.60	0.93(0.73-1.18)	NS	NS
	TC+CC	250(49.4)	305(51.2)				
	TT+TC	455(89.9)	527(88.4)	0.48	0.85(0.58-1.25)	NS	NS
	CC	51(10.1)	69(11.6)				
rs225015 G>A	G	727(71.7)	841(70.7)	0.63	0.95(0.79-1.14)	NS	NS
	A	287(28.3)	349(29.3)				
	GG	263(51.9)	304(51.1)	0.84	0.96(0.76-1.22)	NS	NS
	GA+AA	244(48.1)	291(48.9)				
	GG+GA	464(91.5)	537(90.3)	0.53	0.85(0.56-1.29)	NS	NS
	AA	43(8.5)	58(9.7)				
rs225017 T>A	T	607(59.9)	688(57.6)	0.30	1.09(0.92-1.30)	NS	NS
	A	407(40.1)	506(42.4)				
	TT	185(36.5)	215(36.0)	0.92	1.02(0.79-1.30)	NS	NS
	TA+AA	322(63.5)	382(64.0)				
	TT+TA	422(83.2)	473(79.2)	0.10	1.30(0.95-1.76)	NS	NS
	AA	85(16.8)	124(20.8)				
rs11159446 G>C	G	929(91.8)	1087(91.0)	0.57	0.90(0.67-1.22)	NS	NS
	C	83(8.2)	107(9.0)				
	GG	430(85.0)	495(82.9)	0.39	0.85(0.62-1.18)	NS	NS
	GC+CC	76(15.0)	102(17.1)				
	GG+GC	499(98.6)	592(99.2)	0.56	1.66(0.52-5.26)	NS	NS
	CC	7(1.4)	5(0.8)				
rs7140952 C>T	C	916(90.3)	1084(90.8)	0.77	0.94(0.71-1.26)	NS	NS
	T	98(9.7)	110(9.2)				
	CC	414(81.7)	495(82.9)	0.58	0.91(0.67-1.25)	NS	NS
	CT+TT	93(18.3)	102(17.1)				
	CC+CT	502(99.0)	589(98.7)	0.79	1.36(0.44-4.19)	NS	NS
	TT	5(1.0)	8(1.3)				

Table 18: Association of DIO2 variants with differentiated thyroid cancer risk .The table compares the DIO2 real-time PCR data for the 507 studied DTC patients versus 597 healthy controls. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature. f(%) gives the frequencies of the genotypes of alleles as a percentage. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval; NS, non-significant; P, P value. ¶Each SNP was entered in a multivariate analysis including age, sex, and smoking

3.3.3.4. Iodothyronine Deiodinase type III

Of all the studied deiodinases variants, only one DIO3 variant, the rs945006_T>G was associated with the disease.

SNP ID	Geno-types	Patients f(%)	Control f(%)	Univariate Analysis P	OR(95%CI)	Multivariate Analysis ¶P	OR(95%CI)
rs945006 T>G	T	949(93.6)	1074(90.3)	0.006*	0.63(0.46-0.87)	0.008*	0.64(0.47-0.89)
	G	65(6.4)	116(9.7)				
	TT	447(88.2)	488(82.0)	0.006*	0.61(0.43-0.86)	0.009*	0.62(0.44-0.88)
	GT+GG	60(11.8)	107(18.0)				
	TT+GT	502(99.0)	586(98.5)	0.61	0.64(0.21-1.94)	NS	NS
	GG	5(1.0)	9(1.5)				

Table 19: Association of DIO3 variants with thyroid cancer risk. The table compares the DIO3 real-time PCR data for the 507 studied DTC patients versus 597 healthy controls. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature. f(%) gives the frequencies of the genotypes of alleles as a percentage. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval; NS, non-significant; P, P value. ¶Each SNP was entered in a multivariate analysis including age, sex, and smoking.*P<0.05.

Thereby, the minor allele rs945006_G [0.63(0.46-0.87); p=0.006] was the cause and the association was retained following the multivariate analysis [0.64(0.47-0.89); p=0.008]. Also, the dominant mode of inheritance was associated with the disease.

3.3.3.5. Thyroid Stimulating Hormone-Beta

In the real-time PCR experiment, four SNPs, rs1321108_A>G, rs1321109_A>G, rs28566771_C>G and rs10776792_G>A in the thyroid stimulating hormone-beta TSHβ gene were tested for their possible role in thyroid cancer (Table 20). First, Univariate analysis was performed to test if the 4 variants were related to thyroid cancer. The analysis suggested that the rs1321108_G minor allele [1.26(1.07-1.50); p=0.007] was associated with disease, while the rs1321109_G [0.79(0.66-0.95); p=0.02] was protective against the disease. In order to correct for possible influence of confounders on these observations, the data was subjected to multivariate analysis, including age, sex and smoking. This analysis revealed that the two SNPs retained the significance of their associations, demonstrating the fact that the rs1321108_G minor allele 1.27[(1.07-1.51); p=0.006] was causative of the disease. Similarly, the dominant inheritance mode for rs1321109_(AG+GG) [0.68(0.57-0.82); p<0.00001] was primarily responsible for the

protective property of this SNP. The other two variants, rs28566771 andrs10776792 did not show any delineable relationship with the disease.

SNP ID	Geno-types	Patients f(%)	Control f(%)	Univariate Analysis		Multivariate Analysis(Adjusted)¶	
				P-value	OR(95%CI)	P-value	OR(95%CI)
rs1321108_A>G	A	553 (54.6)	717(60.5)	0.007*	1.26(1.07-1.50)	0.006*	1.27(1.07-1.51)
	G	459 (45.4)	469(39.5)				
	AA	156 (30.8)	220(37.1)	0.03*	1.32 (1.03-1.70)	0.04*	1.30 (1.012-1.69)
	AG+GG	350 (69.2)	373(62.9)				
	AA+AG	397 (78.5)	497(83.8)	0.02*	1.42(1.04-1.92)	0.01*	1.45(1.06-1.99)
	GG	109 (21.5)	96(16.2)				
rs1321109_A>G	A	692 (68.4)	753(63.3)	0.01*	0.79(0.66-0.95)	0.006*	0.77(0.64-0.92)
	G	320 (31.6)	437(36.7)				
	AA	242 (47.8)	237(39.8)	0.008*	0.72(0.56-0.91)	0.00*	0.68(0.57-0.82)
	AG+GG	264 (52.2)	358(60.2)				
	AA+AG	450 (88.9)	516(86.7)	0.26	0.81(0.62-1.05)	NS	NS
	GG	56 (11.1)	79(13.3)				
rs28566771_C>G	C	551 (54.4)	693(58.0)	0.09	0.86(0.72-1.02)	NS	NS
	G	462 (45.6)	501(42.0)				
	CC	153 (30.2)	206(34.5)	0.13	0.82(0.64-1.06)	0.16	0.83(0.64-1.07)
	CG+GG	353 (69.8)	391(65.5)				
	CC+CG	397 (78.5)	487(81.6)	0.19	0.82(0.61-1.11)	0.21	0.82(0.61-1.11)
	GG	109 (21.5)	110(18.4)				
rs10776792_G>A	G	989 (97.7)	1174(98.5)	0.24	1.51(0.81-2.82)	NS	NS
	A	23 (2.3)	18 (1.5)				
	GG	484 (95.7)	578 (97)	0.24	1.46(0.77-2.75)	NS	NS
	AA+GA	22 (4.3)	18 (3.0)				
	GA+GG	(99.8)	(100)	0.93	Undefined	NS	NS
	AA	1 (0.2)	0 (0.0)				

Table 20: Association of TSHβ with differentiated thyroid cancer risk. The table compares the TSHβ real-time PCR data for the 507 studied DTC patients versus 597 healthy controls. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature. f(%) gives the frequencies of the genotypes of alleles as a percentage. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively. OR, odds ratio; CI, confidence Interval; NS, non-significant; P, P value. ¶Each SNP was entered in a multivariate analysis including age, sex, and smoking.*P<0.05.

3.3.3.6. Paired Box gene 8

In the population study, 14 PAX8 SNPs, rs2241975_C>T, rs11123172_T>C, rs13015478_G>T, rs67776659_T>C, rs13007173_A>G, rs3738913_G>A, rs3748915_T>C, rs3748916_A>G, rs1049137_A>G, rs4849176_T>C, rs4849186_G>C, rs1479_T>G, rs1478_T>G and rs874898_G>C were selected from the initial screening and investigated (Table 21). Of these, the initial univariate analysis

suggested that the rs11123172_T>C was associated with the disease. This association appeared to be solely related to the dominant mode of inheritance rs11123172_CC+CT genotype, [1.34(1.5-1.72); p=0.02], and was retained [1.35(1.05-1.75); p=0.01] following the adjustment for confounders.

One variant rs67776659_C [0.39(0.23-0.66); p<0.0001] was protective (Table 21). Similarly, the protective relationship of this SNPs with disease was also retained [0.38(0.22-0.64); p<0.0001] following the adjustment for confounders, showing that the association was attributable to the dominant mode. Two other variants, rs1479_GG [0.35(0.15-0.82); p=0.01] and rs1478_GG [0.24(0.09-0.63); p=0.002] were potentially protective. In addition, the protective relationships were retained following the adjustment for the confounders, showing [0.34(0.14-0.79); p=0.01] for rs1479_GG and [0.26(0.08-0.59); p=0.003] for rs1478_GG.

One other variant rs13015478_(GT+TT) (p=0.06), was only weakly related to the disease in the univariate analyses, but lost the significance after the adjustments.

SNP ID	Geno-types	Patients (%)	Control (%)	Univariate Analysis P	Univariate Analysis OR(95%CI)	Multivariate Analysis ¶P	Multivariate Analysis OR(95%CI)
rs2241975_C>T	C	799(79.0)	971(81.6)	0.13	0.84(0.68-1.04)	NS	NS
	T	213(21.0)	219(18.4)				
	CC	319(63)	403(67.7)	0.11	0.81(0.63-1.04)	NS	NS
	CT+TT	187(37)	192(32.3)				
	CC+CT	480(94.9)	568(95.5)	0.74	0.88(0.50-1.52)	NS	NS
	TT	26(5.1)	27(4.5)				
rs11123172_T>C	T	587(58.0)	740(62.2)	0.05*	1.19(1.00-1.41)	0.067	0.84(0.71-1.01)
	C	425(42.0)	450(37.8)				
	TT	170(33.6)	241(40.5)	0.02*	1.34(1.5-1.72)	0.01*	1.35(1.05-1.75)
	CC+CT	336(66.4)	354(59.5)				
	CT+TT	417(82.4)	499(83.9)	0.57	1.10(0.80-1.52)	NS	NS
	CC	89(17.6)	96(16.1)				
rs13015478_G>T	G	629(62.2)	774(65.0)	0.17	0.88(0.74-1.05)	NS	NS
	T	383(37.8)	416(35.0)				
	GG	196(38.7)	265(44.5)	0.06	0.78(0.62-1.00)	0.08	0.80(0.62-1.03)
	GT+TT	310(61.3)	330(55.5)				
	GG+GT	433(85.6)	509(85.5)	1.0	1.00(0.71-1.40)	NS	NS
	TT	73(14.4)	86(14.5)				

SNP ID	Geno-types	Patients (%)	Control (%)	Univariate Analysis		Multivariate Analysis	
				P	OR(95%CI)	¶P	OR(95%CI)
rs6776659_T>C	T	992(98.0)	1136(95.1)	0.000*	0.39(0.23-0.66)	0.000*	0.38(0.22-0.64)
	C	20(2.00)	58(4.9)				
	TT	486(96.0)	544(91.1)	0.002*	0.42(0.24-0.71)	0.001*	0.40(0.23-0.70)
	TC+CC	20(4.0)	53(8.9)				
	TT+TC	506(100.0)	592(99.2)	0.06	0.000(0.000-)	NS	NS
	CC	0(0.0)	5(0.8)				
rs13007173_A>G	A	853(84.3)	1034(86.7)	0.11	0.82(0.64-1.04)	NS	NS
	G	159(15.7)	158(13.3)				
	AA	362(71.5)	454(76.2)	0.09	0.78(0.60-1.03)	NS	NS
	AG+GG	144(28.5)	142(23.8)				
	AA+AG	491(97.0)	580(97.3)	0.92	0.90(0.44-1.84)	NS	NS
	GG	15(3.0)	16(2.7)				
rs3738913_G>A	G	847(83.5)	1017(85.5)	0.23	1.15(0.92-1.46)	NS	NS
	A	167(16.5)	173(14.5)				
	GG	356(70.2)	441(74.1)	0.16	1.21(0.93-1.58)	NS	NS
	GA+AA	151(29.8)	154(25.9)				
	GG+GA	491(96.8)	576(96.8)	0.97	0.98(0.50-1.94)	NS	NS
	AA	16(3.2)	19(3.2)				
rs3748915_T>C	T	777(76.8)	941(79.1)	0.21	1.14(0.93-1.39)	NS	NS
	C	235(23.2)	249(20.9)				
	TT	302(59.7)	379(63.7)	0.19	1.18(0.92-1.51)	NS	NS
	TC+CC	204(40.3)	216(36.3)				
	TT+TC	475(93.9)	562(94.5)	0.77	1.11(0.67-1.84)	NS	NS
	CC	31(6.1)	33(5.5)				
rs3748916_A>G	A	536(53.0)	626(52.5)	0.86	1.01(0.86-1.20)	NS	NS
	G	476(47.0)	566(47.5)				
	AA	143(28.3)	186(31.2)	0.31	0.86(0.67-1.12)	NS	NS
	AG+GG	363(71.7)	410(68.8)				
	AA+AG	393(77.7)	440(73.8)	0.15	1.23(0.93-1.62)	NS	NS
	GG	113(22.3)	156(26.2)				
rs1049137_A>G	A	792(78.1)	963(81.1)	0.09	0.83(0.67-1.02)	NS	NS
	G	222(21.9)	225(18.9)				
	AA	310(61.1)	394(66.3)	0.08	0.79(0.62-1.02)	NS	NS
	AG+GG	197(38.9)	200(33.7)				
	AA+AG	482(95.1)	569(95.8)	0.66	0.84(0.48-1.49)	NS	NS
	GG	25(4.9)	25(4.2)				
rs4849176_T>C	T	535(52.9)	623(52.4)	0.84	0.98(0.82-1.15)	NS	NS
	C	477(47.1)	567(47.6)				
	TT	144(28.5)	183(30.8)	0.44	1.11(0.86-1.44)	NS	NS
	CC+TC	362(71.5)	412(69.2)				
	TC+TT	391(77.3)	440(73.9)	0.22	0.83(0.63-1.10)	NS	NS
	CC	115(22.7)	155(26.1)				
rs4849186_G>C	G	823(81.3)	971(81.7)	0.84	1.02(0.82-1.27)	NS	NS
	C	189(18.7)	217(18.3)				
	GG	338(66.8)	403(67.8)	0.76	1.04(0.81-1.35)	NS	NS
	CC+GC	168(33.2)	191(32.2)				
	GC+GG	485(95.8)	568(95.6)	0.97	0.94(0.52-1.70)	NS	NS
	CC	21(4.2)	26(4.4)				

SNP ID	Geno-types	Patients (%)	Control (%)	Univariate Analysis		Univariate Analysis	
				P	OR(95%CI)	P	OR(95%CI)
rs1479_T>G	T	905(89.3)	1036(86.8)	0.07	1.27(0.97-1.64)	NS	NS
	G	109(10.7)	158(13.2)				
	TT	405(79.9)	462(77.4)	0.31	0.86(0.64-1.15)	NS	NS
	TG+GG	102(20.1)	135(22.6)				
	TT+TG	500(98.6)	574(96.1)	0.01*	0.35(0.15-0.82)	0.01*	0.34(0.14-0.79)
	GG	7(1.4)	23(3.9)				
rs1478_T>G	T	901(88.9)	1031(86.3)	0.07	1.26(0.98-1.62)	NS	NS
	G	113(11.1)	163(13.7)				
	TT	399(78.7)	458(76.7)	0.43	0.89(0.67-1.18)	NS	NS
	TG+GG	108(21.3)	139(23.3)				
	TT+TG	502(99.0)	573(96.0)	0.002*	0.24(0.09-0.63)	0.003 *	0.26(0.08-0.59)
	GG	5(1.0)	24(4.0)				
rs874898_G>C	G	777(76.6)	911(76.3)	0.85	1.01(0.83-1.24)	NS	NS
	C	237(23.4)	283(23.7)				
	GG	296(58.4)	361(60.5)	0.48	1.09(0.86-1.4)	NS	NS
	GC+GG	211(41.6)	236(39.5)				
	GG+GC	481(94.9)	550(92.1)	0.07	0.63(0.39-1.04)	NS	NS
	CC	26(5.1)	47(7.9)				

Table 21. Association of PAX8 variants with differentiated thyroid cancer risk. The table compares the PAX8 real-time PCR data for the 507 studied DTC patients versus 597 healthy controls. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature. f(%) gives the frequencies of the genotypes of alleles as a percentage. OR, odds ratio; CI, confidence Interval; NS, non-significant; P, P value. [1]Each SNP was entered in a multivariate analysis including age, sex, and smoking.*P<0.05.

3.3.3.7. Sodium Iodide Symporter

Five of the discovered SNPs from the initial screening, rs4808708_G>A, rs4808709_A>G, and rs7250346_C>G, rs45602038 _C>T, and rs12327843_T>C were also tested for their association in the large population with disease in the NIS gene (Table 22). However, one variant, rs45602038 _C>T exon 13 (p.cys542cys) was related to DTC. This association appeared to be solely related to the recessive mode of inheritance rs45602038 _TT genotype, [3.66(1.62-8.21); p=0.001], and was retained [3.50(1.54-7.80); p=0.003] following the adjustment for confounders.

SNP ID	Geno-types	Patients f(%)	Control f(%)	P	OR(95%CI)	P	OR(95%CI)
rs4808708 _G>A	G	687(68.0)	845(71.0)	0.14	1.15(0.96-1.38)	NS	NS
	A	323(32.0)	345(29.0)				
	GG	240(47.5)	307(51.6)	0.19	1.17(0.93-1.49)	NS	NS
	AA+AG	265(52.5)	288(48.4)				
	AG+GG	447(88.5)	538(90.4)	0.31	1.23(0.83-1.80)	NS	NS
	AA	58(11.5)	57(9.6)				
rs4808709 _A>G	A	687(67.8)	824(69.1)	0.51	0.93(0.78-1.12)	NS	NS
	G	327(32.2)	368(30.9)				
	AA	241(47.5)	291(48.8)	0.67	0.95(0.75-1.20)	NS	NS
	AG+GG	266(52.5)	305(51.2)				
	AA+AG	446(88.0)	533(89.4)	0.44	0.86(0.59-1.25)	NS	NS
	GG	61(12)	63(10.6)				
rs7250346 _C>G	C	665(65.7)	766(64.3)	0.50	1.06(0.89-1.27)	NS	NS
	G	347(34.3)	426(35.7)				
	CC	224(44.3)	257(43.1)	0.70	1.05(0.82-1.33)	NS	NS
	CG+GG	282(55.7)	239(56.9)				
	CC+CG	441(87.2)	509(85.4)	0.40	1.16(0.82-1.63)	NS	NS
	GG	65(12.8)	87(14.6)				
rs45602038 _C>T	C	950(93.7)	1155(96.7)	0.001*	1.99(1.33-2.99)	0.002*	1.90(1.26-2.88)
	T	64(6.3)	39(3.3)				
	CC	467(92.1)	566(94.8)	0.07	1.56(0.96-2.53)	NS	NS
	CT+TT	40(7.9)	31(5.2)				
	CC+CT	483(95.3)	589(98.7)	0.001*	3.66(1.62-8.21)	0.003*	3.50(1.54-7.80)
	TT	24(4.7)	8(1.3)				
rs12327843 _T>C	T	692(68.2)	827(69.3)	0.60	0.96(0.80-1.14)	NS	NS
	C	322(31.8)	367(30.7)				
	TT	244(48.1)	294(49.2)	0.71	1.05(0.82-1.32)	NS	NS
	TC+CC	263(51.9)	303(50.8)				
	TT+TC	448(88.4)	533(89.3)	0.63	1.09(0.75-1.59)	NS	NS
	CC	59(11.6)	64(10.7)				

Table 22. Association of NIS gene polymorphism with differentiated thyroid cancer risk .The table displays relationships of rs4808708 G>A, rs4808709 A>G, rs7250346 C>G, rs45602038 C>T and rs12327843T>C with thyroid cancer. ¶Each SNP was entered in a multivariate analysis including age, sex, and smoking

3.3.4. Conclusions

Put together 39 SNPs were evaluated for their association with thyroid cancer. Thereby, two SNPs rs1321108 A>G, rs1321109 A>G in the TSHβ gene, four SNPs rs11123172, rs67776659, rs1479 T>G, rs1478 T>G in the PAX8 gene, one rs945006 in the DIO3 and rs45602038 C>T in NIS gene revealed clearly definable relationships with the disease (Figure 18).

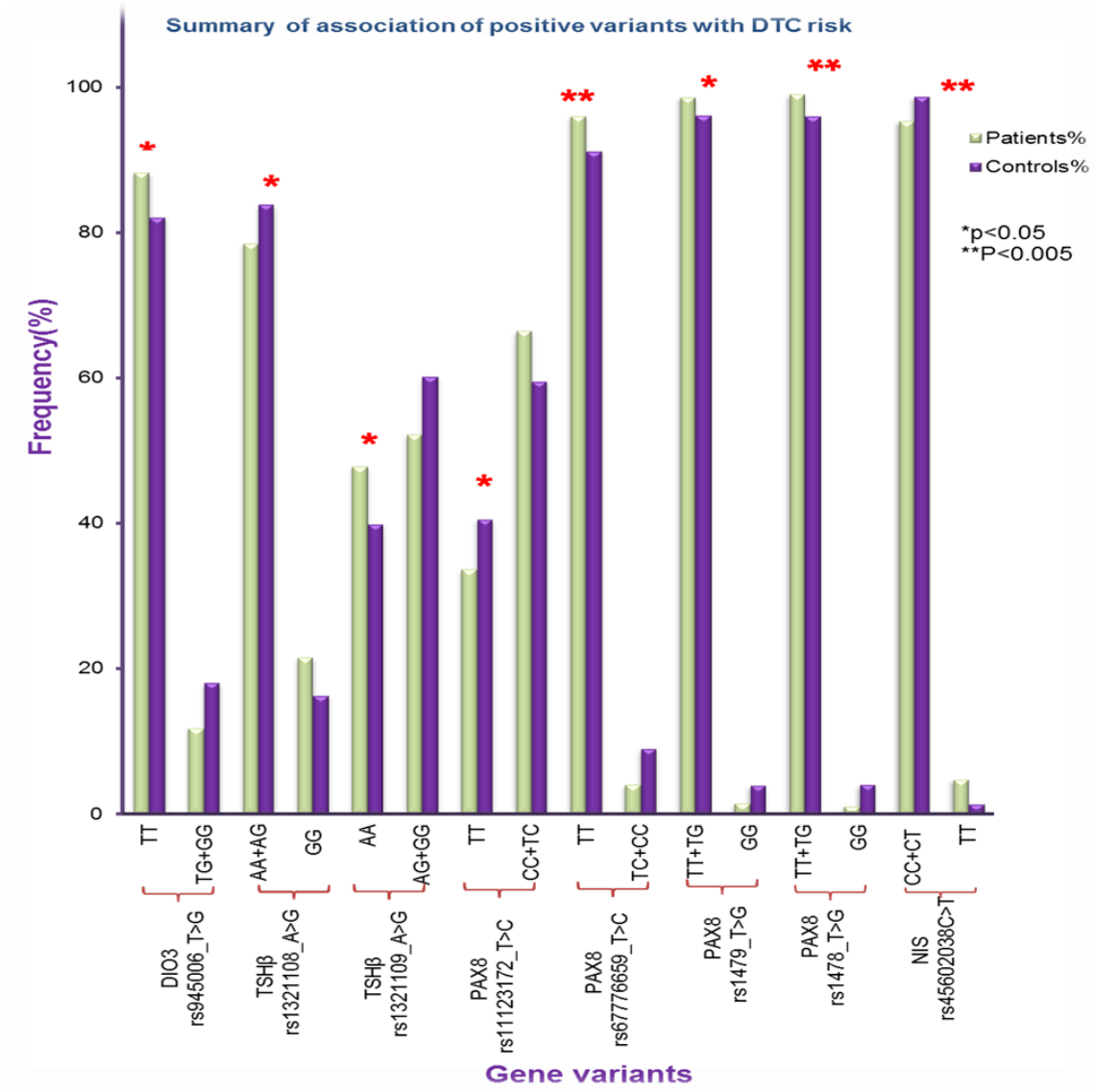


Figure 18: Summary-association of positive variants with differentiated thyroid cancer risk. The figure compares the positive real-time PCR data for the 507 studied DTC patients versus 597 healthy controls. SNP ID gives the single nucleotide identification number denoted with the "rs" nomenclature. f(%) gives the frequencies of the genotypes of alleles as a percentage. The letters A, C, G and T are the codes for the nucleotides adenine, cytosine, guanine and thymine respectively.

3.4. Association of haplotypes with differentiated thyroid cancer

3.4.1. Foreword

Since several SNPs were identified in the different genes, some of which were associated with disease in the study population, it was deemed worthwhile to establish whether or not haplotypes constructed from these variants may be of importance with regard to thyroid cancer manifestation. Hence, the same variants employed for the association studied in five of the genes (DIO1, DIO2, TSHβ, PAX8 and NIS) were subjected to haplotyping using the Haploview software. The DIO3 was excluded, since only one variant had been considered for the individual SNP association study.

3.4.2. Methodology

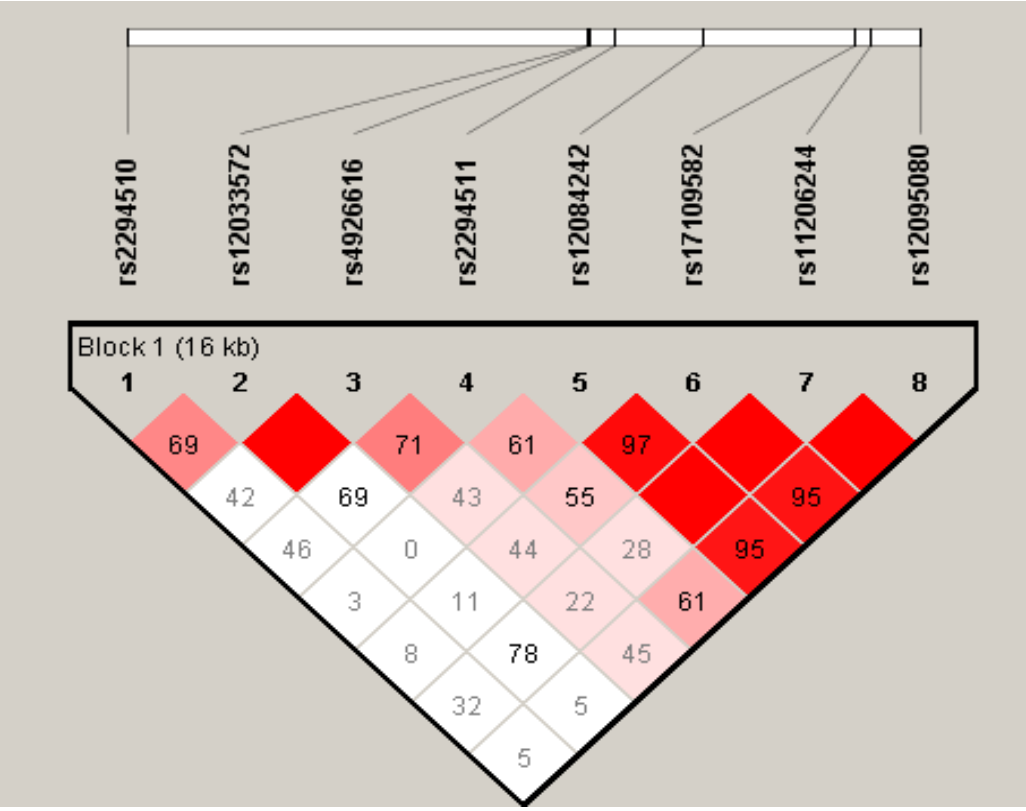
For haplotyping purposes, the maximum possible combinations of all studied SNPs within a particular gene were constructed. The most frequent of the created haplotypes was then used as a baseline for comparing the levels of the other constructs. Furthermore, various combinations of alleles were then derived from the different sets of blocks (nucleotide sequences) of SNPs involved in order to estimate the different haplotypes and to what extent they were linked to disease. The SNPs are denoted numerically and numbered sequentially as they appear on the chromosomes, and blocks represent the nucleotides within the haplotype sequences.

3.4.3. Results

3.4.3.1. Association of deiodinase haplotypes with thyroid cancer

Of the three deiodinases, only DIO1 and DIO2 could be subjected to haplotyping, since only one variant in DIO3 had been analysed for the association study. For the DIO1 eight SNPs rs2294510, rs12033572, rs4926616, rs2294511, rs12084242, rs1710958 rs11206244 and rs12095080 were available for the haplotyping. Figure 19 show the LD plot for these SNPs. In all, at least eleven common 8-nucleotide (8-mer) haplotypes

were described, whereby the most frequent haplotype CGCATTCA, with the frequency of 0.275 was employed as the baseline. None of the eleven haplotypes displayed any delineable relationship with disease. Further haplotypes were then constructed using various shorter sequences of up to three nucleotide blocks, but none of these showed any relationship with disease. These results appeared to be in conformation with the finding that none of the SNPs in the DIO1 was also individually linked to disease in any fashion.

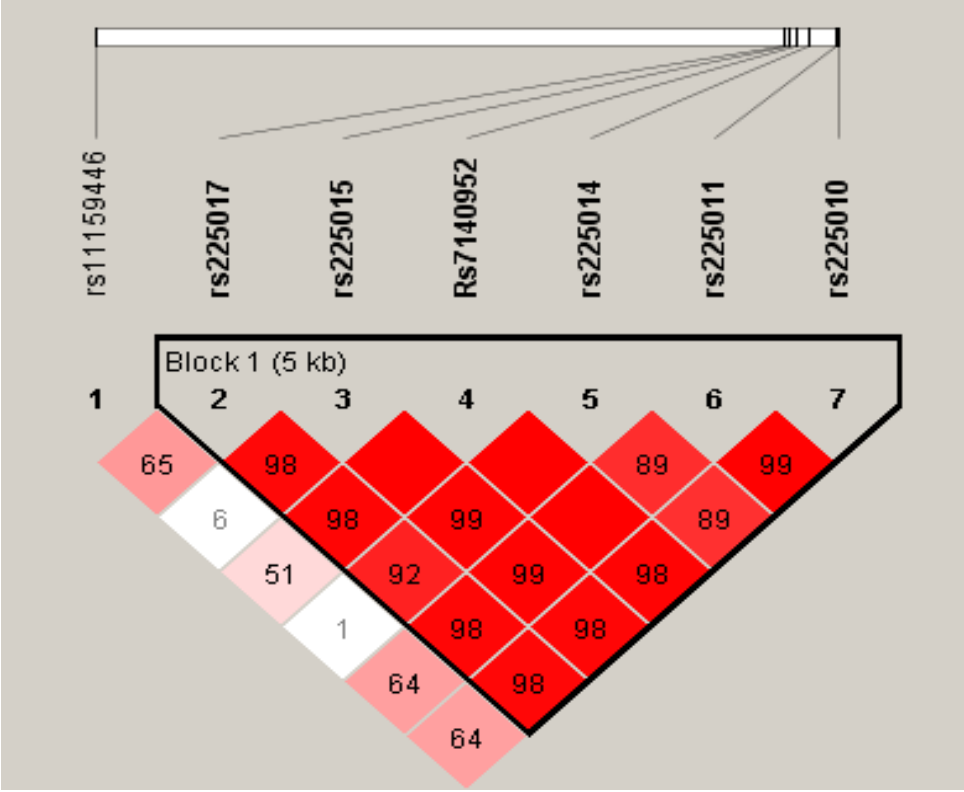


$D' = 0.002 - 1.00$; $r^2 = 0.0 - 0.92$

Figure 19: Linkage disequilibrium (LD) structure of the eight studied DIO1 single nucleotide polymorphisms. The values are expressed in percentage indicating that the higher the value the close the two variants are in LD with each other and smaller the value the further the SNPs are from each other. D' = coefficient of linkage disequilibrium; r = regression coefficient.

Similarly, for the DIO2 seven variants, rs11159446, rs225017, rs225015, rs7140952, rs225014, rs225011 and rs225010 were included in the test. Thereby the haplotype

GAGCTTC with a frequency of 0.563 was the most well-presented. Also, none of the seven primary 7-mer haplotypes showed any significant relationship with the disease. However, the assembling of different derivatives of these primary haplotypes yielded only one 4-mer GTGC (χ^2 = 4.613; p = 0.032) constructed from block 1-4 (rs11159446, rs225017, rs225015, rs7140952), that was associated with thyroid cancer, pointing to the region containing this block as a potential risk for the disease.



D'0.011-0.100; r^2 =0.0-0.98

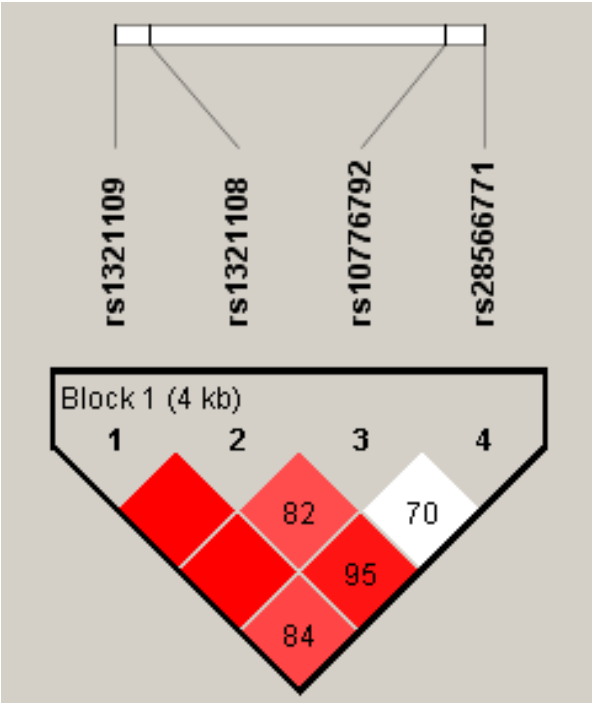
Figure 20: Linkage disequilibrium (LD) structure of the eight studied DIO2 single nucleotide polymorphisms. The values are expressed in percentage indicating that the higher the value the close the two variants are in LD with each other and smaller the value the further the SNPs are from each other. D' = coefficient of linkage disequilibrium; r = regression coefficient.

This block of nucleotides comprises a combination of variants in 3'UTR, indicating that this area of the gene in general is responsible for the observations. None of the other

shorter sequences was associated. Furthermore, like the DIO1 none of the DIO2 variants was individually linked to disease either. As stated above, DIO3 was not studied due to the fact that only one SNP was available from the association study.

3.4.3.2. Association of thyroid stimulating hormone-beta haplotypes with thyroid cancer

For the TSHβ analysis, haplotypes were constructed from the four studied SNPs, (1) rs1321109, (2) rs1321108, (3) rs10776792 and (4) rs28566771, arranged chronologically as the appear on the chromosome.



D' = 0.7-1.0; r^2 = 0.01-0.86

Figure 21: Linkage disequilibrium (LD) structure of the eight studied TSHβ single nucleotide polymorphisms. The values are expressed in percentage indicating that the higher the value the close the two variants are in LD with each other and smaller the value the further the SNPs are from each other. D' = coefficient of linkage disequilibrium; r = regression coefficient.

Figure 21 displays the linkage disequilibrium (LD) scores for the TSHβ studied SNPs whereby the rs28566771 showed the weakest LD properties. Of the (4-mer) haplotypes constructed, the most commonly distributed was AGGT (frequency = 0.411), which was employed as the baseline for comparing the haplotype frequencies. The analysis

indicated that the 4-mer haplotype GAGT (χ^2 = 11.21, p = 0.0008) conferred very significant risk for thyroid cancer (Table 23). Interestingly, a comparison of the two 3-mer derivatives, GAG (χ^2 = 6.31 p=0.012) and AGT (χ^2 = 11.71; p = 0.0006) indicated that the later was much more significantly related than the former, possibly pointing to the rs1321108G>A rs1077692G>A and rs28566771C>T as the primary entities involved in this causative relationship.

Block	Haplotype	Pooled	Control	Cases	χ^2	P-value
1-4	AGGT	0.411	0.442	0.385	7.36	0.0067*
	GAGT	0.024	0.012	0.034	11.21	0.0008*
1-3	AGG	0.421	0.453	0.395	7.45	0.0063*
	GAG	0.344	0.316	0.367	6.31	0.012
2-4	AGC	0.536	0.512	0.556	4.15	0.042
	GGT	0.41	0.442	0.383	7.78	0.0054*
	AGT	0.024	0.012	0.035	11.71	0.0006**

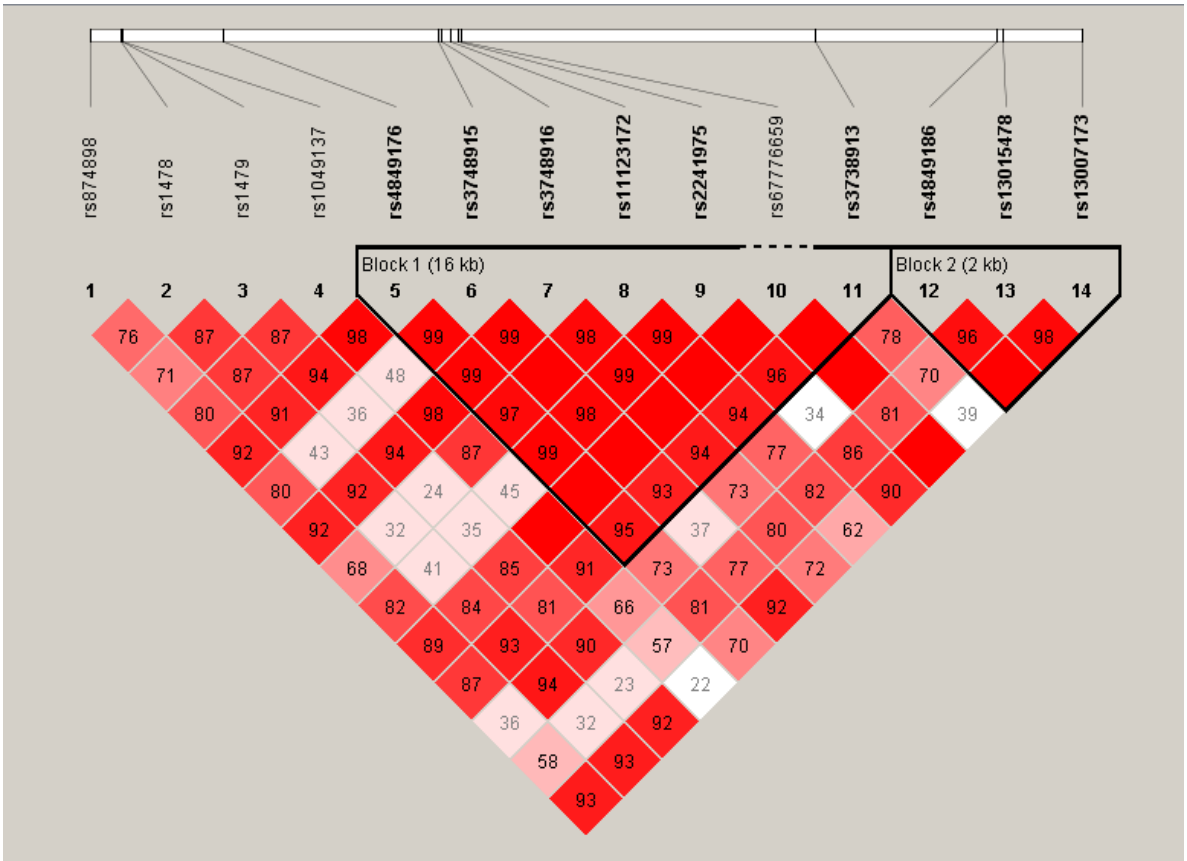
Table 23. Association of thyroid stimulating hormone -β haplotypes with thyroid cancer. The table shows selected haplotypes associated with disease. The most frequent 4-mer haplotype AGGT, (frequency = 0.411) was employed as the baseline to determine the relative effects of the other haplotypes. The studied SNPs are (1) rs1321109, (2) rs1321108, (3) rs10776792 and (4) rs28566771 arranged sequentially by their chromosomal positions, and blocks represent the range of variants constituting the respective haplotypes. *P<0.01; **P<0.001 by χ^2 test.

Besides, the data also shows that another 4-mer AGGT (χ^2 = 7.36; p=0.0067) was protective against disease. Similar comparison of the flanking 3-mers AGG (χ^2 = 7.45; p=0.0063) and GGT (χ^2 = 7.78; p=0.0054) did not seem to indicate any bias towards the one or the other. Notably, however, the difference between the potentially causative GAGT and protective AGGT lies in the complementary nucleotides of the variants rs1321109_A>G and rs1321108_G>A. This observation clearly points to the changes at these two loci as the determinant of these properties.

3.4.3.3. Association of paired box 8 haplotypes with thyroid cancer

For the paired box gene, we first constructed a haplotype from all 14 variants, as depicted in Figure 22. Several haplotypes were created using the haplotype GTTATTATCTGGGA (frequency =0.356) as the baseline exhibiting the highest prevalence of them all. In all 12 haplotypes were generated from the combinations of the 14 variants. However, none of these 14-mer haplotypes showed any delineable relationship with thyroid cancer. Further in-depth analysis showed that in fact, none of the haplotypes involving SNPs in block 1-4 showed any delineable characteristics. Differences were observed only after separating this region from the rest. Thus, significant associations only emerged following the creation of haplotypes within the block 5-14. In this region, the most abundant 10-mer was TTATCTGGGA with a frequency of 0.365, which was employed as the baseline. Interestingly, different combinations of these variants led to the identification of the 10-mer CCGCTTGGGA (χ^2=4.61; p=0.032) as the longest most significantly associated stretch of nucleotides with the disease, albeit in a protective fashion (Table 24).

Several shorter constructs derived from this haplotype showed consistently the same level of significance. The greatest level of significance was displayed by the 5-mer derivative TTGGG (χ^2 = 5.71; p = 0.017) constructed from the block 9-13, pointing therefore to the SNPs in this region as the likely basis for this protective action. Notably, this locus harbours the rs67776659T>C that exhibited strong protective properties against acquiring the disease.



$D' = 0.22\text{-}1.00$; $r^2 = 0.01\text{-}0.73$

Figure 22: Linkage disequilibrium (LD) structure of the eight studied PAX8 single nucleotide polymorphisms. The values are expressed in percentage indicating that the higher the value the close the two variants are in LD with each other and smaller the value the further the SNPs are from each other. D' = coefficient of linkage disequilibrium; r = regression coefficient.

Interestingly, on the other hand, haplotypes created from the variants in the first segment of the gene seemed to suggest that one 5-mer TTACT (block 2-6) and another 4-mer GTAC (block 2-5) were only weakly linked to thyroid cancer in a potentially causative fashion. These associations appear to depend on the fact that this block embraces the SNP rs1478T>G and rs1479 T>G, whose recessive modes of inheritance were also significantly associated with disease. Put together, it would appear that changes in the first part of the gene might be related to potentially protective properties, while the later would be causative against disease.

Block	Haplotype	Pooled	Controls	Cases	χ^2	P-value
5-14	CCGCTTGGGA	0.015	0.018	0.005	4.61	0.032*
6-14	CGCTTGGGA	0.015	0.018	0.005	4.53	0.033*
5-13	CCGCTTGGG	0.015	0.019	0.005	5.07	0.024*
7-14	GCTTGGGA	0.015	0.018	0.005	4.70	0.030*
6-13	CGCTTGGG	0.016	0.019	0.005	4.99	0.025*
7-13	GCTTGGG	0.015	0.019	0.005	4.94	0.026*
8-14	CTTGGGA	0.014	0.017	0.005	4.27	0.039*
9-14	TTGGGA	0.015	0.018	0.005	4.67	0.031*
8-13	CTTGGG	0.016	0.019	0.005	5.26	0.023*
9-13	TTGGG	0.017	0.020	0.005	5.71	0.017**
2-6	TTACT	0.02	0.017	0.030	3.41	0.065
2-5	GTAC	0.01	0.008	0.017	3.31	0.069

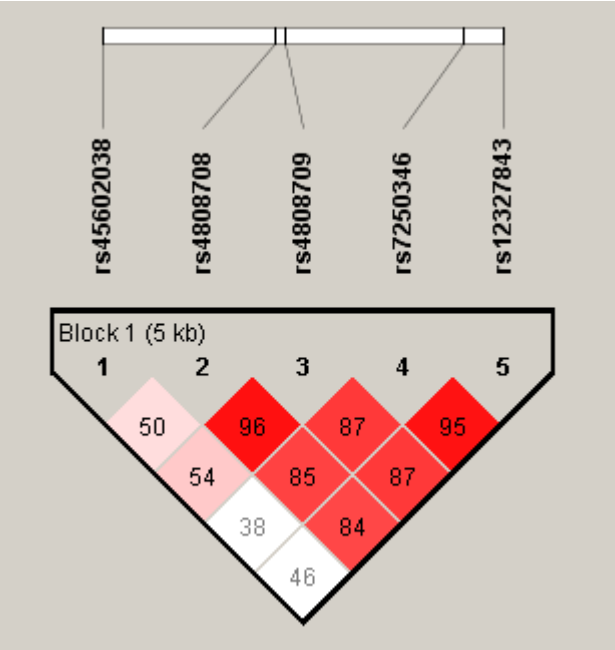
Table 24. Association of paired box 8 haplotypes with thyroid cancer. The table shows selected haplotypes associated with disease. The most frequent 14-mer haplotype GTTATTATCTGGGA (frequency= 0.356) was employed as the baseline to determine the relative effects of the other haplotypes. The studied SNPs are rs874898 denoted as 1, rs1478(2), rs1479(3), rs1049137(4), rs4849176(5), rs3748915(6), rs3748916(7), rs11123172(8), rs2241975(9), rs67776659(10), rs3738913(11), rs4849186(12), rs13015478(13), rs13007173(14) arranged sequentially by their chromosomal positions, and blocks represent the range of variants constituting the respective haplotypes. *P<0.05; **P<0.02by χ^2 test.

3.4.3.4. Association of the sodium iodide symporter haplotypes with thyroid cancer.

Five variants (1) rs45602038, (2) rs4808708, (3) rs4808709, (4) rs7250346 and (5) rs12327843 were employed to create haplotypes, with the most common 5-mer as the CGACT (frequency= 0.580) as the baseline for the NIS gene. Interestingly, of all the studied genes, the NIS exhibited the most significant relationships with differentiated thyroid cancer. To begin with, the 5-mer haplotype CGAGT (χ^2 = 10.98; p = 0.0009) and CGGGC (χ^2 = 5.70; p = 0.017) conferred significant risk for the disease.

This association trickled down through the 4-mer CGAG (χ^2 =13.25; p = 0.0003) displaying the most significant properties to the 3-mer GAG (χ^2 =11.80; p= 0.0006) which was equally strongly linked to the disease. Comparison of the flanking derivatives of the 5-mer CGAGT, showed that the 3-mer CGA (χ^2 = 4.04; p = 0.045) constructed from block 1-3 was a lot weaker than that of the AGT (χ^2 = 6.73; p =

0.0095), from the other end of the gene. Put together this data implicate the three nucleotide changes at the rs4808708, rs4808709 and rs7250346 loci (block 2-4) as the core for this relationship.



$D' = 0.38\text{-}0.96$; $r^2 = 0.01\text{-}0.88$

Figure 23. Linkage disequilibrium (LD) structure of the eight studied NIS single nucleotide polymorphisms. The values are expressed in percentage indicating that the higher the value the close the two variants are in LD with each other and smaller the value the further the SNPs are from each other. D' = coefficient of linkage disequilibrium; r = regression coefficient.

Interestingly, apart from the potentially causative CGAGT and its other similar derivatives, we established another 5-mer haplotype TGACT ($\chi^2 = 9.38$; $p = 0.0022$) that was protective against the disease. This haplotype differ from the potentially causative at the rs45602038 and rs7250346 loci. Closer analysis shows further that, while the smaller derivatives of the CGGGC do not show any difference from the mother haplotype, those of the TGACT actually show interesting differences. Briefly, the TGA ($\chi^2 = 8.90$; $p = 0.0028$) retained the level of significance, suggesting that the first nucleotides are important.

Block	Haplotype	Pooled	Controls	Cases	χ^2	P-value
1-5	CGAGT	0.035	0.021	0.048	10.98	0.0009**
	TGACT	0.035	0.048	0.024	9.38	0.0022**
	CGGGC	0.016	0.009	0.022	5.70	0.017
1-4	CGAG	0.054	0.035	0.070	13.25	0.0003**
	TGAC	0.036	0.049	0.024	9.80	0.0017**
	CGGG	0.019	0.012	0.025	4.67	0.0307
2-5	GAGT	0.038	0.025	0.048	8.45	0.0037*
	GGGC	0.016	0.009	0.022	5.72	0.017
	GAGC	0.02	0.014,	0.025	3.04	0.081
2-4	GAG	0.058	0.039	0.074	11.80	0.0006**
	GGG	0.019	0.012	0.025	4.68	0.031
1-3	CGA	0.638	0.616	0.657	4.04	0.045
	TGA	0.039	0.053	0.028	8.90	0.0028*
	CGG	0.019	0.012	0.025	4.61	0.032
3-5	AGT	0.039	0.028	0.049	6.73	0.0095*
	AGC	0.022	0.016	0.027	3.04	0.081

Table 25. Sodium iodide symporter haplotype associations. The table shows selected haplotypes associated with disease. The most frequent 5-mer haplotype CGACT (frequency = 0.580) was employed as the baseline to determine the relative effects of the other haplotypes. The studied SNPs are (1) rs45602038, (2) rs4808708, (3) rs4808709, (4) rs7250346 and (5) rs12327843 arranged sequentially by their chromosomal positions, and blocks represent the range of variants constituting the respective haplotypes. *P<0.01; **P<0.005 by χ^2 test.

3.4.4. Conclusion

In this part of the study, an important question was addressed, as to whether or not haplotypes constructed from the various SNPs identified in the Saudi population might be of importance with respect to the thyroid disease manifestation. Hence, haplotypes were created using various combinations of variants that had been included in the SNP association study in the earlier part of the study. Interestingly, several haplotypes of interest were identified in at least four of the five investigated genes, TSHβ, PAX8, NIS and the DIO2.

The observations unequivocally implicate some of the studied genes in thyroid cancer at haplotype level. Thus, among the DIOs, one haplotype in DIO2 was implicated in disease, while none of the DIO1 sequences was linked at all to disease. By far the most explicit relationships were observed in the relationships for the TSHβ and NIS genes

with the disease. Thus, at least one haplotype created from the maximum possible number of SNPs involved conferred risk for thyroid cancer. Moreover, not only did the study reveal potentially causative haplotypes, but the results seemed also to identify the important regions of the gene that were responsible for these actions, as indicated by clear difference between the causative and protective effects in the complementary nucleotides. Furthermore, the results also reflected on the trends observed for the individual variants. On the other hand, however, the significance levels displayed for the haplotype associations were distinctively greater than those shown by the individual variants, a point discussed further below.

3.5. Relationship of gene variants with L-thyroxine dose requirement

3.5.1. Foreword

As stated previously, the third aim of the study was to evaluate the possibility that the polymorphisms in the selected genes may play an important role in the way in which different patients respond to therapy of differentiated thyroid cancer (DTC) with L-thyroxine. Since 39 SNPs had been evaluated for their association with disease, it was thought worthwhile to focus on the same variants for this part of the study. Furthermore, the study was directed at the same patient population involved in the disease association study. Accordingly 453 out of the 507 DTC patients were included in the analysis. Fifty four patients were excluded from the analysis because they failed to meet the targeted therapeutic range of TSH and/or FT4 levels [(TSH $\leq$ 4.3 mU/L); (FT4 = 12.0-22.0 pmol/L)] suggesting non-compliance with thyroxine intake or inappropriate dose prescription. The analysis was performed in two stages. First stage involved whole patient group (ALL) in order to test whether this was general phenomenon with respect to patient response to drug therapy. In the second part, the patients were grouped based on their TSH level into the near suppressed (NSG) (0.1mUI$\leq$TSH<0.5 mUI) and suppressed (SG) (TSH<0.1 mUI) TSH groups to test whether such associations may be related to certain delineable levels of dose required.

3.5.2. Methodology

The analysis for the association of genotypes and thyroxine dose requirement for the case-control study was done by the Taqman chemistry using the ABI Prism 7900HT Sequence Detection System (Applied Biosystems Inc. CA, USA) as described in methodology (section 2.5.).

3.5.3. Results

3.5.3.1. Population Characteristics

The demographic data of patients enrolled in the thyroxine dose association study are displayed in Table 26.

	All	Male	Female
N	453	82/453 (18.1%)	371/453 (81.9%)
Age(years)	45.57±12.90	47.74±14.57	45.09±12.47
BMI	30.34±6.57	28.11±5.72	30.84±6.65
TSH(mU/l)	0.16±0.44	0.21±0.46	0.15±0.43
FT4(pmol/l)	20.70±1.70	20.69±1.82	20.70±1.67
L-T4 dose(µg/kg)	2.05±0.45	2.09±0.51	2.04±0.44

Table 26. Demographic data of individuals involve in the thyroxine dose association study. N, number of individuals in the group; BMI, body mass index; TSH, thyroid stimulating hormone; FT4, free thyroxine level; L-T4; Dose is give as µg/kg

As show in Figure 24, the distribution of the thyroxide doses among the studies patients followed near normal distibution.

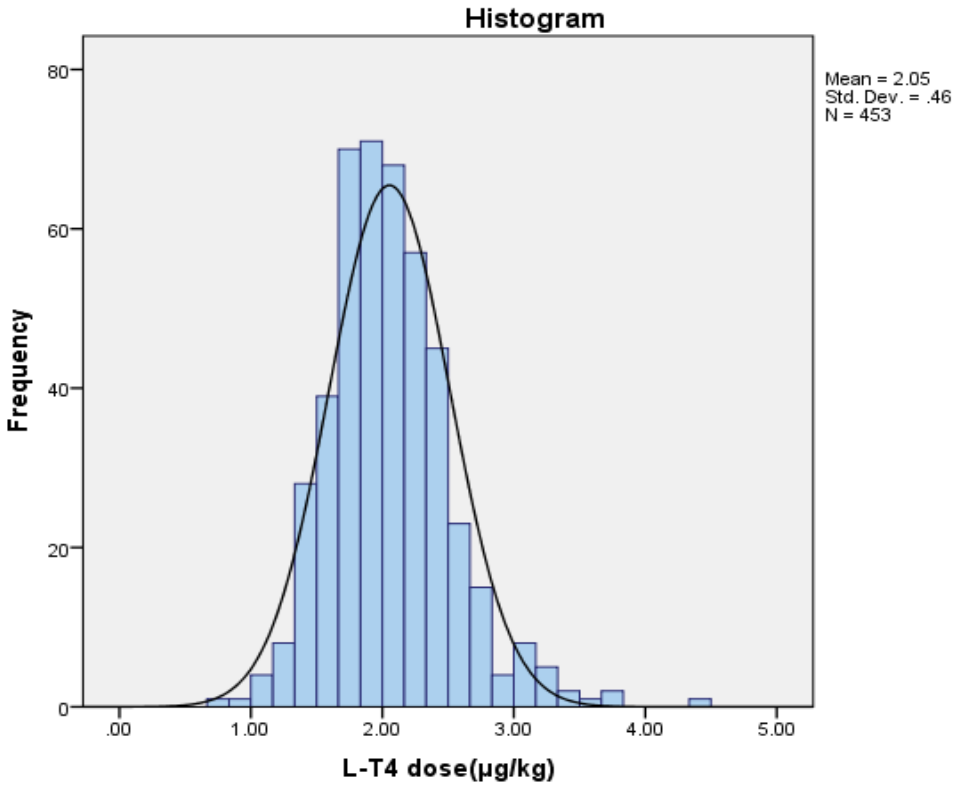


Figure24. Distribution of L-T4 dose among differentiated thyroid cancer patients.

Analysis of the relationship between thyroxine dose requirement and patient response to the therapy was then done in 453 patients. The same variants that were tested for their association with disease were also engaged in this analysis. Thus, analysis was performed first for the whole patient group (ALL), and then patients were grouped based on their TSH level into the near suppressed (NSG) (0.1mU/l≤TSH<0.5mU/l) and suppressed (SG) (TSH<0.1mU/l) TSH groups. Similar to the cancer association study, the same 39 variants were studied for their possible influence on patient thyroxine dose requirement variability.

3.5.3.2. Iodothyronine Deiodinase type I

First, several variants of the three deiodinases were evaluated for their role in thyroxine dose requirements. Eight SNPs rs12033572_G>C, rs12084242_T>A, rs2294510_C>T, rs2294511_A>T, rs12095080_A>G, rs11206244_C>T, rs4926616_C>T, rs17109582_T>C were chosen from the iodothyronine deiodinase type I (DIO1). Only one variant, rs12095080_GG (p=0.02) was associated with higher thyroxine dose requirement, while three others, rs12084242_AA (p=0.07), rs11206244_TT (p=0.06) and rs17109582_C (p=0.06) showed borderline association in the ALL group.

None of the variants was related to the dose requirement in the NSG group. Interestingly, however, rs12084242_AA (p=0.01) and rs17109582_CC (p=0.03) turned significantly associated with higher dose requirement, while the rs12095080_GG (p=0.01) retained its significant association in the SG (Table 27).

SNP ID	Geno-type	All (TSH<4.3)			NSG(0.1≤TSH<0.5)			SG(TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
rs12033572_G>C	G	892	2.05±0.45	0.59	258	2.00±0.42	0.52	575	2.09±0.46	0.65
	C	14	1.98±0.53		6	1.89±0.31		7	2.01±0.70	
	GG	14	1.98±0.53	0.58	126	2.01±0.42	0.51	284	2.09±0.45	0.65
	GC+CC	439	2.05±0.45		6	1.89±0.31		7	2.01±0.70	
	GG+GC	453	2.05±0.45	NS	132	2.00±0.42	NS	291	2.09±0.46	NS
	CC	0	0.00±0.00		0	0.00±0.00		0	0.00±0.00	
rs12084242_T>A	T	775	2.04±0.45	0.17	224	2.01±0.43	0.47	496	2.07±0.44	0.06
	A	129	2.10±0.49		40	1.96±0.33		84	2.17±0.55	
	TT	332	2.04±0.45	0.34	95	2.02±0.45	0.52	212	2.06±0.44	0.19
	AA+TA	120	2.08± 0.47		37	1.96±0.34		78	2.14±0.52	
	TA+TT	443	2.04±0.45	0.07	129	2.00±0.42	0.62	284	2.08±0.45	0.01*
	AA	9	2.32±0.75		3	1.88±0.04		6	2.54±0.86	
rs2294510_C>T	C	881	2.05±0.46	0.47	253	2.00±0.42	0.55	569	2.09±0.46	0.52
	T	25	1.98±0.40		11	1.93±0.24		13	2.00±0.51	
	CC	429	2.05±0.46	0.47	122	2.01±0.43	0.52	278	2.09±0.46	0.51
	CT+TT	24	1.98±0.41		10	1.92±0.25		13	2.00±0.51	
	CC+CT	452	2.05±0.46	0.94	131	2.00±0.42	0.97	291	2.09±0.46	NS
	TT	1	2.02±0.00		1	2.02±0.00		0		
rs2294511_A>T	A	569	2.04±0.44	0.38	168	1.98±0.41	0.43	367	2.08±0.45	0.71
	T	335	2.07±0.47		96	2.03±0.44		213	2.10±0.47	
	AA	178	2.04±0.43	0.61	54	1.96±0.38	0.42	117	2.08±0.45	0.84
	AT+TT	274	2.06±0.47		78	2.02±0.45		173	2.09±0.47	
	AA+AT	391	2.04±0.45	0.30	114	1.99±0.42	0.70	250	2.08±0.46	0.67
	TT	61	2.11±0.48		18	2.04±0.42		40	2.12±0.49	
rs12095080_A>G	A	761	2.03±0.44	0.04*	222	1.99±0.42	0.77	484	2.07±0.44	0.02*
	G	131	2.12±0.50		38	1.97±0.34		88	2.19±0.55	
	AA	325	2.03±0.44	0.13	94	2.00±0.43	0.81	206	2.06±0.44	0.10
	AG+GG	121	2.10±0.48		36	1.98±0.35		80	2.16±0.53	
	AA+AG	436	2.04±0.44	0.02*	128	1.99±0.41	0.76	278	2.08±0.45	0.01*
	GG	10	2.37±0.72		2	1.91±0.01		8	2.48±0.77	
rs11206244_C>T	C	640	2.03±0.46	0.14	197	1.99±0.40	0.41	402	2.08±0.48	0.88
	T	260	2.08±0.43		67	2.04±0.46		174	2.09±0.42	
	CC	227	2.03±0.47	0.41	72	1.97±0.36	0.35	141	2.10±0.50	0.63
	CT+TT	223	2.07±0.44		60	2.04±0.48		147	2.07±0.43	
	CC+CT	413	2.04±0.46	0.06*	125	2.00±0.43	0.85	261	2.07±0.47	0.25
	TT	37	2.18±0.37		7	2.03±0.22		27	2.18±0.35	
rs4926616_C>T	C	521	2.04±0.44	0.44	156	1.98±0.41	0.26	335	2.08±0.46	0.82
	T	381	2.06±0.47		106	2.04±0.44		245	2.09±0.47	
	CC	148	2.03±0.41	0.48	44	1.95±0.35	0.38	96	2.07±0.44	0.69
	CT+TT	303	2.06±0.48		87	2.02±0.45		194	2.09±0.47	
	CC+CT	373	2.04±0.46	0.58	112	1.98±0.43	0.29	239	2.09±0.46	0.94
	TT	78	2.07±0.46		19	2.09±0.36		51	2.08±0.46	

SNP ID	Geno-type	All (TSH<4.3)			NSG(0.1≤TSH<0.5)			SG(TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
rs17109582_T>C	T	767	2.04±0.45	0.06	223	2.01±0.43	0.55	489	2.07±0.44	0.02*
	C	133	2.12±0.50		39	1.96±0.33		89	2.19±0.55	
	TT	329	2.03±0.45	0.13	95	2.01±0.45	0.61	209	2.05±0.43	0.06
	TC+CC	121	2.10±0.48		36	1.97±0.35		80	2.17±0.52	
	TT+TC	438	2.04±0.45	0.08	128	2.00±0.42	0.63	280	2.08±0.45	0.03*
	CC	12	2.28±0.68		3	1.88±0.04		9	2.41±0.75	

Table 27: Influence of DIO1 variants (rs12033572_G>C , rs12084242_T>A, rs2294510_C>T, rs2294511_A>T, rs12095080_A>G, rs11206244_C>T, rs4926616_C>T, rs17109582_T>C) on thyroxine dose requirement. N, number of patients in the group; Concentration is given as µg/kg body mass. NSG, near-suppressed group; SG, suppressed group. *P<0.05

3.5.3.3. Iodothyronine Deiodinase type II

On the other hand, analysis for the Iodothyronine Deiodinase type II (DIO2) showed no definable relationship for any of the SNPs with thyroxine dose requirement (Tables 28).

SNP ID	Geno-type	ALL(TSH<4.3)			NSG(0.1≤TSH<0.5)			SG(TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
rs225010_C>T	C	552	2.04±0.46	0.72	152	1.99±0.43	0.70	363	2.08±0.47	0.63
	T	354	2.06±0.44		112	2.01±0.40		219	2.10±0.44	
	CC	167	2.05±0.48	0.95	49	2.03±0.47	0.60	109	2.07±0.48	0.60
	CT+TT	286	2.05±0.44		83	1.99±0.39		182	2.10±0.45	
	CC+CT	385	2.04±0.46	0.54	103	1.98±0.41	0.21	254	2.08±0.47	0.82
	TT	68	2.08±0.44		29	2.09±0.44		37	2.10±0.41	
rs225011_T>C	T	551	2.04±0.46	0.70	153	1.99±0.43	0.61	361	2.08±0.47	0.67
	C	353	2.06±0.44		111	2.02±0.40		219	2.10±0.44	
	TT	167	2.05±0.48	0.96	49	2.03±0.47	0.60	109	2.07±0.48	0.61
	TC+CC	285	2.05±0.44		83	1.99±0.39		181	2.10±0.45	
	TT+TC	384	2.04±0.46	0.50	104	1.97±0.41	0.14	252	2.08±0.47	0.90
	CC	68	2.08±0.44		28	2.10±0.44		38	2.09±0.41	
rs225014_T>C	T	629	2.04±0.46	0.65	177	2.00±0.42	0.97	412	2.08±0.47	0.47
	C	275	2.06±0.44		87	2.00±0.41		168	2.11±0.44	
	TT	224	2.03 ±0.47	0.50	62	2.00±0.45	0.98	150	2.06±0.48	0.37
	TC+CC	228	2.06±0.44		70	2.00±0.39		140	2.11±0.45	
	TT+TC	405	2.05±0.45	0.89	115	2.00±0.41	0.92	262	2.09±0.47	0.94
	CC	47	2.04±0.47		17	2.01±0.47		28	2.09±0.43	

SNP ID	Geno-type	ALL(TSH<4.3) N	ALL(TSH<4.3) T4 dose	P	NSG(0.1≤TSH<0.5) N	NSG(0.1≤TSH<0.5) T4 dose	P	SG(TSH<0.1) N	SG(TSH<0.1) T4 dose	P
rs225015_G>A	G	643	2.04±0.46	0.40	181	2.00±0.42	0.87	421	2.07±0.47	0.36
	A	263	2.07±0.44		83	2.01±0.41		161	2.11±0.44	
	GG	230	2.03±0.47	0.39	65	1.99±0.44	0.83	153	2.06±0.48	0.33
	GA+AA	223	2.07±0.44		67	2.01±0.40		138	2.11±0.44	
	GG+GA	413	2.05±0.45	0.70	116	2.00±0.41	0.98	268	2.08±0.46	0.72
	AA	40	2.08±0.46		16	2.00±0.49		23	2.12±0.46	
rs225017_T>A	T	537	2.05±0.46	0.86	147	2.00±0.43	0.82	353	2.08±0.47	0.74
	A	369	2.05±0.45		117	2.01±0.41		229	2.09±0.45	
	TT	159	2.05±0.47	0.91	47	2.04±0.47	0.38	103	2.07±0.46	0.61
	TA+AA	294	2.05±0.45		85	1.98±0.39		188	2.10±0.46	
	TT+TA	378	2.04±0.46	0.65	100	1.97±0.42	0.18	250	2.09±0.47	0.95
	AA	75	2.07±0.43		32	2.09±0.42		41	2.08±0.41	
rs11159446_G>C	G	829	2.05±0.46	0.47	242	2.00±0.42	0.59	529	2.09±0.47	0.78
	C	75	2.01±0.43		22	1.95±0.44		51	2.07±0.40	
	GG	383	2.06±0.46	0.38	111	2.01±0.41	0.73	244	2.09±0.47	0.61
	CG+CC	69	2.00±0.44		21	1.97±0.45		46	2.05±0.40	
	GG+CG	446	2.05±0.46	0.75	131	2.00±0.42	0.35	285	2.08±0.46	0.55
	CC	6	2.11±0.45		1	1.61		5	2.21±0.42	
rs 140952_C>T	C	817	2.05±0.46	0.70	237	1.99±0.42	0.43	523	2.09±0.46	0.40
	T	89	2.03±0.42		27	2.06±0.38		59	2.04±0.44	
	CC	369	2.05±0.46	0.67	108	1.99±0.42	0.44	234	2.10±0.46	0.38
	CT+TT	84	2.03±0.43		24	2.06±0.40		57	2.04±0.45	
	CC+CT	448	2.05±0.46	0.95	129	2.00±0.42	0.80	289	2.09±0.46	0.93
	TT	5	2.06±0.10		3	2.06±0.03		2	2.06±0.20	

Table 28. Influence of DIO2 variants (rs225010 C>T, rs225011 T>C, rs225014 T>C, rs225014 T>C, rs225015 G>A, rs225017 T>A, rs11159446 G>C, rs140952C>T) on thyroxine dose requirement. TSH, thyroid-stimulating hormone; N, number of patients in the group; Concentration is given as μg/kg body mass. *P<0.05

3.5.3.4. Iodothyronine Deiodinase type III

The association study between the iodothyronine deiodinase type III (DIO3) variant rs945006 G>T and thyroxine dose requirement showed no relationship (Tables 29).

SNP ID	Geno-type	ALL(TSH<4.3) N	ALL(TSH<4.3) T4 dose	P	NSG(0.1≤TSH<0.5) N	NSG(0.1≤TSH<0.5) T4 dose	P	SG(TSH<0.1) N	SG(TSH<0.1) T4 dose	P
rs945006_G>T	G	850	2.05±0.46	0.68	249	1.99±0.42	0.34	545	2.09±0.46	0.87
	T	56	2.07±0.40		15	2.10±0.42		37	2.07±0.40	
	TT	401	2.05±0.46	0.86	117	1.99±0.42	0.32	257	2.09±0.47	0.70
	GT+GG	52	2.06±0.41		15	2.10±0.42		34	2.06±0.41	
	GG	4	2.26±0.10	0.36	132	2.00±0.42	NS	288	2.08±0.46	0.48
	TT+GT	449	2.05±0.46		0			3	2.27±0.12	

Table 29. Influence of DIO3, rs945006_G>T variants on thyroxine dose requirement. N, number of patients in the group; Concentration is given as μg/kg body mass. *P<0.05

3.5.3.5. Thyroid Stimulating Hormone-Beta

For the Thyroid Stimulating Hormone-Beta (TSHβ) gene, the same variants rs1321108_A>G, rs1321109_A>G, rs28566771_C>G and rs10776792_G>A evaluated in the cancer study were also assessed for the dose-dependence study.

SNP ID	Geno-type	All (TSH<4.3)			NSG(0.1≤TSH<0.5)			SG (TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
rs1321108_A>G	A	495	2.07±0.45	0.32	142	2.00±0.40	0.99	330	2.11±0.46	0.18
	G	409	2.03±0.45		122	2.00±0.44		250	2.06±0.45	
	AA	140	2.08±0.44	0.37	41	2.00±0.37	0.96	96	2.12±0.47	0.36
	AG+GG	312	2.04±0.46		91	2.00± 0.44		194	2.07±0.45	
	AA+AG	355	2.06±0.45	0.48	101	2.00±0.41	0.94	234	2.11±0.46	0.20
	GG	97	2.02±0.44		31	2.00±0.46		56	2.02±0.44	
rs1321109_A>G	A	622	2.04±0.45	0.32	176	2.01±0.44	0.68	405	2.06±0.45	0.059
	G	282	2.07±0.46		88	1.99±0.38		175	2.14±0.48	
	AA	220	2.04±0.47	0.60	63	2.02±0.47	0.65	144	2.05±0.47	0.16
	AG+GG	232	2.06±0.44		69	1.98±0.37		146	2.12±0.45	
	AA+AG	402	2.04±0.44	0.21	113	2.00±0.42	0.90	261	2.07±0.44	0.07
	GG	50	2.12±0.54		19	1.99±0.41		29	2.23±0.61	
rs28566771_C>G	C	492	2.07±0.45	0.17	142	2.00±0.40	0.91	326	2.12±0.46	0.06
	G	412	2.03±0.46		122	2.00±0.44		254	2.05±0.46	
	CC	137	2.09±0.44	0.20	41	2.00±0.37	0.99	93	2.14±0.47	0.17
	CG+GG	315	2.03±0.46		91	2.00± 0.44		197	2.06±0.46	
	CC+CG	355	2.06±0.46	0.36	101	2.00±0.41	0.87	233	2.11±0.46	0.09
	GG	97	2.01±0.45		31	1.99±0.46		57	1.99±0.46	
rs10776792_G>A	G	882	2.05±0.45	0.20	259	2.00±0.42	0.74	563	2.08±0.46	0.28
	A	22	2.17±0.49		5	2.06±0.11		17	2.21±0.56	
	GG	431	2.04±0.45	0.33	127	2.00±0.43	0.74	274	2.08±0.46	0.45
	GA+AA	21	2.14±0.49		5	2.06±0.11		16	2.17±0.56	
	GG+GA	451	2.05±0.45	0.11	132	2.00±0.42	NS	289	2.08±0.46	0.13
	AA	1	2.78		0			1	2.78	

Table 30: Influence of TSHβ variants rs1321108_(A>G); rs1321109_A>G, rs28566771_C>G, rs10776792_G>A on thyroxine dose requirement. N, number of patients in the group; Concentration is given as µg/kg body mass. TSH, thyroid-stimulating hormone; NSG, near-suppressed group; SG, suppressed group. *P<0.05

For both the "ALL" and "Near-suppressed" group analyses, none of the variants was related in any way to the dose requirement. However, for the suppressed group, two SNPs rs1321109_A>G (p=0.059) and rs28566771_C>G (p=0.06) showed a weak association, whereby the first one was related to an increase in thyroxine dose while the other SNP was related to lower thyroxine dose requirement (Table 30).

3.5.3.6. Paired Box gene 8

In the Paired Box gene 8 (PAX8) gene, fourteen SNPs were analyzed for their role in differential hormone dose requirement. Thereby, in the ALL group, two variants rs2241975_CT+TT (p=0.01) and rs4849186_CC (p=0.01) were associated with thyroxine dose requirement. Thereby, the first variant was associated with an increased dose requirement, while the second one linked to lower dose requirement. On the other hand, another rs3738913_A (p=0.07) was only weakly linked to decreased dose requirement. The association for rs4849186_CC was retained in the NSG (p=0.03) group, but became weaker in the SG (p=0.06). Furthermore, rs1478_GG was associated with lower thyroxine dose requirement in NSG. Otherwise, none of the other variants showed definable relationship in the NSG

SNP ID	Geno-type	All (TSH<4.3)			NSG (0.1≤TSH<0.5)			SG (TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
rs2241975_C>T	C	712	2.03±0.44	0.03*	210	1.99±0.39		454	2.07±0.45	
	T	192	2.11±0.51		54	2.05±0.51	0.37	126	2.15±0.50	0.08
	CC	285	2.01±0.41	0.01*	85	1.98±0.36		181	2.04±0.42	
	CT+TT	167	2.12±0.52		47	2.04±0.51	0.41	109	2.16±0.51	0.03*
	CC+CT	427	2.05±0.46		125	2.00±0.41		273	2.09±0.46	
	TT	25	2.06±0.46	0.94	7	2.08±0.55	0.60	17	2.08±0.42	0.97
rs11123172_T>C	T	526	2.05±0.43	0.77	153	1.99±0.37	0.50	346	2.09±0.44	0.84
	C	378	2.04±0.49		111	2.02±0.48		234	2.08±0.49	
	TT	152	2.07±0.41	0.50	39	1.99±0.36	0.92	108	2.10±0.42	
	TC+CC	300	2.04±0.48		93	2.00±0.44		182	2.08±0.49	0.70
	TT+TC	374	2.05±0.44		114	1.98±0.37		238	2.08±0.46	0.9
	CC	78	2.06±0.53	0.74	18	2.12±0.66	0.21	52	2.09±0.49	

SNP ID	Geno-type	All (TSH<4.3)			NSG (0.1≤TSH<0.5)			SG (TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
rs13015478_G>T	G	571	2.05±0.43	0.89	165	2.00±0.37	0.85	374	2.08±0.44	0.52
	T	333	2.05±0.49		99	1.99±0.49		206	2.10±0.49	
	GG	180	2.03±0.40	0.52	46	2.03±0.34	0.51	126	2.04±0.40	0.18
	GT+TT	272	2.06±0.49		86	1.98±0.46		164	2.12±0.50	
	GG+GT	391	2.05±0.45	0.51	119	1.99±0.38	0.53	248	2.09±0.46	0.51
	TT	61	2.01±0.50		13	2.07±0.69		42	2.04±0.44	
rs67776659_T>C	T	885	2.05±0.46	0.47	260	2.00±0.42	0.84	567	2.09±0.47	0.32
	C	19	1.97±0.23		4	2.04±0.22		13	1.96±0.26	
	TT	433	2.05±0.46	0.46	128	2.00±0.42	0.84	277	2.09±0.47	0.31
	TC+CC	19	1.97±0.23		4	2.04±0.22		13	1.96±0.26	
	TT+TC	452	2.05± 0.46	NS	132	2.00±0.42		290	2.09±0.46	NS
	CC	0	-		0	-		0	-	
rs13007173_A>G	A	755	2.04±0.46	0.22	223	2.01±0.43	0.25	482	2.07±0.46	0.10
	G	149	2.09±0.43		41	1.93±0.34		98	2.16±0.46	
	AA	317	2.03±0.46	0.21	94	2.02±0.44	0.31	201	2.06±0.46	0.10
	AG+GG	135	2.09±0.44		38	1.94±0.35		89	2.15±0.47	
	AA+AG	438	2.05±0.46	0.72	129	2.01±0.42	039	281	2.08±0.46	0.52
	GG	14	2.09±0.37		3	1.79±0.28		9	2.18±0.37	
rs3738913_G>A	G	765	2.06±0.45	0.07	216	2.00±0.41	0.80	504	2.10±0.46	0.06*
	A	141	1.99±0.45		48	2.01±0.47		78	1.99±0.45	
	GG	324	2.07±0.46	0.08	88	1.98±0.39	0.41	220	2.12±0.47	0.05*
	GA+AA	129	1.99±0.44		44	2.04±0.47		71	1.99±0.42	
	GG+GA	441	2.05±0.45	0.43	128	2.01± 0.42	0.13	284	2.09±0.45	0.65
	AA	12	1.95±0.64		4	1.69±0.31		7	2.01±0.78	
rs3748915_T>C	T	691	2.04±0.44	0.20	206	1.99±0.39	0.68	441	2.07±0.45	0.20
	C	213	2.08±0.51		58	2.02±0.51		139	2.13±0.51	
	TT	269	2.02±0.41	0.41	81	1.99±0.35	0.80	171	2.05±0.42	0.17
	TC+CC	183	2.09±0.51		51	2.01±0.51		119	2.13±0.51	
	TT+CT	422	2.05±0.45	0.88	125	2.00± 0.41	0.60	270	2.08±0.46	0.70
	CC	30	2.06±0.50		7	2.08±0.55		20	2.12±0.50	
rs3748916_A>G	A	479	2.06±0.43	0.51	139	1.98±0.37	0.46	316	2.10±0.45	0.55
	G	425	2.04±0.48		125	2.02±0.47		264	2.07±0.48	
	AA	128	2.07±0.42	0.61	31	1.97±0.33	0.61	93	2.09±0.44	0.84
	AG+GG	324	2.04±0.47		101	2.01±0.44		197	2.08±0.47	
	AA+AG	351	2.06±0.44	0.57	108	1.99±0.38	0.43	223	2.10±0.45	0.44
	GG	101	2.03±0.51		24	2.06±0.58		67	2.05±0.49	

SNP ID	Geno-type	All (TSH<4.3)			NSG (0.1≤TSH<0.5)			SG (TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
rs1049137_A>G	A	710	2.06±0.45	0.16	201	1.99±0.40	0.41	469	2.10±0.46	0.19
	G	196	2.01±0.48		63	2.04±0.48		113	2.03±0.47	
	AA	279	2.08±0.45	0.11	75	1.97±0.38	0.28	193	2.11±0.47	0.21
	AG+GG	174	2.01±0.45		57	2.05±0.47		98	2.04±0.44	
	AA+AG	431	2.05±0.44	0.79	126	2.00±0.41	0.85	276	2.09±0.45	0.52
	GG	22	2.02±0.64		6	1.97±0.69		15	2.01±0.66	
Rs489176_TC	T	478	2.06±0.43	0.66	140	1.98±0.37	0.38	315	2.10±0.45	0.55
	C	426	2.05±0.48		124	2.02±0.47		265	2.07±0.48	
	TT	129	2.06±0.41	0.75	32	1.95±0.34	0.47	93	2.09±0.43	0.91
	CC+TC	323	2.05±0.47		100	2.02±0.44		197	2.08±0.47	
	TT+TC	349	2.05±0.44	0.70	108	1.99±0.38	0.43	222	2.10±0.45	0.39
	CC	103	2.03±0.51		24	2.06±0.58		68	2.04±0.49	
rs4849186_G>C	G	744	2.06±0.44	0.31	217	2.00±0.41	0.98	482	2.09±0.45	0.42
	C	160	2.02±0.51		47	2.00±0.47		98	2.05±0.52	
	GG	309	2.05±0.43	0.86	89	1.98±0.39	0.44	85	2.08±0.52	0.92
	CC+GC	143	2.04±0.51		43	2.04±0.47		205	2.09±0.43	
	GC+GG	435	2.06±0.45	0.01*	128	2.01±0.42	0.03*	277	2.10±0.46	0.06*
	CC	17	1.78±0.44		4	1.55±0.17		13	1.85±0.48	
rs1479_T>G	T	808	2.04±0.45	0.15	238	2.00±0.40	0.97	516	2.08±0.46	0.15
	G	98	2.11±0.49		26	2.00±0.54		66	2.16±0.47	
	TT	362	2.03±0.45	0.14	108	2.00±0.38	0.90	230	2.06±0.46	0.14
	TG+GG	71	2.11±0.49		24	2.01±0.56		61	2.16±0.47	
	TT+TG	446	2.05±0.46	0.82	130	2.00±0.42	0.77	286	2.08±0.46	0.73
	GG	7	2.09±0.48		2	1.92±0.02		5	2.16±0.57	
rs1478_T>G	T	806	2.04±0.45	0.27	240	2.01±0.40	0.28	513	2.07±0.45	0.08
	G	100	2.10±0.51		24	1.91±0.55		69	2.18±0.50	
	TT	358	2.04±0.44	0.26	109	2.01±0.39	0.49	226	2.06±0.45	0.12
	TG+GG	95	2.10±0.50		23	1.95±0.53		65	2.16±0.49	
	TT+TG	448	2.05±0.45	0.74	131	2.01±0.42	0.04*	287	2.08±0.46	0.23
	GG	5	2.12±0.75		1	1.14		4	2.36±0.59	
rs874898_G>G	G	692	2.04±0.44	0.55	208	2.00±0.40	0.72	439	2.08±0.45	0.55
	C	214	2.01±0.50		56	2.02±0.49		143	2.11±0.49	
	GG	264	2.04±0.42	0.56	82	2.00±0.38	0.92	166	2.07±0.43	0.43
	GC+GG	189	2.06±0.50		50	2.00±0.48		125	2.11±0.49	
	GG+GC	428	2.05±0.45	0.73	126	1.99±0.41	0.22	273	2.09±0.46	0.91
	CC	25	2.08±0.49		6	2.21±0.55		18	2.07±0.47	

Table 31: The influence of PAX8 variants rs2241975_C>T, rs11123172_T>C, rs13015478_G>T, rs67776659_T>C, rs13007173_A>G, rs3738913_G>A, rs3748915_T>C, rs3748916_A>G, rs1049137_A>G rs4849176_T>C and rs4849186_G>C on thyroxine dose requirement. N, number of patients in the group; Concentration is given as μg/kg body mass. TSH, thyroid-stimulating hormone; NSG, near-suppressed group; SG, suppressed group. *P<0.05

Furthermore, in the SG, the rs2241975_CT+TT (p=0.03) was significantly associated with elevated thyroxine dose and the rs3738913_GA+AA (p=0.05) was significantly

associated with lower thyroxine dose requirement. Thus, various SNPs in PAX8 gene displayed varying tendencies in their relationship with altered thyroxine dose requirement in the three patient groups (Table 31).

3.5.3.7. Sodium Iodide Symporter

In the case of Sodium Iodide Symporter (NIS), the same five SNPs subjected to disease association study, rs4808708_G>A, rs4808709_A>G, rs7250346_C>G, rs45602038_C>T and rs12327843_T>C were also investigated for their potential role in dose requirement. However, no definable relationship could be drawn between the SNPs and thyroxine dose requirement for all three groups (Table 32).

SNP ID	Geno-type	ALL(TSH<4.3)			NSG(0.1≤TSH<0.5)			SG(TSH<0.1)		
		N	T4	P	N	T4 dose	P	N	T4 dose	P
rs4808708_G>A	G	613	2.05±0.46	0.91	187	2.01±0.41	0.51	386	2.09±0.47	0.94
	A	289	2.05±0.45		77	1.97±0.45		192	2.09±0.44	
	GG	216	2.03±0.46	0.49	67	2.01±0.42	0.70	135	2.06±0.48	0.37
	GA+AA	235	2.06±0.45		65	1.99±0.42		154	2.11±0.45	
	GG+GA	397	2.06±0.46	0.39	120	2.01±0.40	0.43	251	2.10±0.47	0.24
	AA	54	2.00±0.43		12	1.91±0.59		38	2.00±0.39	
rs808709_A>G	A	612	2.04±0.46	0.65	183	2.01±0.41	0.58	389	2.08±0.47	0.63
	G	294	2.06±0.45		81	1.98±0.44		193	2.10±0.44	
	AA	216	2.03±0.46	0.48	65	2.02±0.42	0.67	137	2.06±0.47	0.34
	AG+GG	237	2.06±0.45		67	1.98±0.42		154	2.11±0.45	
	AA+AG	396	2.05±0.46	0.87	118	2.01±0.40	0.63	252	2.09±0.47	0.64
	GG	57	2.04±0.45		14	1.95±0.55		39	2.05±0.43	
rs7250346_C>G	C	595	2.05±0.45	0.95	174	2.00±0.40	0.80	380	2.09±0.46	0.86
	G	309	2.05±0.46		90	1.99±0.45		200	2.08±0.46	
	CC	202	2.04±0.45	0.80	57	2.01±0.40	0.87	130	2.09±0.46	0.84
	CG+GG	250	2.05±0.46		75	2.00±0.43		160	2.08±0.46	
	CC+CG	393	2.05±0.45	0.80	117	2.00±0.40	0.77	250	2.09±0.46	0.54
	GG	59	2.04±0.47		15	1.97±0.54		40	2.04±0.47	
rs45602038_C>T	C	852	2.04±0.45	0.08	253	1.99±0.41	0.22	543	2.08±0.46	0.23
	T	54	2.15±0.51		11	2.15±0.60		39	2.17±0.52	
	CC	419	2.04±0.45	0.22	124	2.00±0.41	0.63	267	2.08±0.46	0.30
	CT+TT	34	2.14±0.49		8	2.07±0.54		24	2.18±0.50	
	CC+CT	433	2.04±0.45	0.22	129	1.99±0.41	0.12	276	2.08±0.46	0.54
	TT	20	2.17±0.55		3	2.37±0.81		15	2.16±0.55	
rs12327843_T>C	T	616	2.04±0.46	0.37	182	2.00±0.40	0.80	392	2.08±0.47	0.59
	C	290	2.07±0.45		82	2.01±0.46		190	2.10±0.44	
	TT	218	2.03±0.45	0.53	64	2.03±0.40	0.49	139	2.06±0.47	0.32
	TC+CC	235	2.06±0.46		68	1.98±0.44		152	2.11±0.55	
	TT+TC	398	2.04±0.46	0.40	118	1.98±0.40	0.09	253	2.09±0.47	0.66
	CC	55	2.10±0.44		14	2.18±0.56		38	2.05±0.40	

Table 32: The influence of NIS variants rs4808708_G>A, rs4808709_A>G, rs7250346_C>G on thyroxine dose requirement. N, number of patients in the group; Concentration is given as µg/kg body mass. TSH, thyroid-stimulating hormone; NSG, near-suppressed group; SG, suppressed group. *P<0.05

3.5.4. Conclusion

Thus, altogether thirty-nine SNPs were evaluated for their role in DTC therapy and thyroxine dose requirement in DTC patients. The most significant findings are summarized in association Table 33. As demonstrated in this table, in the ALL group one DIO1 variant, rs12095080_GG (p=0.02) was associated, while three others showed borderline association. Two PAX8 variants rs2241975_(CT+TT) (p=0.01) and rs4849186_CC (p=0.01) were associated while another variant was only weakly linked to decreased dose requirement. The association for PAX8 variant, rs4849186_CC was retained in NSG (p=0.03) group and became weaker in the SG (p=0.06). Another PAX8 variant, rs1478_GG (p=0.04) was positively associated with elevated thyroxine dose in NSG. However, in the SG, two DIO1 variants, rs12084242_AA (p=0.01) and rs17109582_CC (p=0.03) turned significant while the rs12095080_GG (p=0.01) retained its significant association in the SG. Two PAX8 SNPs, rs2241975_(CT+TT) (p=0.03), and the rs3738913_GA+AA (p=0.05) were significantly associated with elevated dose in rs2241975_(CT+TT) and lower dose requirement in rs3738913_GA+AA. Furthermore, another TSHβ SNP, the rs1321109_A>G (p=0.059) showed a weak association with thyroxine dose requirement (Table 33).

SNP ID	Genotype	All (TSH<4.3)			NSG(0.1≤TSH<0.5)			SG(TSH<0.1)		
		N	T4 dose	P	N	T4 dose	P	N	T4 dose	P
DIO1 rs12084242 T>A	TA+TT	4439	2.04±0.45	0.07	1293	2.00±0.42	0.62	2846	2.08±0.45	0.01*
	AA		2.32±0.75			1.88±0.04			2.54±0.86	
DIO1 rs12095080 A>G	A	761	2.03±0.44	0.04*	222	1.99±0.42	0.77	484	2.07±0.44	0.02*
	G	131	2.12±0.50		38	1.97±0.34		88	2.19±0.55	
	AA+AG	436	2.04±0.44	0.02*	128	1.99±0.41	0.76	278	2.08±0.45	0.01*
	GG	10	2.37±0.72		2	1.91±0.01		8	2.48±0.77	
DIO1 rs17109582 T>C	T	767	2.04±0.45	0.06*	223	2.01±0.43	0.55	489	2.07±0.44	0.02*
	C	133	2.12±0.50		39	1.96±0.33		89	2.19±0.55	
	TT+TC	438	2.04±0.45	0.08	128	2.00±0.42	0.63	280	2.08±0.45	0.03*
	CC	12	2.28±0.68		3	1.88±0.04		9	2.41±0.75	
TSHβ rs1321109 A>G	A	622	2.04±0.45	0.32	176	2.01±0.44	0.68	405	2.06±0.45	0.059*
	G	282	2.07±0.46		88	1.99±0.38		175	2.14±0.48	
PAX8 rs2241975 C>T	C	712	2.03±0.44	0.03*	210	1.99±0.39		454	2.07±0.45	
	T	192	2.11±0.51		54	2.05±0.51	0.37	126	2.15±0.50	0.08
	CC	285	2.01±0.41	0.01*	85	1.98±0.36	0.41	181	2.04±0.42	0.03*
	CT+TT	167	2.12±0.52		47	2.04±0.51		109	2.16±0.51	
PAX8 rs3738913 G>A	GG	324	2.07±0.46	0.08	88	1.98±0.39	0.41	220	2.12±0.47	0.05*
	GA+AA	129	1.99±0.44		44	2.04±0.47		71	1.99±0.42	
PAX8 rs4849186 G>C	GC+GG	435	2.06±0.45	0.01*	128	2.01±0.42	0.03*	277	2.10±0.46	0.06*
	CC	17	1.78±0.44		4	1.55±0.17		13	1.85±0.48	
PAX8 rs1478 T>G	TT+TG	448	2.05±0.45	0.74	131	2.01±0.42	0.04*	287	2.08±0.46	0.23
	GG	5	2.12±0.75		1	1.14		4	2.36±0.59	

Table 33: Summary-association of positive variants with thyroxine dose requirement. N, number of patients in the group; Concentration is given as μg/kg body mass. NSG, near-suppressed group; SG, suppressed group. *P<0.05

3.6. Generalized Linear Models for estimating the best association model

3.6.1. Foreword

In the present study, several parameters being tested showed significant association by applying different statistical methods. However, it was thought to estimate the minimum adequate model for the most significant association of these variables with disease or dose requirement.

3.6.2. Methodology

The modeling approach adopted for thyroid cancer and L-thyroxine dose was generalized linear models with fixed effects implemented using R (statistical programming language www.R-project.org/) according to (McCullagh *et al.*, 1989). Binomial error distributions were assumed for the thyroid cancer model and logistic regression used to select the best fitting model. Gaussian error distributions were assumed for L-thyroxine dose models. Initial maximal models included all main effects and parameters were estimated by maximum likelihood. The significance of each term was assessed by comparing the deviance value on removal of the term, sequentially in order to arrive at a minimal adequate model.

3.6.3. Results

3.6.3.1. Generalized linear model for differentiated thyroid cancer association

Initially all 39 studied SNPs were entered together with age, gender and smoking to evaluate their effect on disease. The minimum adequate model arrived at was case/control ~ PAX8 rs11123172+PAX8 rs3748916+ TSHβ rs1321108 + NIS rs45602038 + TSHβ rs28566771 + gender + smoking. Thus, each of the above variables was significant (with PAX8 rs3748916 just on the edge). The R output is given in table 34.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			1042	1436.1	
gender	1	12.73	1041	1423.4	0.00036***
smoking	1	7.30	1040	1416.1	0.0069**
TSHβ rs1321108	2	7.41	1038	1408.7	0.025*
TSHβ rs28566771	2	8.06	1036	1400.6	0.018 *
NIS rs45602038	2	8.17	1034	1392.4	0.017 *
PAX8 rs11123172	2	6.99	1032	1385.5	0.03*
PAX8 rs3748916	2	5.07 .	1028	1371.1	0.079*

Table 34. R output of the Generalized Linear Model for the association of the variable with differentiated thyroid cancer. Df, degree of freedom; Resid. Df, residual degrees of freedom; Resid. Dev; residual deviance Pr; probability

By placing each term at the end of the model and calculating the proportion of the total deviance explained (1436.1), the following results were obtained.

	Df	Parameter estimate	Deviance	Pr(>Chi)
PAX8 rs11123172	2	-	0.0098	0.0008705 ***
PAX8 rs3748916	2	-	0.0090	0.0019985 **
TSHβ rs1321108	2	-	0.0075	0.0047183 **
NIS rs45602038	2	-	0.0067	0.0083142 **
TSHβ rs28566771	2	-	0.0044	0.0416637 *
gender	1	-0.3953	0.0038	0.0201209 *
Smoking	1	-0.5663	0.0040	0.0161908 *

Table 35. Summary of minimum adequate model of differentiated thyroid cancer with genotypes (binomial response, number of observations = 1043). Total deviance explained by main effects = 3.8%. For each term the deviance explained refers to the change in deviance attributed to the term in question when fitted last, as a proportion of the total deviance explained by the main effects in the minimum adequate model. P-values were estimated by comparison with the reduced model not containing the term in question. Signif. codes: ***P<0.001; **P<0.01; *P<0.05

As a result, total deviance explained by main effects = 3.8%, demonstrating that PAX8 rs11123172, PAX8 rs3748916, TSHβ rs1321108, NIS rs45602038, TSHβ rs28566771, gender and smoking have a small effect (3.8%) on the DTC risk

3.6.3.2. Generalized linear model for L-thyroxine dose association

For the dose requirement test, the minimum adequate model arrived at was T4 dose~PAX8 rs2241975 +PAX8 rs11123172+ PAX8rs4849186 +DIO1rs17109582. Thus, PAX8 rs2241975 was significantly associated (p=0.016), while PAX8 rs4849186 and DIO1 rs17109582 and PAX8 rs11123172 were just on the edge. The R output is given in table 36.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			419	86.886	
PAX8rs4849186	2	1.07196	411	82.553	0.069
PAX8rs11123172	2	1.16986	411	82.553	0.054
PAX8rs2241975	2	1.66199	411	82.553	0.016
DIO1rs17109582	2	1.08022	411	82.553	0.068

Table 36. R output of the Generalized Linear Model for the association of the variables with thyroxine dose. Df, degree of freedom; Resid. Df, residual degrees of freedom; Resid. Dev; residual deviance; Pr; probability

Summary of minimum adequate model of L-thyroxine dose with genotypes (binomial response, number of observations = 420). Total deviance explained by main effects = 6%. For each term the deviance explained refers to the change in deviance attributed to the term in question when fitted last, as a proportion of the total deviance explained by the main effects in the minimum adequate model. P-values were estimated by comparison with the reduced model not containing the term in question.

As a result, total deviance explained by main effects = 6%, shows that PAX8 rs2241975 +PAX8 rs11123172+ PAX8rs4849186 + DIO1rs17109582, have a small effect on the

L-thyroxine dose requirement. Notably, smoking, age and gender did not influence the model for the dose requirement.

3.6.4. Conclusion

The generalized linear model analysis showed that the minimum adequate model for DTC association included 2 PAX8, 2 TSHβ and 1 NIS SNPs in addition to gender and smoking have a small effect of 3.8 % on DTC, while 3 PAX8 and 1 DIO1 SNPs had an effect of 6.0% on thyroxine dose requirement.

4. General Discussion

4.1. Prevalence of gene polymorphisms in the Saudi population by sequencing

The present study focuses on two important questions with respect to the role of gene polymorphisms in the manifestation and management of differentiated thyroid cancer. The first describes the potential causative relation of six selected genes involved in the functional or metabolic pathways of thyroxine with DTC manifestation, while the second evaluates their possible involvement in patient response to therapy of the disease. Prior to tackling these issues, it was deemed essential to establish first which variants in the selected genes would be informative as a basis for selecting the SNPs for further studies in the Saudi population. Employing such population-based prevalence data, rather than randomly selecting variants based only on available literature would facilitate a more direct approach to the problem.

The six genes selected for the study included the three deiodinases (DIO1, 2 and 3), one anion transporter (NIS), one transcription factor (PAX8) and a thyroxine regulatory gene (TSHβ) involved in various synthetic and metabolic aspects as well as pathways of the thyroid hormone signalling. Sequencing of these genes in 96 healthy Saudi individuals was accomplished for all exons and their surrounding regions to ensure the identification of the maximum number of potentially functional variants. Following this, another population of 96 patients with DTC was also sequenced, to facilitate a preliminary association study for the discovered variants, on which selection of SNPs would be based for further evaluation of their role in disease and drug therapy in larger population sizes.

As indicated in the data described above, the sequencing results demonstrate the prevalence of varying degrees of polymorphic changes in the studied genes in the

studied population. The changes were ranging from less than five variants in DIO3 and TSHβ to about twenty in the PAX8 gene. This observation may be partly explained by the difference in the sizes of these genes. Overall however, the study captured approximately 69 SNPs in the six genes in the healthy study individuals. More importantly, 24 (35%) out of these variants are novel discoveries, pointing to possibly several of them being uniquely found in the Saudi population. Furthermore, the 3 DIO2 SNPs, rs191477357, rs187037837, and rs181554842 were originally discovered as novel in 2010 at the time of the experiment but were later on reported in the database. Among the novel SNPs were a number of non-synonymous variants, with a relatively significant minor allele presentation. Considering the small population size involved, it can be assumed that these variants are informative enough to warrant further investigation. Besides, several of the identified familiar coding SNPs were also non-synonymous, providing further support to the notion that this population harbours a large number of potential functional SNPs that need to be studied further with respect to their relevance for disease manifestation.

Perhaps the most notable finding in the study is the presence of indels in a number of these genes, some of which were present only in the patient groups. This finding is likely to be of high significance, since this type of mutations tends to lead to premature termination of protein synthesis and could therefore be lethal. A likely possibility is that some of these indels particularly those found only in disease cases might indeed be disease-causing in a (Mendelian) monogenic fashion. Hence, the notion of their impact on disease manifestation is attractive and demands further scrutiny.

It should be stated however, that a great number of SNPs were either intronic or resided in untranslated regions of the gene. On the other hand, it has also become increasingly evident that the UTRs, particularly the 3'-UTR, harbour important sequences for gene regulation and mRNA maturation processes. Hence, the presence of these SNPs in

these regions is in fact of particular interest with regard to disease pathways or mechanisms regulated by such entities. Furthermore, searching of the literature thus far has yielded negative results with regard to the prevalence of these variants in other ethnic Arab groups, except 2 variants in DIO2 gene (rs225011 and rs225012) that were studied in the Jordanian population (Al-Azzam *et al.*, 2013). However, it can be assumed that some of the novel findings may be unique for this or the Arab population in general. The observation that some SNPs were present in only one of the two groups is also likely to find explanation in the small sample sizes. Interestingly, some of the SNPs have so far been studied in other populations.

As pointed out in the Results section 19 of the SNPs described in the present study show significant deviation from the HWE principle. Criteria for fulfilling the HWE assume that there is no mutation, gene migration, selection or genetic drift in the population, whereby the latter two are usually a product of consanguinity in the population. In the Saudi population, consanguinity is an embedded cultural factor, which exists in the vast majority of families. Thus, the fact that a number of variants show such a drift from the Hardy-Weinberg's principle is likely to be partly explained by this cultural factor in our population. This scenario obviously allows preferential multiplication of any existing polymorphic traits which would influence the HWE in the population, particular when small numbers are employed to represent the general population. Therefore, the deviation shown in the present study partly indicates that the prevalence of certain SNPs was a result of preferential selection in this population. Hence, it is generally believed that very large study population sizes are normally required to provide accurate assessment of the HWE in the given general population.

The important question was whether this population might differ significantly from other ethnic groups. In order to address this question, it was deemed worthwhile to

compare the study data on the familial variants with that found in existing databases (Table 37).

SNP ID	Genotype	MA	Study	Europe	China	Japan	Africa	Database
DIO1								
rs2235544	C>A	A	0.27	0.53	0.51	0.36	0.12	HapMap
rs11206244	C>T	T	0.25	0.34	0.10 0	0.18	0.13	HapMap
rs114860598	G>A	A	0.04	NA	NA	NA	0.06	NCBI
rs113019354	C>A	A	0.02	NA	NA	NA	0.50	NCBI
rs78848743	C>T	T	0.02	NA	0.02	0.02	0.02	HapMap
rs12095080	A>G	G	0.13	0.098	0.00	0.00	0.23	HapMap
DIO2								
rs56033314	T>C	C	0.02	0.107	0.13	0.13	NA	NCBI
rs225011*	A>G	G	0.42	0.42	0.84	0.66	0.73	HapMap
rs225014	A>G	G	0.34	0.39	0.57	0.44	0.43	HapMap
rs7140952*	G>A	A	0.08	0.08	0.29	0.29	0.26	HapMap
rs225016*	A>T	T	0.46	0.57	NA	NA	0.03	NCBI
rs225017*	T>A	A	0.36	0.59	0.16	0.34	0.05	HapMap
rs75196191	A>T	T	0.01	NA	NA	NA	0.96	NCBI
rs17110436	G>A	A	0.03	0.00	0.00	0.00	0.09	HapMap
DIO3								
rs113370931	G>T	T	0.01	NA	NA	NA	0.50	NCBI
rs945006	T>G	G	0.10	0.12	0.02	0.00	0.52	HapMap
TSHβ								
rs72695872	C>T	T	0.39	0.46	0.43	0.43	0.24	NCBI
rs1321108	A>G	G	0.31	0.50	0.42	0.49	0.25	HapMap
rs10776792	G>A	A	0.01	0.03	0.01	0.01	0.01	HapMap
rs17477369	T>A	A	0.07	0.05	0.01	0.01	0.00	HapMap

SNP ID	Genotype	MA	Study	Europe	China	Japan	Africa	Database
PAX8								
rs1867763*	T>C	C	0.25	0.32	0.11	0.11	1.00	NCBI
rs13007173	T>C	C	0.11	0.11	0.00	0.00	0.01	HapMap
rs4849186	C>G	G	0.15	0.76	0.88	0.95	0.98	HapMap
rs77365796	A>T	T	0.05	0.92	NA	NA	NA	NCBI
rs1110839*	A>C	C	0.12	0.50	0.37	0.33	0.75	HapMap
rs67776659	A>G	G	0.05	0.50	NA	NA	0.50	NCBI
rs11123172	A>G	G	0.32	0.42	0.35	0.31	0.37	HapMap
rs3748916	A>G	G	0.5	0. 5	0.40	0.32	0.79	HapMap
rs3748915*	A>G	G	0.05	0.09	0.19	0.19	0.30	NCBI
rs4849176*	A>G	G	0.18	0.52	0.30	0.30	0.83	NCBI
rs1049137	T>C	C	0.12	0.03	0.06	0.06	0.50	NCBI
rs1479*	A>C	C	0.17	0.50	NA	NA	0.50	NCBI
rs874898	C>G	G	0.19	0.76	0.74	0.68	0.45	HapMap
NIS								
rs118133504	C>T	T	0.01	0.025	NA	NA	NA	NCBI
rs112077649	C>T	T	0.01	NA	NA	NA	0.06	NCBI
rs73518702	C>A	A	0.02	NA	NA	NA	1	NCBI
rs76616163	C>G	G	0.13	0.09	NA	NA	0.05	NCBI
rs7250061	C>T	T	0.03	0.02	0.23	0.17	0.19	HapMap
rs4808708	G>A	A	0.15	0.17	0.00	0.05	0.09	HapMap
rs4808709	A>G	G	0.26	0.18	0.00	0.05	0.24	HapMap
rs35036312	G>A	A	0.14	0.17	0.02	0.02	0.16	NCBI
rs12327843	T>C	C	0.26	0.17	0.00	0.05	0.22	HapMap
rs76103356	G>C	C	0.02	NA	NA	NA	0.18	NCBI

Table 37. Comparison of minor allele frequency distribution in the SNPs discovered in the six studied genes among various populations. The table compares minor allele (MA) frequency data from the current study with that in the HapMAp and NCBI database for DIO1,DIO2,DIO3,TSHβ, PAX8 and NIS genes. Study, current study; NA, not available. *Denotes that the SNP deviates from the Hardy-Weinberg equilibrium.

Based on available literature, it was interesting to establish in the current study that the majority of the minor alleles distribution results lie within the same average range of the values found in the HapMap database and other sources for various other ethnic populations. This was generally true especially with respect to those variants that deviated significantly from the Hardy-Weinberg's in the study population. These

include, for example, the frequencies of rs225011 (intronic) A>G (MAF = 0.42), rs225017 (3'-UTR) T>A (0.36) and rs225016 (3'-UTR) A>T (0.46) in the DIO2 to name a few, which were identical with some of those in the given databases. In fact, it appears that most of the study data matches most closely that of the Japanese. Therefore, these observations suggest that, in general, the current data is not dissimilar from that established for other ethnical groups, particularly the Asian populations. However, a few SNPs differed somehow from available data of various unverified sources in other populations, which might be explained by phenomena beyond the scope of this discussion.

Put together, the fact that the majority of the SNPs found in the present study meet Hardy-Weinberg's principle actually points to the likelihood that gene polymorphisms in our general population do not deviate significantly from the general expectation.

4.2. Evaluation of the role of gene variants in differentiated thyroid cancer

Since thyroid disorders present a heterogeneous group of defects arising mainly as a result of mutations in various genes engaged in the development of the thyroid gland, the focus of the present study was directed at investigating the potential role of some of these genes in DTC, to enable the determination of potential genetic variations that important in this regard in the study population. The genes of importance in this field include the transcription factors, such as the thyroid transcription factor-1 and -2 (TTF-1 and TTF-2) and paired box gene 8 (PAX8), as well as genes coding for the proteins involved in thyroid hormone biosynthesis, including thyroglobulin (TG), TPO, DUOX2 (Liang *et al.*, 2005; Smanik *et al.*, 1994), NIS (Tonacchera *et al.*, 2002a), TSH and TSHR (Rivolta *et al.*, 2005). Accordingly, six genes were selected from these families, partly because the literature was indicative of the possible relevance of polymorphic changes in these genes, and partly because they represent a spectrum of various

pathways that regulate the thyroid function. To achieve this, the role of the selected genes in DTC was investigated in two stages. The first set of experiments involved the association analysis by sequencing a small group of controls and DTC patients, but large enough to provide some statistically meaningful inferences. Accordingly, following the discovery of several SNPs in the six genes in healthy individuals, then a similar number of DTC patients were sequenced, in order to compare the prevalence of the discovered SNPs in the two groups as a first step. This stage served two purposes.

Firstly, the two groups put together would allow a better assessment of the SNPs in relatively larger sample size than the minimum normally required to make statistically meaningful inferences on the prevalence of the discovered SNPs. Secondly, this data set would also enable making initial assessment of the possible role of the discovered SNPs, which would lay the foundation for a larger population study. To begin with, interestingly, the two groups appeared to differ in the numbers of the identified SNPs. Specifically, a remarkably larger number of gene variants were discovered in the disease cases than in the healthy individuals. While it is tempting to suggest a link to the presence of the disease, these variations are also likely to be explained by the fact that some of the genes were sequenced in only a smaller portion of the healthy control group (i.e. the number of sequenced healthy individuals was not always the same as the cases), rather than by any particular genetic phenomenon. For this reason, the initial analysis of the data focused primarily on comparing those variants that were present in both groups.

The statistical analysis indicated that a number of the studied SNPs, including the rs1321108 in TSHβ and the rs11123172 in the PAX8, are associated with DTC in a potentially causative fashion. In both cases, it was primarily the minor allele that was most significantly associated, leading to the conclusion that simple polymorphic changes in these genes constitute the primary culprits. On the other hand, three other

SNPs, the rs1321109 in TSHβ, rs67776659 in the PAX8 and the rs945006 in the DIO3 genes were associated with DTC in a protective way. Some other SNPs showed borderline relationships with DTC. This laid the foundation for further interest to try and replicate these observations in larger population sizes. Besides, combining the data of the two groups revealed that several of these SNPs would still be prevalent with MAFs greater than 0.1, rendering their presence as a potential susceptibility factor for the disease. In addition, the discovery of indels in a number of genes in the population, some of which were novel and only present in the patients, can be conceived as potentially very important for DTC manifestation.

Interestingly, a number of SNPs discovered in the present study, including the rs11206244, rs12095080, rs2235544 in DIO1 as well as the rs225014, rs225012, and rs225010 in DIO2 have recently been discussed with respect to their possible role in regulating thyroid hormone levels and thyroxine dose requirement (Panicker *et al.*, 2008; Panicker *et al.*, 2009a; Torlontano *et al.*, 2008). Other variants in these genes have also been studied recently by various investigators with respect to their role in cancer and thyroid dyshormonogenesis. Accordingly, some substitution polymorphisms and missense mutations, such as the rs2235544 in DIO1(Panicker *et al.*, 2008), the rs225014 (Thr92Ala) in the DIO2 (Torlontano et al., 2008) and Q49X (rs121918668G>A) in TSHβ (Vuissoz *et al.*, 2001) have all been previously discussed as influencing T3 levels, causing thyroid cancer or congenital central hypothyroidism (Partsch *et al.*, 2006). Other studies have implicated the PAX8 gene polymorphism in various thyroid disorders, such as unusual small thyroid gland size, and in thyroid cancer in various ways (Komatsu *et al.*, 2001). Using this data as a basis for population-based association study, a larger case-control cohort then employed to analyse further the observed relationship by real-time PCR techniques. As shown in Figure 25, the study population was drawn from all regions of the country. This process

reduces the possible effect of biased selection and should be representative of the Saudi population in general. Besides, since the populations is homogeneous, the information obtained should be specifically relevant to Arab ethnic communities.

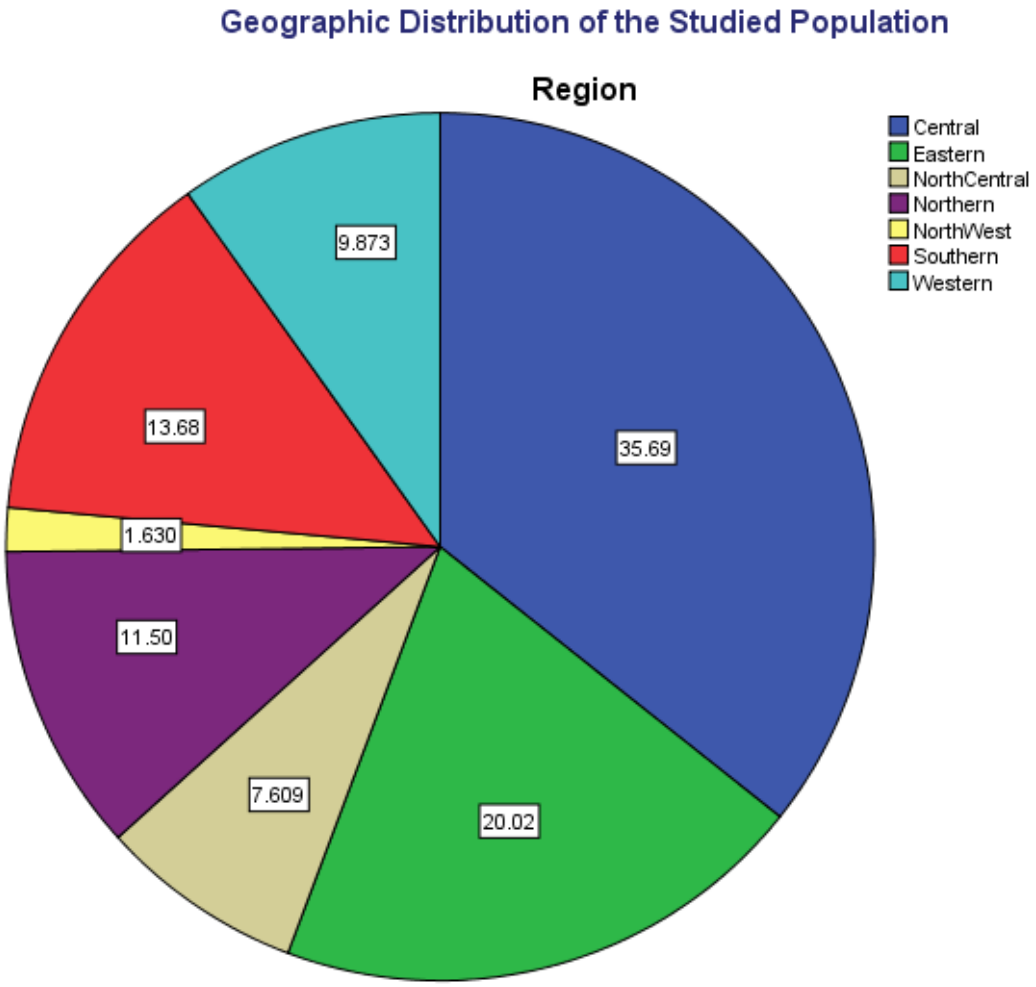


Figure 25: Pie chart showing the regional distribution of the 1104 studied individuals. These values also reciprocate the relative proportions of sizes the different regions of the country. Values are gives as percentage.

Hence, in addition to sequencing results, the study focussed on population-based analysis targeting particularly these variants in order to establish whether some of the observations of these studies could be replicated. A couple of variants that had been previously discussed in relationship with DTC in other ethnic groups were also, but not necessarily so in the study population. Their inclusion was primarily based on their prevalence in general population and partly on the availability of data in the literature on

their role in disease. Also, novel SNPs were omitted because they needed to be verified first and special probes needed to be designed in order to determine them by the same technique. Hence, only well-defined SNPs were employed.

4.2.1. Association of iodothyronine deiodinase variants with differentiated thyroid cancer risk

The sequencing experiments initially implicated several of the studied DIO1 SNPs in DTC. However, as mentioned above, this was partly because of their lack of expression in the healthy controls. These results did not stand the test of replication in the larger studies, and only a couple became weakly associated with the disease. Genetic alterations in DIOs have been a subject of a number studies with respect to their role in thyroid disorders. To date, polymorphisms described in DO1 include two SNPs, the C785T and A1814G (Peeters *et al.*, 2003b; van der Deure *et al.*, 2009), which appear to exert opposite effects with respect to their impact on disease, in general. In a number of studies, the Cys785Thr was found to affect the serum T3/rT3 ratio(de Jong *et al.*, 2007; Peeters *et al.*, 2003a; Peeters *et al.*, 2003c; van der Deure *et al.*, 2008a; van der Deure *et al.*, 2009), whereby the 785Thr variant resulted in decreased activity of DIO1 (Peeters *et al.*, 2006; Peeters *et al.*, 2003a). Specifically, some investigators suggested that the rs2235544 in the DIO1 gene alters deiodinase function, leading to an alteration in the balance of circulating free T3 to free T4 (Panicker *et al.*, 2008). Also in another study, two DIO1 SNPs, DIO1a-C/T and DIO1b-A/G, were implicated in changes in plasma T4, free T4, T3, rT3, and TSH levels (Peeters *et al.*, 2003b). Thereby, the DIO1a-T was associated with a higher plasma rT3, a higher plasma rT3/T4, and a lower T3/rT3 ratio in a dose-dependent fashion, while the DIO1b-G allele was linked to lower plasma rT3/T4 and higher T3/rT3 ratios (Peeters *et al.*, 2003b). Reduced DIO1 expression has also been reported in the non-thyroid syndrome and in several human neoplasmas (Arnaldi *et al.*, 2005; Maia *et al.*, 2011a). However, some investigators could not

establish similar associations for these SNPs with serum rT3 levels (Bianco *et al.*, 2002). It can therefore be inferred that alterations in the DIO1 gene may indeed lead to changes in serum thyroid hormone levels. Notably, the two SNPs are located in the 3'-UTR region of the gene. This fact seems to point to some gene regulatory function related to mRNA expression/maturation, rather than a result of changes in the DIO1 protein itself, as the cause.

In the DIO2 gene, one functional mutation, rs225014 A>G residing in exon 4, was found in the studied healthy individuals. However, no delineable relationship could be established for any of the other DIO2 variants with disease. On the other hand, in the literature DIO2 mutations have been associated with diverse types of various disorders, including insulin resistance, mental retardation and neurological disorders. Perhaps the most exhaustively-studied variant is the Thr92Ala (rs225014) that has been implicated in decreased DIO2 enzymatic activity and an increased risk of type 2 diabetes mellitus (T2DM), obesity and insulin resistance (Canani *et al.*, 2005a; Grarup *et al.*, 2007; Maia *et al.*, 2007). Data on the influence of the Thr92Ala polymorphism on serum T3 level is somewhat conflicting, but the SNP has been overwhelmingly associated with insulin resistance in different populations (Canani *et al.*, 2005b; Mentuccia *et al.*, 2005), with some other investigators failing to establish any association in various populations (Guo *et al.*, 2004). In the Amish population, an association of the Thr92Ala was observed with increased rather than decreased insulin sensitivity (Mentuccia *et al.*, 2005). Also, observations of reduced DIO2 in muscle and thyroid homogenates of the Ala92 allele carriers has led to the speculation that the ultimate effects are a product of linkage with other polymorphisms, possibly in form of a haplotype (Canani *et al.*, 2005a). Like in the case of DIO1, few of the implicated DIO2 mutation reside in the intronic regions of the gene. Some investigators have suggested that intronic mutations of the DIO2 gene may affect the amount of T3 available and partly determine the overall risk of mental

retardation in an iodine-deficient environment (Guo *et al.*, 2004). However, the role of the gene on neurological actions of the thyroid hormone remains unclear. Another interesting observation is the description of possible involvement of another variant in the 5'UTR of the gene (ORFa-Gly3Asp) in insulin resistance in various populations (Guo *et al.*, 2004). This SNP has been associated with the variations in the serum T3/T4 ratio in young compared to elderly subjects (Peeters *et al.*, 2005). Several studies have also focussed on the expression rather than mutational changes as cause of the disease. In one study, both DIO1 and DIO2 were found to be under-expressed in nearly all papillary thyroid carcinomas, leading to the notion that this could not only help to better define a tumour signature for thyroid tumours, but may also become useful as targets for thyroid tumour treatment (Arnaldi *et al.*, 2005).

Notably, in the DIO3 gene, one SNP, rs945006 T>G, was associated with DTC in the present study, and stood the test of the multivariate analysis in the larger study cohort. This SNP resides in the 3'UTR, and therefore, no speculation can be advanced with any possible relevance to protein functional changes. To date, although only limited information is currently available with respect to the functional role of these genomic regions, the untranslated regions are now known to be very active and dynamic entities, with respect to gene regulatory function, transcription factor activities and mRNA maturation processes. It is now well-acknowledged that these regions, particularly the untranslated regions contain several elements involved in gene regulatory mechanisms associated with mRNA maturation and various transcriptional activities. One important function of the 3'UTR is that of the micro RNA, which are engaged in the control of several factors related to gene regulation and protein maturation processes. In particular, it is now acknowledged that translational control mediated by the UTRs regulates cell proliferation, metabolism and responses to cellular stress, processes that are involved in carcinogenesis. Besides, several studies have shown association of

diseases such as the different forms of cancer with mutations in either the 5' or 3'-UTR of various genes. This leads to the conclusion that the changes in these regions remain nonetheless of potential interest and further studies are needed ascertain the role of the individual SNPs in cancer.

Although DIO3 is the major T3 and T4 inactivating enzyme, which contributes to TH homeostasis by protecting tissues from excess hormone, currently there is a little information on the possible role of DIO3 gene polymorphism in the management of thyroid disorders. A study involving 156 healthy blood donors previously described some SNPs located in the 3'-UTR region of the gene, albeit with no definable relationship with plasma T4, free T4, T3, rT3, and TSH levels (Peeters *et al.*, 2003b). Furthermore, it is now well acknowledged that the expression and activity of iodothyronine deiodinases are altered in different types of neoplasia (Piekielko-Witkowska *et al.*, 2011), but no inactivating mutations have been described in these genes to date. Another study has shown that increased DIO3 expression may lead to enhanced tumour proliferation, in support of the notion of the deiodinases having the potential to affect progression of cancer. This has in turn, led to the suggestion to exploit the potential of this information as cancer markers and modulators of tumour progression (Piekielko-Witkowska *et al.*, 2011).

4.2.2. Association of Thyroid Stimulating Hormone- Beta variants with differentiated thyroid cancer risk

By virtue of its stimulatory function on the thyroid gland, and ultimately the metabolism of almost every tissue, alterations in the thyroid stimulating hormone regulatory pathways are likely to produce direct overwhelming consequences on thyroid cancer manifestation. Understandably, great interest continues to evolve around efforts to establish the mode by which genetic modifications of components of this pathway might play a role in carcinogenesis and influence the treatment of the disease. However, in

general, information on the impact of the TSHβ gene polymorphism in thyroid cancer is still scanty. In the present study, as described in Results chapter, one of the four studied TSHβvariants (rs1321108) was strongly associated with DTC, while another (rs1321109) displayed a significantly protective property against acquiring the disease in the population-based study. Notably, both rs1321108 and rs1321109 are located in the promoter region of the gene (Appendix F), suggesting that the impact of this SNP may manifest itself through mechanisms associated with the function of this genomic region rather than changes in the protein itself. This is discussed further below. Hence, the results unequivocally point to changes in the sequence of this gene as being important in the manifestation of the disease.

While no reports are currently available in the literature pertaining to the impact of these two variants in thyroid cancer per se, other variants of this gene have also been studied with various outcomes. For example, a homozygous missense mutation Gly85Arg (rs190110651) in the TSHβ gene was described in a 15-month-old male of consanguineous parents with congenital hypothyroidism and in three Japanese cases with a similar phenotype (Muthukrishnan *et al.*, 2010). Furthermore, a Danish twin study also implicated the -G9A in harbouring elevated THRβ with higher serum TSH (Sorensen *et al.*, 2008). However, some other studies suggested that changes in the TSHβ gene do not present a relevant mechanism for human thyroid carcinogenesis (St Germain *et al.*, 1997). Apart from being implicated in thyroid cancer, mutations in TSHβ may cause a rare disease defined as congenital central hypothyroidism (CCH) (Partsch *et al.*, 2006). One such mutation described in the German population is the homozygous deletion, delta 313T, in which the TSH serum levels remained unaltered (Partsch *et al.*, 2006). Severe isolated TSH deficiency causing a protein change from glutamine (CAG) to a premature TAG stop codon has also been reported at position 49 of the TSHβ subunit gene in two children from a consanguineous Turkish family

showing homozygosity for a C>T (Q49X) transition (rs121918668) (Vuissoz *et al.*, 2001). Since serum TSH is thought to be a sensitive indicator of thyroid function, abnormalities in this function are believed to lead to common endocrine disorders affecting approximately 10% of individuals over a life span (Arnaud-Lopez *et al.*, 2008).

Perhaps one of the most well-studied genes in the TSH pathway is the TSHR, in which somatic gain-of-function mutations resulting in a phenotype of toxic adenoma, toxic multinodular goitre and congenital hyperthyroidism have also been described (Kohrle, 1999). However, only a few TSHR polymorphisms resulting in amino acid substitutions have been identified. Two of these, Asp36His and Pro52Thr, are located in the extracellular and another, Asp727Glu, in the intracellular domain of the receptor (Leonard *et al.*, 2001; Peeters *et al.*, 2003a). In one study, the Thr variant of the Pro52Thr was found to be significantly more potent in hyperthyrotropinaemic infants ($p=0.03$) compared to newborn controls (Visser, 1975). Mutations in the TSHR have also been described in a study involving children of consanguineous families with a non-syndromic autosomal recessively inherited non-goitrous congenital hypothyroidism phenotype (Cangul *et al.*, 2012; Cangul *et al.*, 2010), while other loss-of-function mutations have been associated with TSH resistance and congenital hypothyroidism (Barbaro *et al.*, 2003; Cassio *et al.*, 2013). It has been suggested that individuals harbouring a heterozygous loss-of-function mutation have a dominant transmission of partial TSH resistance, possibly due to intracellular entrapment and reduced maturation of the wild-type TSHR by inactive mutants (Calebiro *et al.*, 2005; Cassio *et al.*, 2013). The Leu512Arg found in the autonomously functioning papillary carcinoma has been associated with thyrotoxicosis and a hot thyroid (Barbaro *et al.*, 2006). Besides, a number of microsatellite markers and intronic SNPs residing in the TSHR gene, such as Asp727Glu, have also been linked to lower plasma TSH levels in healthy blood donors,

without influencing free thyroxine (FT4) levels (Bianco *et al.*, 2005; Peeters *et al.*, 2003b). However, their contribution to the genetic variance is thought to be small. One study has suggested that the variant triggers an increased cAMP response of the receptor to TSH (Leonard *et al.*, 2001). Another study pointed to an association of TSHR mutations with iodine supply (Curcio-Morelli *et al.*, 2003), while in non-diabetic elderly men, variation in serum thyroid levels and the Asp727Glu have been associated with relative insulin resistance (Callebaut *et al.*, 2003). On the other hand, these findings could not be replicated by some studies, suggesting therefore that this variant may be linked to other polymorphisms in a haplotype form (Braverman, 1994). Also polymorphisms leading to a relative loss of function may exert an effect through this mechanism, and thus account for some of the so-called 'euthyroid outliers' with elevated TSH determinations (Hesch *et al.*, 1975).

A case has been recently described in which a child and her mother presenting with different clinical phenotypes, but were both affected by Resistance to thyroid hormone (RTH) and carried an identical missense variation in exon 8 of THRB gene in the heterozygous state I276N (Monzani *et al.*, 2012). A similar scenario was also reported in a family in which the patient had hyperthyroidism, and the father was positive for both anti-Tg antibody and anti-TPO antibody, yet both carried the Pro453Thr (Sato *et al.*, 2006) and showed elevated levels of TSH, FT3 and FT4. The I276N variant was also reported in a case of postpartum thyroiditis (PPT) (Paragliola *et al.*, 2013), a syndrome characterized by the development of postpartum thyroid dysfunction, which may occur up to 12 months after delivery and usually presents with transient thyrotoxicosis, followed by transient hypothyroidism. Gene expression studies have indicated that loss of the thyroid hormone receptor is common in tumours. In mouse models, a truncated THRβ gene leads to thyroid cancer (Jazdzewski *et al.*, 2011) .

4.2.3. Association of paired box 8 variants with differentiated thyroid cancer risk

Like the TSHβ gene, at least two PAX8 variants were identified as predisposing individuals to thyroid cancer in the preliminary investigation. However, these variants lost their significant association in the larger population study. Instead, a different variant rs11123172_T>C was implicated, while another was only weakly related in the same fashion. Put together, these results point to the gene as posing a risk for the manifestation of the disease. Furthermore, four PAX8 variants were associated with DTC, whereby one intronic variant rs11123172 T>C was potentially causative and the other three rs6776659 (intronic), rs1479 (3'-UTR) rs1478 (3'-UTR) were protective, against differentiated thyroid cancer. The clarity over the borderline relationship might be resolved in larger population sizes. The question as to the relevance of the intronic SNPs being important for the disease follows the same argument as for the variants residing the promoter and 3'UTR of the gene, whereby it is now well accepted that these regions may play regulatory roles which may ultimately be important in disease. Moreover, there could also be unidentified genes in these regions to which these positions may be ascribable. These points are discussed further below.

Currently, there is also lack of data on the role of the PAX8 gene in cancer, in general. However, some mutations have been described to varying extent. Thereby, some mutations are thought to cause congenital hypothyroidism (Congdon *et al.*, 2001), while others are thought to moderately reduce thyroid hormone levels, possibly as a result of the unusual small thyroid gland size in people with the PAX8 mutation (Komatsu *et al.*, 2001). For example, a girl suffering from thyroid dysgenesis revealed a maternally inherited heterozygous mutation (-456C>T) in the promoter region (Villanueva *et al.*, 2000). Particularly interesting is the report that the PAX8 gene, which resides on chromosome 2, fuses with the peroxisome proliferator-activated receptor (*PPAR*) gene on chromosome 3, leading to a translocation called PAX8/PPAR gamma (PPAR-γ)

rearrangement. FTC is thought to be characterized by PAX8/PPAR-γ rearrangements and RAt sarcoma (RAS) mutations, whereas B-Raf (BRAF) mutations, RET/PTC and TRK oncogenes have been detected in PTC (Diallo-Krou *et al.*, 2009; Greco *et al.*, 2009b). Expression of such a PAX8- PPAR-γ fusion protein, PPFP has been reported in about 35% of FTCs and a small fraction of follicular adenomas (Kroll *et al.*, 2000). The most common molecular alterations associated with this phenomenon in thyroid cancer include the BRAF and RAS point mutations and RET/PTC and PAX8/PPAR-γ rearrangements in more than 70% of papillary and follicular thyroid carcinomas (Nikiforov, 2011). In one study, the frequency of PAX8-PPAR-γ rearrangement was found to be similar in the follicular variant of papillary thyroid carcinoma (FVPTCs), follicular thyroid carcinomas (FTCs) (45.5%), and follicular thyroid adenomas (FTAs) (33.3%) (Castro *et al.*, 2006). A PAX8-PPAR-γ rearrangement has also been reported patients with encapsulated FVPTC (Rivera *et al.*, 2010). In FVPTCs, the PAX8-PPAR-γ rearrangement was significantly associated with multifocality and vascular invasion (Castro *et al.*, 2006). The mechanism by which PPFP contributes to follicular thyroid neoplasia is not fully understood. On the other hand, the biological pathways controlled by this transcription factor remain largely unknown, due to lack of identification of its additional targets, despite the fact that it is well-characterized with respect to its role in regulating genes involved in thyroid differentiation. Nonetheless, the transcription factor has recently been investigated as a potential marker of thyroid epithelial neoplasms (Schmitt *et al.*, 2010), these findings have been shown to be a potential diagnostic tool, particularly in thyroid fine-needle aspiration samples with indeterminate cytology (Nikiforov, 2011). The use of these and other molecular markers is believed to improve the accuracy of cancer diagnosis, thereby permitting more individualized management of thyroid cancer patients (Nikiforov, 2011).

4.2.4. Association of anion transporter variants with differentiated thyroid cancer

The sixth gene investigated in this study was the anion transporter, NIS. One of the studied NIS variants, rs45602038 (p.cys542cys) residing in exon 13, was associated with DTC in population-based study in a causative fashion. Currently, no study is available on the SNPs investigated in the present study. Furthermore, only a few studies are available in the literature in general pertaining to the role of thyroid transporters in thyroid disease as a whole. This does not, however, preclude an important role of this family of genes in carcinogenesis, as some studies have indicated that changes in the anion transporters may have a noticeable impact on the manifestation of different forms of cancer. However, the majority of studies have focused on the changes in expression levels, rather than genetic changes, in the NIS gene in thyroid cancer (Filetti *et al.*, 1999; Joba *et al.*, 1999; Lin *et al.*, 2001; Neumann *et al.*, 2004; Scipioni *et al.*, 2007; Sodre *et al.*, 2008; Tonacchera *et al.*, 2002a; Tonacchera *et al.*, 2004; Trouttet-Masson *et al.*, 2004). Thereby, low radioiodide uptake due to a defect in iodine transport observed in the majority of nonfunctioning thyroid nodules has been attributed to the defect in the expression and therefore targeting of human NIS (hNIS) protein to the cell membrane (Tonacchera *et al.*, 2002b). One study showed lack of NIS mRNA expression in some papillary thyroid cancers, but distinct presence in follicular thyroid adenoma (Arturi *et al.*, 1998), leading to the conclusion that early detection of such loss of the NIS gene expression in the primary cancer may provide useful therapeutic information in DTC patient (Arturi *et al.*, 1998). Another study showed a significant diminishing in NIS expression in thyroid carcinomas and cold adenomas and over expressed in toxic adenomas and Graves thyroid tissues (Lazar *et al.*, 1999), in inverse relationship to the tumour stage (Lazar *et al.*, 1999). This led to the conclusion that reduction in NIS gene expression occurs in most hypofunctional, benign and malignant

thyroid tumour, possibly activated by DNA methylation in critical regulatory regions (Venkataraman *et al.*, 1999b).

Furthermore, a silent polymorphism, C544G, in hNIS was described in one nonfunctioning nodule among patients that were originally subjected to surgery for a solitary nonfunctioning thyroid nodule originating in an otherwise normal gland after the oral administration of a tracer dose of iodine-131 (^{131}I) (Tonacchera *et al.*, 2002b). Recent studies reported that the most prevalent genetic event found in papillary thyroid carcinoma, the oncogene BRAF (V600E) mutation, correlates with a less differentiated tumour stage due to reduced expression of key genes involved in iodine metabolism. The BRAF has also been suggested as the most frequent genetic event and prognostic factor in PTC, correlating with a high risk of recurrences and less differentiated tumours due to the loss of NIS-mediated ^{131}I uptake (Riesco-Eizaguirre *et al.*, 2006). Hence some studies have suggested that BRAF mutational status and decreased NIS and TSHR expression in such incidences may reduce radioiodine uptake and lead to a negative response to radioiodine therapy (Bikker *et al.*, 1996). A Gly543Glu substitution in the NIS gene has also been described in Japanese siblings with I^2-transport defect, an observation which was supported by the finding of minimal I^2-uptake in NIS constructs carrying this mutation in COS-7 cells (Kosugi *et al.*, 1998). Another study observed that, the homozygous mutant (C272X, rs121909175) leading to the stop codon at the nt5868) causes congenital hypothyroidism (Pohlenz *et al.*, 1997). A G1530A transition of the NIS gene, resulting in a Gly395Arg substitution, has also been described in a large Hutterite family with an autosomal recessive form of congenital hypothyroidism due to an iodide transport defect, pointing to this mutation as the direct cause of the iodide transport defect in these patients (Couch *et al.*, 1985; Kosugi *et al.*, 1999). Thus, although changes in the NIS do not appear to have a great impact on DTC, they appear to be nonetheless important in various forms of cancer disorders. On the other hand,

some studies failed to establish major alterations in the hNIS gene in nonfunctioning nodules, but described a silent polymorphism (C544G, exon 13) found in one nodule (Tonacchera *et al.*, 2002a). The authors concluded that two mechanisms may contribute to the reduced radioiodide uptake typical of benign nonfunctioning thyroid nodules involving reduced expression of the hNIS protein, as well as defective targeting of hNIS to the cell membrane (Tonacchera *et al.*, 2002a).

4.3. Association of haplotypes with differentiated thyroid cancer

Apart from changes at single loci, one important possibility of how the gene variants may be involved in disease is by being part of a disease causing region as a haplotype in the genomic/chromosomal region, rather than the individual variants or genes per se. As describe above, a haplotype constitutes a combination of alleles that are inherited together at adjacent loci on a chromosome. Depending on the number of recombination events that have evolved between a given set of loci, this may represent a single locus, several loci or entire chromosome. The other definition of haplotype is a set of SNPs on one of a pair of chromosomes that is statistically associated with one other sequence of nucleotides. Such associations of nucleotides as well as the identification of a few alleles of a haplotype sequence can uniquely characterize all other polymorphic sites in its region. Such information has proven to be valuable for investigating the genetics of common diseases, such as cancer.

Having established that several of the studied SNPs in the different genes were linked to DTC, the next obvious step was to evaluate the possible role of haplotypes constructed from these variants in disease manifestation. Only five of the six studied genes were analysed. Exclusion of the DIO3 was due to the fact that only one SNP of this gene had been included in the SNP association study. However, the analysis for the two deiodinases did not reveal much either in the direction of haplotyping associations. In particular, it was interesting to establish that although several haplotypes could be created at different levels in the DIO1, none of them revealed any definable relationship with disease in any direction. A similar pattern was observed for DIO2, except that one 4-mer haplotype conferred risk for the disease, in the absence of any link between individual SNPs and the disease. It appears therefore that the study revealed isolated potentially important risk variants and haplotypes for DTC in this family of protein. The

role of DIO haplotypes in disease has received some attention previously. Thus, observations of reduced DIO2 in muscle and thyroid homogenates of the Ala92 allele carriers has led to the speculation that these were an ultimate result of linkage of this change with other polymorphisms, possibly in form of a haplotype (Canani *et al.*, 2005a). However, like in the case of DIO1, several of the DIO2 mutations implicated in the present or other studies reside primarily in the intronic regions of the gene, which may offer some explanation as why only limited delineable relationships could be established with the individual variants or haplotypes. On the other hand, some investigators have suggested that intronic mutations of the DIO2 gene may affect the amount of T3 available and partly determine the overall risk of mental retardation in an iodine-deficient environment (Guo *et al.*, 2004). In the case of the DIO2, in our study, the haplotype associated with disease was constructed with a combination of coding and non-coding SNPs, with no individual SNPs showing any such trends. Therefore, it is likely that there are intronic components that may be essential for the function of these enzymes that were not picked in our study, possibly due to the small sample size involved in the study. Thus, the selection of 39 SNPs among the six studied genes involved in the case-control study was based on the results obtained from the sequencing of 48-96 Saudi individuals. However, although based on the power analysis calculations, 48 individuals were enough to detect informative SNPs in the studied population; it is still possible that some important SNPs might have been missed due to small sample size. As was evident from the haplotype analysis, in many cases, the regions containing some of the 39 SNPs were associated with the disease, while the individual SNPs themselves were not. For example, in the NIS gene, the 5-mer CGAGT was positively associated with DTC whereby the analysis indicated that the SNP block 2-4 (rs4808708, rs4808709 and rs750346) was the core of that relationship, although none of the three constituent SNPs was individually linked to the disease.

Unlike the DIOs, the TSHβ was unequivocally associated with disease at haplotype level. A similar scenario was observed for the NIS gene whereby the two haplotypes showing opposite trends could be distinguished by two SNPs. Importantly, the associations were similarly greater than the various associations of the individual variants. The literature is still lacking with respect to the possible influence of haplotypes in these gene on DTC. On the other hand, the present findings indeed point to these genomic regions in these two genes as potentially important with respect to DTC manifestation. Interestingly, the only notable 10-mer haplotype in the PAX8 analysis exhibited protective properties. This haplotype originated from the later part of the gene, comprising primarily SNPs in the last exon of the gene. Variants in the first portion of the gene displayed only a weak link with the disease. Notably also the minor allele of three SNPs in this region rs67776659, rs1479 and rs1478 were all protective against the disease, and none of the studied variants was linked in a causative fashion. Thus, based on the present observation, it appears that changes in this gene in this study population are concentrated in the latter part of the gene.

Put together, a number of observations can be inferred from the results of the present study. First, at least three genes were associated with disease at the haplotype level. As stated above, these associations were in general more significant than those of the individual variants. The important question is therefore whether haplotypes might be more informative than the individual variants. Based on the present results there appears to be overwhelming evidence that this may be the case. Besides, these haplotypes also involved variants that were otherwise not identified as individual risk loci, suggesting therefore that stretches of nucleotides rather than individual nucleotide changes may indeed be the responsible entities for the association of the gene polymorphisms with disease. Indeed, there is some evidence from the currently

available data suggesting that the characterization of these patterns of genetic variations, in the form of haplotypes, rather than individual variations per se, may indeed be more informative regarding the influence of genetic changes on disease pathways. Hence, several haplotypes have already been described to date making a convincing case for searching for such genomic sequences rather than single point mutation as primary causes for diseases. This is particularly important for complex disease, which are for most part a result of interactions of various factors rather than a change at a single locus. These stretches of nucleotides may reside within one gene, non-coding areas of a gene or simply a genomic region encompassing more than one gene.

The second important observation was that the constituent nucleotides of the haplotypes were both coding and non-coding SNPs. In the present study, the fact that most of the variants involved were non-coding may therefore imply also that yet undiscovered genes encompassing these loci are actually responsible for the observed actions. In fact, the majority of these variants reside mostly in the 3'UTR of the genes. Currently, only limited information is available with respect to the functional role of specifically the 3'UTR of most of gene. Nonetheless, it is now well-acknowledged that these regions, particularly the untranslated regions contain several elements involved in gene regulatory mechanisms associated with mRNA maturation and various transcriptional activities related translational control, cell proliferation, metabolism and responses to cellular stress, processes that are involved in carcinogenesis. Hence, the manifestation of non-coding regions as predisposing to disease has a number of potential implications. In particular, it can be speculated that these genomic regions bear specific relevance for various complex disease pathways and mechanisms, as they are now known to have a number of functions yet to be fully appreciated. These mechanisms involve certain gene regulatory entities, genetic hitchhiking and regional stretches of nucleotide in form of haplotypes rather that single variants.

To date, several studies have evaluated the possible association of diseases such as different types of cancer with mutation in either the 5'- or 3'-UTR. For example, a number of genes, such as the DIO2, are known to undergo complex 5'-UTR transcriptional mechanisms, with several transcriptional start sites and therefore alternative splice position, which would alter the open reading frame, rendering its expression to be tightly controlled and tissue specific (Bartha *et al.*, 2000). Other studies showed that the mechanism by which the transcription of certain genes such the thyroid hormone response element is regulated in the 3'UTR is different from that through the regulator element located in the promoter region (Bigler *et al.*, 1995). Pointing to multiple yet undefined modes by the way genes may be regulated in these regions.

Recently the literature has also demonstrated inhibition of certain genes, such as the expression of THRβ in PTC through the upregulation of mRNA, for example, (Jazdzewski *et al.*, 2011). It should be noted that thus far, the knowledge about the relevance of the non-coding genome to human disease has been acquired primarily through studies involving the widespread disruption of microRNA (miRNA) expression and function that is observed in human cancer. MicroRNAs (miRNAs) are small noncoding RNA genes composed of 21–25 nucleotides, whose targets are located in 3'UTR of the genes. They inhibit expression of genes at the transcriptional and posttranslational levels (Zhou *et al.*, 2012), and play an important role in proliferation, differentiation and apoptosis (Nikiforova *et al.*, 2008; (Nikiforova *et al.*, 2009) Pallante *et al.*, 2006). It is thought that changes in the expression of certain miRNAs can induce initiation and/or progression of human cancers (Greco *et al.*, 2009a). Several studies have demonstrated changes in miRNA profiles in thyroid cancers, dependent on cellular origin and tumour differentiation (Nikiforova *et al.*, 2008a; Pallante *et al.*, 2006; Schwertheim *et al.*, 2009). In particular, miR-222, miR-221, miR-146b and miR-187,

are thought to play pivotal roles in PTC initiation and development (He *et al.*, 2005). It was shown the miRNA-221 was also up regulated in unaffected thyroid tissue of PTC patients, leading to the conclusion that this is an early event in carcinogenesis (Chou *et al.*, 2010; He *et al.*, 2005). Moreover, the association between increased expression of miR-222, miR-221, miR-146b and miR-135b have been shown to be strongly related with pathological features of PTC in samples of PTC patients with extrathyroidal invasion (Wang *et al.*, 2013b). The candidate targets of these miRNAs could be p27kip1, a cell cycle progression inhibitor, and KIT, a receptor tyrosine kinase that plays important roles in cell growth and differentiation (Pallante *et al.*, 2006; Wang *et al.*, 1998). Therefore, alterations in miRNAs may contribute to diseases such as thyroid cancer. A recent study also demonstrated that miR-146b could trigger thyroid cancer development by regulating TGF-β signalling, although the mechanisms are largely unknown (Geraldo *et al.*, 2012). In the present study, one DIO3 SNP, rs945006G>T and two PAX8 SNPs (rs1479 T>G, rs1478 T>G) were all negatively associated with DTC. Hypothetically, since these SNPs are located in the 3'UTR, their presence might lead to changes in the miRNA targets in this region, leading to changes in their cell proliferation and differentiation activities, and therefore possibly leading to cancer. Importantly, apart from BRAF and RAS point mutations as well as RET/PTC and PAX8/PPARγ rearrangements, miRNAs are also thought to be among the most promising markers for diagnostic purposes in many cancers including lung, pancreas, breast, colorectal and thyroid cancers (Gomez Saez, 2011; Lu *et al.*, 2005; Nikiforov *et al.*, 2009; Nikiforova *et al.*, 2008b; Pallante *et al.*, 2006; Volinia *et al.*, 2006).

In the present study, two PAX8 SNPs (rs11123172 and rs67776659), that were associated with disease reside in intron 10 of the gene. As discussed above, these regions contain functional entities involved in the regulation of gene expression. For example, the mutations may be residing in a region influencing spliceosome function

and regulation of the mRNA synthesis. This would ultimately lead to a non-functional variant of the protein. Besides, they may constitute part of a genomic region encompassing a coding area of the gene that may be involved in pathways leading to disease manifestation, just by virtue of simple hitchhiking. Furthermore, two of the studied SNPs (rs1321108 and rs1321109), also associated with disease in different ways, reside in the promoter region of the TSHβ gene, whereby the rs1321108 is found in close vicinity to the start codon (about 50 bp) for the gene. It can be speculated that alterations in the region may influence the function of the transcriptional machinery possible leading to abnormal splicing of the gene. However, more specific work needs to be done to support these notions. This may involve testing the function of the related regulatory entities, such as the transcription factors, in presence of these gene variants, or the impact of their cumulative effect on the extent of disease.

Another study in clear renal cell cancer recently suggested that the TSHβ expression in these cells was maintained by posttranscriptional mechanisms involving 5'UTRs and miRNA-204, leading to the conclusion that its reduced expression and tissue hypothyroidism in the tumours is likely to be involved in the process of carcinogenesis or in maintaining a proliferative advantage to malignant cells (Master *et al.*, 2010). However, we are only beginning to understand the nature and extent of the involvement of these non-coding RNAs (ncRNAs) in disease. Apart from the miRNA, other ncRNAs, such as PIWI-interacting RNAs (piRNAs), small nucleolar RNAs (snoRNAs), transcribed ultraconserved regions (T-UCRs) and large intergenic non-coding RNAs (lincRNAs) are also emerging as important entities involved in cellular homeostasis that are likely to play a great role in disease (De Benedictis *et al.*, 1996).

Recently it has also been suggested that non-coding SNPs may show the association with a disease by virtue of them residing in the vicinity of an important variant through genetic drifting during the course of time, a phenomenon known as genetic hitchhiking

(also called genetic draft). Through this process the frequency of an allele may increase due to the fact that it is linked to a positively selected gene. Hence it is now believed that deleterious mutations can inhibit the spread of linked adaptive mutations through a population or adaptive substitutions thereby increasing the frequency of linked deleterious mutations present at particular locus. As a result, the distribution of disease-associated alleles will also be altered in the regions harbouring hitchhiking. Such alleles have been associated with several disorders including, cancers metabolic disorders, autoimmune and mental disorders. Accordingly, it has been suggested that positive selection has had a significant impact on deleterious polymorphism and may be partly responsible for the high frequency of certain human disease alleles (Li *et al.*, 2008b). Deleterious mutations are thought to reduce genetic fitness within natural populations. Therefore they need to be continually removed by natural selection. On the other hand, it is thought that some of them may reach unexpectedly high frequencies. Several mechanisms have been ascribed to this phenomenon. These include changes in genetic or environmental constraints. For example, it has been suggested that some deleterious mutations have hitchhiked to high frequency due to linkage to sites that have been under positive selection (Li *et al.*, 2008b). This concept might also explain the association of particularly the intragenic non-coding SNPs, as a result of the variants being linked with causative mutation on these genes through positive selection at this locus (Barton, 2000; Mäkinen *et al.*, 2008; Yeaman, 2013) or some spatial selection pressures (Barton *et al.*, 2013; Desai *et al.*, 2013; Flaxman *et al.*, 2012; Kim, 2013; Schneider *et al.*, 2013). However, further speculation on this possibility is beyond the scope of the present work. Hence further studies are needed to look into these possibilities. However, such changes in allele frequency could be due to any form of selection regulating organism. Thus, when an organism inherits some new and beneficial gene, it inherits everything in the vicinity of the gene also. Hence, if that new gene becomes fixed in the population,

then everything around it will also become fixed. The consequence is that when a strong selection process leads a new gene to being fixed at a locus in a relatively short time, there remains a mark on the genome, resulting in the locus displaying distinctively reduced genetic diversity than other genomic regions. This action defined as selective sweep in turn results in fixation of many factors, most of which will be present simply by virtue of being in the neighbourhood of the beneficiary gene. This may provide an explanation for the differences in genomic characteristics among the different ethnic population.

Theoretically, haplotypes may in turn also be a result of genetic hitchhiking. Hence, the impact of such genomic changes on disease may vary among ethnic group. This makes it mandatory to characterize these changes for individual ethnic groups in order to realize the clinical impact for any given populations.

Put together, the implication of non-coding SNPs in DTC in present observations appears to point to the importance of entities for manifestation of the disease. While their exact roles in disease need to be deciphered more accurately, it nonetheless can be speculated that the untranslated or non-coding regions of a gene are equally important in causing disease. There are many areas to be exploited with regard to the relevance of structural changes occurring in these genomic regions. Understanding of these processes will lead to development of therapeutic strategies to counteract these perturbations of variants.

4.4. Summary: Gene polymorphisms and differentiated thyroid cancer risk

In summary, the present study identified several coding and non-coding SNPs in a number of genes related to thyroid function that were implicated in differentiated thyroid cancer disease. Most importantly, the study described several haplotypes particularly in the TSHβ NIS and PAX8 genes that were associated with DTC. Notably, the vast majority of the SNPs discovered in the studied population were non-coding with respect to the selected genes. These variants were partly in the 3 prime untranslated, intronic or promoter regions of the genes, and were equally implicated in disease. It can therefore be concluded that the observed relationships of these variants with disease are not necessarily a result of functional changes in the respective proteins, but primarily linked to the functional activities of the genomic regions of these suspected gene. As discussed briefly in preceding paragraphs, it is becoming increasingly evident that the non-coding regions of many genes, including the introns, may equally contribute to disease like those residing in the coding regions. There are a number of possibilities that may explain how these mutations may exert their impact. These may be related to mechanisms involved in certain gene regulatory pathways, involving entities that are associated with the different functions of the untranslated regions of the gene, rather than changes at level of individual single nucleotide changes.

4.5. Association of gene polymorphism with L-thyroxine dose requirement

Following thyroidectomy, thyroid hormone quickly declines and TSH increases in a fashion similar to that of spontaneous hypothyroidism state. Hence, levothyroxine becomes part of the therapeutic strategy. In contrast to simple hypothyroidism, the goal of treatment in DTC is not only to replace thyroxine, but also to give a thyroxine dose high enough to suppress the growth promoting effect of TSH on any residual tumour cells. The treatment of choice for correction of hypothyroidism is synthetic L-thyroxine (L-T4), a prohormone with very little intrinsic activity. The L-thyroxine is normally deiodinated in peripheral tissues to form T3, the more active thyroid hormone. This deiodination process accounts for about 80% of the total daily production of T3 in normal subjects. Appropriate treatment reverses all the clinical manifestations of hypothyroidism. However, the therapeutic outcome may differ depending on the type of thyroidectomy performed. It is also appreciated that considerable interindividual variations exist in L-thyroxine dose required for TSH suppression in patients with DTC. Furthermore, the regulation of thyroxine metabolism is subject to many interlinked signalling pathways.

In the present study, the investigation of the variants engaged in the disease association study was extended to further investigate their possible role in the L-thyroxine dose requirement in the DTC athyrotic patients receiving thyroxine. For simplicity, the data for the individual genes/gene families are discussed separately.

4.5.1. Association of iododeiodinases with dose response requirement to in DTC therapy

The local availability of various iodothyronine metabolites is mediated through the thyroid hormone metabolism by the three deiodinase selenoproteins, DIO1, DIO2, and DIO3 (Kohrle, 2007), thereby mediating their effects on many important reactions, such as gene expression, thermoregulation, energy metabolism during the development

and maintenance of an adult organism (Kohrle, 2007). The circulating serum levels of the thyroid hormones and TSH are used as a combined indicator of thyroid hormone status. These include thyroid secretion as well as tissue-specific production of T3 by DIO1 and DIO2 activity, both of which contribute to T3 circulating levels and degradation of the T4, T3, rT3 and other iodothyronine metabolites with a lower iodine content as well as thyroid hormone conjugates. These degradation reactions are catalysed by either DIO1 or DIO3. Hence, mutations in DIOs and the gene regulators are likely to show more complex picture of the way in the adaptation of systemic and local bioavailability of thyroid hormones (Kohrle, 2007).

Of all the studied genes, the DIO1 showed the most convincing data, implicating at least three of the studied SNPs in the different ways in which patients respond to therapy with thyroxine. Notably, all three SNPs showed differences specifically with respect to the suppressed group dose requirement. The protein encoded by DIO1 gene is a thiol-requiring propyl thiouracil-sensitive oxidoreductase, which activates thyroid hormone by converting the T4 by outer ring deiodination (ORD) to bioactive 3,3',5-triiodothyronine T3, as well as degrading both hormones by inner ring deiodination (IRD). Understandably, functional mutations of the encoding gene are likely to inhibit the conversion and hence influence therapy using this prohormone.

Unlike the DIO1, none of the DIO2 or DIO3 showed any clear relationship with thyroxine dose requirement. In general there is lack of literature on the possible relationship between gene polymorphism and thyroxine dose requirement. Hence, the present study offers some insight in this area that need to be exploited further. Recently, some studies have shown that DIO2 Thr92Ala polymorphism predicts L-thyroxine dose to achieve target TSH (thyrotropin) levels in thyroidectomized patients. Since DIO2 converts T4 to T3 in several human tissues, including hypothalamus and pituitary, it therefore plays a pivotal role in the negative feedback regulation of TSH secretion. The

Thr92Ala has previously been identified and associated with decreased DIO2 enzymatic activity. It was also suggested that the DIO2 variant Thr92Ala predicts the T4 dose needed to achieve target thyrotropin levels in thyroidectomized patients (Torlontano *et al.*, 2008). While no association between the Thr92Ala polymorphism and serum thyroid hormone levels has been actually observed in humans, some studies in euthyroid patients have suggested that patients with the Ala92Ala phenotype needed higher thyroxine doses to achieve TSH suppression (Torlontano *et al.*, 2008). The same study observed that Ala/Ala homozygous patients needed a higher T4 dose as compared to patients carrying the Thr92 variant (X/Thr patients) according to a recessive genetic model (2.08±0.43 vs. 1.90±0.35 μg/kg; $P<0.05$). This difference was observable in the near-suppressed group ($P=0.002$), but not in the suppressed group. DIO2 Thr92Ala polymorphism seems to predict the need for higher T4 intake in thyroidectomized patients. If this finding is confirmed in additional studies, it may predict the T4 requirement to suppress TSH on the basis of the individual genetic background. Panicker et al suggest that the rare CC genotype of the rs224015(T>C) in the DIO2 gene is associated with impaired baseline psychological well-being as assessed by general health questionnaire in patients on T_4 and enhanced response to T_4/T_3 combination therapy, but did not affect serum thyroid hormone levels (Panicker *et al.*, 2009b).

A couple of studies have recently addressed the impact of DIO2 gene polymorphism on thyroxine treatment (Al-Azzam *et al.*, 2013; Peltsverger *et al.*, 2012). A study by Al-Azzam et al found no correlation between 4 DIO2 SNPs (rs225011, rs225012, rs2839858, rs7140952) and replacement doses of L-T4, but reported a statistical significance between rs7140952 and central obesity, as well as systolic and diastolic blood pressure. The dose of L-T4 was associated with lower levels of TSH, FT4, central obesity, body mass index and waist circumference (Al-Azzam *et al.*, 2013). Peltsverger et al found an association of the -258G/x DIO2 polymorphism variant with a decreased

rate of acute TSH-stimulated FT4 secretion with a normal T3 release from the thyroid gland consistent with a shift in the reaction equilibrium toward the product. They concluded that this polymorphism causes changes in the pattern of hormone secretion, pointing to the likelihood that common polymorphisms in DIO2 can finely affect the circulating levels of TH and might modulate the TH homeostasis. (Peltsverger *et al.*, 2012) .

BRAF V600E mutation in PTCs is associated with reduced expression of genes that are involved in iodine metabolism, such as the deiodinases. This may lead to alteration of the effectiveness of diagnostic and/or therapeutic use of radioiodine in BRAF PTCs (Durante *et al.*, 2007).

4.5.2. Association of thyroid stimulating hormone-beta with thyroxine dose requirement in DTC therapy

The potential role of the TSHβ gene in patient response to DTC therapy with thyroxine was investigated. Interestingly, while the majority of the variants did not appear to play a significant role, the rs1321109 showed some borderline association in the suppressed group. Notably, this SNP was strongly protective against acquiring DTC. Further studies in larger sample sizes are warranted to establish to what extent this variant may be involved in thyroxine dose requirement. Nonetheless, these findings seem indeed to suggest a potentially important role of this variant in the regulation of TSH. The potential role of TSH pathway in patient response to the therapy has also been addressed previously. So far, much research has focused mainly on changes in the TSHR and their impact on different thyroid parameters and/or clinical endpoints, such as insulin resistance. In particular activating mutation have been described which may interact with other genes in the TSH signalling pathways in causing cancer (Cetani *et al.*, 1999; Lado-Abeal *et al.*, 2010; Russo *et al.*, 1995; Russo *et al.*, 1999). Hence, it is now common knowledge that the TSHR and transthyretin gene polymorphisms play an

important role in intra-individual variation in TH bioactivity (Bartalena *et al.*, 1992; Curtis *et al.*, 1994; Episkopou *et al.*, 1993; Haymart *et al.*, 2008; Moses *et al.*, 1990; Sorensen *et al.*, 2008). Previous data on the response of the TSHR- Pro52Thr variant (rs2234919) to TSH stimulation in healthy blood donors is somewhat conflicting, with some studies failing to find any association with changes in serum TSH or iodothyronine levels (Peeters *et al.*, 2003b). This has been ascribed to subtle effects or low frequency of the 52Thr variant (Peeters *et al.*, 2003b). In healthy subjects, serum thyroid parameters show substantial inter-individual variability, whereas the intra-individual differences lie within a narrow range (Andersen *et al.*, 2002). Since the window between drug toxicity and suboptimal therapy for cancer treatment is often narrow, it can be assumed that inter-individual variations in drug metabolism might complicate therapy often as a result of genetic variations, in addition to environmental factors including food or iodine intake. For example, genetic heritability is thought to account for a greater portion of the variation in serum TSH, FT4 and FT3 levels, in a study of thyroid function tests involving Danish twins (Hansen *et al.*, 2004), and in serum TSH, free T4 and free T3 levels in another study from a Mexican-American population(Samollow *et al.*, 2004). Thus, mutations in the TSHβ have been linked to TSH resistance. Hence, assuming that mutations in TSHβ can cause RTH, this could explain why patients with such mutation may fail to respond to T4 therapy.

4.5.3. Association of Paired Box gene 8 with dose requirement in DTC therapy

The PAX8 gene presented the most interesting scenario whereby, fourteen SNPs were studied. In the ALL group, two variants rs2241975_CT+TT (p=0.01) and rs4849186_CC (p=0.01) were associated, while another rs3738913_A (p=0.07) was only weakly linked to decreased dose requirement. The association for rs4849186_CC was retained in NSG (p=0.03) group and became weaker in the SG (p=0.06). Also, rs1478_GG was significantly associated with elevated thyroxine dose in NSG. On the

other hand, in the SG, rs2241975_CT+TT (p=0.03), and the rs3738913_GA+AA (p=0.05) were significantly associated with elevated dose in rs2241975_CT+TT and lower dose requirement in rs3738913_ GA+AA.

This finding is interesting as it does not reflect on any predictable pattern as to how the different groups might be expected to respond in presence of an association of a variant with any one of them. Nonetheless, the results clearly demonstrate an association of the PAX8 gene with the variation in the patient requirement for thyroxine in the study population. There appears to be no data in the literature pertaining to the role of this gene in DTC therapy.

4.5.4. Association of sodium iodide symporter with dose response requirement in DTC therapy

In the NIS gene five variants were selected for the association study. However, no definable relationship could be drawn between the SNPs and thyroxine dose requirement for all three sub-groups. The literature is also lacking in this regard. Nonetheless, polymorphisms in these genes should still be considered to potentially influence the availability of the thyroid hormones at their various functional sites.

4.6. Summary: Gene polymorphism and dose response requirement to in DTC therapy

The present study has identified association of SNPs in three genes, DIO1, TSHβ and PAX8 genes with dose requirement. While these associations show variations between patients with partially and completely suppressed TSH hormones, it can nonetheless be inferred that these variants may play a role thyroxine dose requirement. Five of the associated SNPs (rs12084242 and rs1710952 in DIO1, rs2241975, rs3738913, and rs4849186 in PAX8) were intronic, two others (rs12095080 in DIO1 and rs1478 in PAX8) resided in the 3'UTR regions, and one (rs1321109) was found in the promoter region of TSHβ gene. As discussed above with respect to the association of gene

polymorphism with disease (Section 4.3.), mutations in these genic regions can influence the function of various gene and mRNA regulatory elements, such as transcriptosomes and miRNAs, leading to changes in the transcribed mRNA, and therefore possibly the synthesis of a non-functional protein product. Besides, since DIO1 is involved in the conversion of T4 to the more active T3, mutations in this gene may lead to changes in the production rate of the later.

4.7. Estimating effect size by the generalized linear models

The present study suggests that different genetic variants are associated with the risk of thyroid cancer and thyroxine dose requirements in the Saudi population. A question following the discovery of an association of parameters with disease remains as to what the actual effect size is of the influential factors. Generalized linear models (GLMs) have been designed to facilitate such estimations, and some studies have indicated that they are a useful method of estimating such effects (Baccini *et al.*, 2006; Carlin *et al.*, 2005; Hodge *et al.*, 1995). In the present study, GLM was employed to estimate the size of the effect of the variants identified to associate with both DTC manifestation and thyroxine dose requirement in DTC patients. The analysis led to a deviance of 3.8% for the association of the variables with disease and 6.0% for the association with dose requirement. These suggest therefore that the presence of some of the studied PAX8, TSHβ, and NIS variants as well as gender and smoking status could predict the chances of developing the disease and PAX8 and DIO1 variants could predict patient response to thyroxine therapy. These effects, although small, demonstrate that the studied polymorphisms explain some of the observed variances, although there are other, unknown, factors that may explain a greater proportion. Whether screening for these polymorphisms for the prediction of DTC risk or thyroid hormone replacement dose is necessary remains unclear.

4.8. Clinical impact on patient outcome

The present study identified a number of SNPs that were associated with DTC risk. Some previous studies have already indicated that mutations in some of the studied genes are linked to the histopathological features and prognosis of DTC. Thus, further studies are needed to understand the specific mechanisms/pathways and possible interactions with other known pathways by which these SNPs contribute to the risk of DTC. This would facilitate the exploration of the possibility of their involvement in such diseases and their clinical use as markers of DTC risk. Furthermore, if proven to be replicable, also the discovered variants associated with thyroxine dose requirement could be potentially used to screen patients in determining the appropriate dose for personalized medicine purposes.

5. General Summary and Conclusions

The present study was performed in three parts. The first was intended to establish the variants in six genes that are important in the thyroid hormone regulatory pathways. Several SNP with MAFs of >0.1 in six genes related to thyroxine functional and metabolic pathway were identified, presenting a large selection of variants of interest for further studies. A number of the discovered variants are non-synonymous SNP as well as indels which pointed to the likelihood that some of the studied genes may be related to disease in potentially causative fashion. Association studies were then performed on some of these variants by sequencing small DTC patient population and corresponding number of controls. These preliminary studies initially implicated a number of these variants in disease. However, these findings were somewhat distorted by the fact that a number of SNP were present only among the DTC patients which would mask the actual presence in the general populations. Then these findings were replicated using larger population sizes. Although the majority of the SNP failed the test of multivariate analysis, there were nonetheless some variants that remained significantly associated with disease. It was also interesting to test what role the discovered SNP may play in drug therapy of DTC. Interestingly, a number of SNPs also showed some association with patient response to standard thyroxine dose requirement. Importantly, the vast majority of the discovered SNPs were intronic or reside in the untranslated regions of the respective genes. This fact does not necessarily render them less important in their potential impact on disease and drug therapy thereof, since it is now well-appreciated that there is still a lot to learn about the importance of these genomic entities. Most importantly, this study has comprehensively investigated a large number of key genes for variations that probably play a decisive role in thyroid cancer pathogenesis and thyroxine dose requirements. This work was essential to define the potential role or lack of such a role of these genes in the development of thyroid

cancer and thyroxine dose requirement. Efforts from molecular thyroid cancer medicine have now shown great promises in helping improve the management of thyroid cancer by using gene mutation as a prognostic molecular marker. Also, the results showed that thyroid cancer is unlike simple Mendelian traits, such as cystic fibrosis, where the genetic factor may explain all the deviance. Other factors, such as SNPs, gender and smoking were found to influence the risk of thyroid cancer and response to thyroxine dose. Hence, there are clearly many other contributory factors that require exploration.

To date, the proven published mutation can only detect probably about 70% of DTC cases. Hence, our extensive search for SNPs led to the finding of a number of important polymorphisms that are associated with thyroid cancer or showed a strong association with thyroxine dose that may potentially detect the 30% cases in the future. Besides, this study also sets the stage for further work in these genes and their roles in thyroid cancer pathogenesis as well as thyroid hormone Pharmacogenetics.

The present study has raised a number of questions with respect to the potential relevance of the studied genes in thyroid cancer risk and differences in the way patients respond to the drug therapy of DTC. Several approaches can be envisaged to try and answer some of these questions. Notably, several of the associated variants reside in non-coding regions of the genes. One approach to understanding the role of such SNPs may involve testing the function of the related regulatory entities, such as the transcription factors, in presence of these gene variants. Besides, a number of haplotypes created from these SNPs were even more significantly associated with disease than the individual SNPs themselves. I am interested in testing the impact of their cumulative effect on the extent of disease. The prognosis of PTC patients is considered to be good. Accordingly it would be of interest to investigate whether in the presence of more than one associated SNP in form of the observed haplotypes would provide an insight into the aggressiveness of the tumour, in consideration of parameters,

such as frequency of radiation therapy and metastasis. Another set of studies may involve creating animal models harbouring these mutations to evaluate the possibility of their acquiring the disease.

References

Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, *et al.* (2010). A map of human genome variation from population-scale sequencing. *Nature* **467**(7319)**:** 1061-1073.

Abubaker J, Jehan Z, Bavi P, Sultana M, Al-Harbi S, Ibrahim M, *et al.* (2008). Clinicopathological analysis of papillary thyroid cancer with PIK3CA alterations in a Middle Eastern population. *The Journal of clinical endocrinology and metabolism* **93**(2)**:** 611-618.

Al-Azzam SI, Alkhateeb AM, Al-Azzeh O, Alzoubi KH, Khabour OF (2013). The Role of Type II Deiodinase Polymorphisms in Clinical Management of Hypothyroid Patients Treated with Levothyroxine. *Exp Clin Endocrinol Diabetes*.

Albrechtsen N, Dornreiter I, Grosse F, Kim E, Wiesmuller L, Deppert W (1999). Maintenance of genomic integrity by p53: complementary roles for activated and non-activated p53. *Oncogene* **18**(53)**:** 7706-7717.

Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, *et al.* (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* **467**(7311)**:** 52-58.

Alvarez JC, Mendez JC, Moris C (2000). [Thyroid cancer treatment]. *Annales d'oto-laryngologie et de chirurgie cervico faciale : bulletin de la Societe d'oto-laryngologie des hopitaux de Paris* **117**(1)**:** 40-44.

Alves IL, Divino CM, Schussler GC, Altland K, Almeida MR, Palha JA, *et al.* (1993). Thyroxine binding in a TTR Met 119 kindred. *J Clin Endocrinol Metab* **77**(2)**:** 484-488.

Andersen S, Pedersen KM, Bruun NH, Laurberg P (2002). Narrow individual variations in serum T(4) and T(3) in normal subjects: a clue to the understanding of subclinical thyroid disease. *J Clin Endocrinol Metab* **87**(3)**:** 1068-1072.

Ansari M, Krajinovic M (2007). Pharmacogenomics of acute leukemia. *Pharmacogenomics* **8**(7)**:** 817-834.

Arnaldi LA, Borra RC, Maciel RM, Cerutti JM (2005). Gene expression profiles reveal that DCN, DIO1, and DIO2 are underexpressed in benign and malignant thyroid tumors. *Thyroid* **15**(3)**:** 210-221.

Arnaud-Lopez L, Usala G, Ceresini G, Mitchell BD, Pilia MG, Piras MG, *et al.* (2008). Phosphodiesterase 8B gene variants are associated with serum TSH levels and thyroid function. *Am J Hum Genet* **82**(6)**:** 1270-1280.

Arturi F, Russo D, Schlumberger M, du Villard JA, Caillou B, Vigneri P, *et al.* (1998). Iodide symporter gene expression in human thyroid tumors. *The Journal of clinical endocrinology and metabolism* **83**(7)**:** 2493-2496.

Azad RM (2011). Abnormal serum thyroid hormones concentration with healthy functional gland: a review on the metabolic role of thyroid hormones transporter proteins. *Pak J Biol Sci* **14**(5)**:** 313-326.

Baccini M, Biggeri A, Accetta G, Lagazio C, Lerxtundi A, Schwartz J (2006). Comparison of alternative modelling techniques in estimating short-term effect of air pollution with application to the Italian meta-analysis data (MISA Study). *Epidemiol Prev* **30**(4-5)**:** 279-288.

Bai JL, Zheng MH, Xia X, Ter-Minassian M, Chen YP, Chen F (2009). MTHFR C677T polymorphism contributes to prostate cancer risk among Caucasians: A meta-analysis of 3511 cases and 2762 controls. *Eur J Cancer* **45**(8)**:** 1443-1449.

Barbaro D, Boni G, Meucci G, Simi U, Lapi P, Orsini P, *et al.* (2003). Radioiodine treatment with 30 mCi after recombinant human thyrotropin stimulation in thyroid cancer: effectiveness for postsurgical remnants ablation and possible role of iodine content in L-thyroxine in the outcome of ablation. *J Clin Endocrinol Metab* **88**(9)**:** 4110-4115.

Barbaro D, Boni G, Meucci G, Simi U, Lapi P, Orsini P, *et al.* (2006). Recombinant human thyroid-stimulating hormone is effective for radioiodine ablation of post-surgical thyroid remnants. *Nucl Med Commun* **27**(8)**:** 627-632.

Baron Y, Corre S, Mouchet N, Vaulont S, Prince S, Galibert MD (2012). USF-1 is critical for maintaining genome integrity in response to UV-induced DNA photolesions. *PLoS Genet* **8**(1)**:** e1002470.

Bartalena L, Robbins J (1992). Variations in thyroid hormone transport proteins and their clinical implications. *Thyroid* **2**(3)**:** 237-245.

Bartha T, Kim SW, Salvatore D, Gereben B, Tu HM, Harney JW, *et al.* (2000). Characterization of the 5'-flanking and 5'-untranslated regions of the cyclic adenosine 3',5'-monophosphate-responsive human type 2 iodothyronine deiodinase gene. *Endocrinology* **141**(1)**:** 229-237.

Barton NH (2000). Genetic hitchhiking. *Philos Trans R Soc Lond B Biol Sci* **355**(1403)**:** 1553-1562.

Barton NH, Etheridge AM, Kelleher J, Veber A (2013). Genetic hitchhiking in spatially extended populations. *Theor Popul Biol*.

Bianco AC, Kim BW (2006). Deiodinases: implications of the local control of thyroid hormone action. *J Clin Invest* **116**(10)**:** 2571-2579.

Bianco AC, Larsen PR (2005). Cellular and structural biology of the deiodinases. *Thyroid* **15**(8)**:** 777-786.

Bianco AC, Salvatore D, Gereben B, Berry MJ, Larsen PR (2002). Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases. *Endocr Rev* **23**(1)**:** 38-89.

Bigler J, Eisenman RN (1995). Novel location and function of a thyroid hormone response element. *EMBO J* **14**(22)**:** 5710-5723.

Bikker H, Waelkens JJ, Bravenboer B, de Vijlder JJ (1996). Congenital hypothyroidism caused by a premature termination signal in exon 10 of the human thyroid peroxidase gene. *J Clin Endocrinol Metab* **81**(6)**:** 2076-2079.

Biondi B, Cooper DS (2010). Benefits of thyrotropin suppression versus the risks of adverse effects in differentiated thyroid cancer. *Thyroid : official journal of the American Thyroid Association* **20**(2)**:** 135-146.

Bishop JA, Sharma R, Westra WH (2011). PAX8 immunostaining of anaplastic thyroid carcinoma: a reliable means of discerning thyroid origin for undifferentiated tumors of the head and neck. *Human pathology* **42**(12)**:** 1873-1877.

Boutin A, Allen MD, Geras-Raaka E, Huang W, Neumann S, Gershengorn MC (2011). Thyrotropin receptor stimulates internalization-independent persistent phosphoinositide signaling. *Mol Pharmacol* **80**(2)**:** 240-246.

Braverman LE (1994). Deiodination of thyroid hormones. A 30 year perspective (Berthold Memorial Award Lecture 1994). *Exp Clin Endocrinol* **102**(5)**:** 355-363.

Bray F, Ren JS, Masuyer E, Ferlay J (2013). Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *International journal of cancer. Journal international du cancer* **132**(5)**:** 1133-1145.

Brockmoller J, Tzvetkov MV (2008). Pharmacogenetics: data, concepts and tools to improve drug discovery and drug treatment. *Eur J Clin Pharmacol* **64**(2)**:** 133-157.

Caillou B, Troalen F, Baudin E, Talbot M, Filetti S, Schlumberger M, *et al.* (1998). Na+/I-symporter distribution in human thyroid tissues: an immunohistochemical study. *J Clin Endocrinol Metab* **83**(11)**:** 4102-4106.

Calebiro D, de Filippis T, Lucchi S, Covino C, Panigone S, Beck-Peccoz P, *et al.* (2005). Intracellular entrapment of wild-type TSH receptor by oligomerization with mutants linked to dominant TSH resistance. *Hum Mol Genet* **14**(20)**:** 2991-3002.

Callebaut I, Curcio-Morelli C, Mornon JP, Gereben B, Buettner C, Huang S, *et al.* (2003). The iodothyronine selenodeiodinases are thioredoxin-fold family proteins containing a glycoside hydrolase clan GH-A-like structure. *J Biol Chem* **278**(38)**:** 36887-36896.

Canani LH, Capp C, Dora JM, Meyer EL, Wagner MS, Harney JW, *et al.* (2005a). The type 2 deiodinase A/G (Thr92Ala) polymorphism is associated with decreased enzyme velocity and increased insulin resistance in patients with type 2 diabetes mellitus. *J Clin Endocrinol Metab* **90**(6)**:** 3472-3478.

Canani LH, Capp C, Dora JM, Meyer EL, Wagner MS, Harney JW, *et al.* (2005b). The type 2 deiodinase A/G (Thr92Ala) polymorphism is associated with decreased enzyme velocity and increased insulin resistance in patients with type 2 diabetes mellitus. *J Clin Endocrinol Metab* **90**(6)**:** 3472-3478.

Cangul H, Aycan Z, Saglam H, Forman JR, Cetinkaya S, Tarim O, *et al.* (2012). TSHR is the main causative locus in autosomal recessively inherited thyroid dysgenesis. *J Pediatr Endocrinol Metab* **25**(5-6)**:** 419-426.

Cangul H, Morgan NV, Forman JR, Saglam H, Aycan Z, Yakut T, *et al.* (2010). Novel TSHR mutations in consanguineous families with congenital nongoitrous hypothyroidism. *Clin Endocrinol (Oxf)* **73**(5)**:** 671-677.

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, *et al.* (1999). Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* **22**(3)**:** 231-238.

Carlin JB, Gurrin LC, Sterne JA, Morley R, Dwyer T (2005). Regression models for twin studies: a critical review. *Int J Epidemiol* **34**(5)**:** 1089-1099.

Carre A, Hamza RT, Kariyawasam D, Guillot L, Teissier R, Tron E, *et al.* (2014). A Novel FOXE1 Mutation (R73S) in Bamforth-Lazarus Syndrome Causing Increased Thyroidal Gene Expression. *Thyroid*.

Cassio A, Nicoletti A, Rizzello A, Zazzetta E, Bal M, Baldazzi L (2013). Current loss-of-function mutations in the thyrotropin receptor gene: when to investigate, clinical effects, and treatment. *J Clin Res Pediatr Endocrinol* **5 Suppl 1:** 29-39.

Castro MR, Bergert ER, Goellner JR, Hay ID, Morris JC (2001). Immunohistochemical analysis of sodium iodide symporter expression in metastatic differentiated thyroid cancer: correlation with radioiodine uptake. *J Clin Endocrinol Metab* **86**(11)**:** 5627-5632.

Castro P, Rebocho AP, Soares RJ, Magalhaes J, Roque L, Trovisco V, *et al.* (2006). PAX8-PPARgamma rearrangement is frequently detected in the follicular variant of papillary thyroid carcinoma. *The Journal of clinical endocrinology and metabolism* **91**(1)**:** 213-220.

Cetani F, Tonacchera M, Pinchera A, Barsacchi R, Basolo F, Miccoli P, *et al.* (1999). Genetic analysis of the TSH receptor gene in differentiated human thyroid carcinomas. *J Endocrinol Invest* **22**(4)**:** 273-278.

Chakravarti A (1999). Population genetics--making sense out of sequence. *Nat Genet* **21**(1 Suppl)**:** 56-60.

Chen ML (2006). Ethnic or racial differences revisited: impact of dosage regimen and dosage form on pharmacokinetics and pharmacodynamics. *Clin Pharmacokinet* **45**(10)**:** 957-964.

Chong E, Ensom MH (2003). Pharmacogenetics of the proton pump inhibitors: a systematic review. *Pharmacotherapy* **23**(4)**:** 460-471.

Chopra IJ (1976). An assessment of daily production and significance of thyroidal secretion of 3, 3', 5'-triiodothyronine (reverse T3) in man. *J Clin Invest* **58**(1)**:** 32-40.

Chou CK, Chen RF, Chou FF, Chang HW, Chen YJ, Lee YF, *et al.* (2010). miR-146b is highly expressed in adult papillary thyroid carcinomas with high risk features including extrathyroidal invasion and the BRAF(V600E) mutation. *Thyroid* **20**(5)**:** 489-494.

Christoffolete MA, Ribeiro R, Singru P, Fekete C, da Silva WS, Gordon DF, *et al.* (2006). Atypical expression of type 2 iodothyronine deiodinase in thyrotrophs explains the thyroxine-mediated pituitary thyrotropin feedback mechanism. *Endocrinology* **147**(4)**:** 1735-1743.

Cobi R, Gharib H (2001). AACE/AAES MEDICAL/SURGICAL GUIDELINES FOR CLINICAL PRACTICE: MANAGEMENT OF THYROID CARCINOMA. *ENDOCRINE PRACTICE* **Vol. 7 No. 3** (May/June)**:** 202-220.

Collins FS, Guyer MS, Charkravarti A (1997). Variations on a theme: cataloging human DNA sequence variation. *Science* **278**(5343)**:** 1580-1581.

Congdon T, Nguyen LQ, Nogueira CR, Habiby RL, Medeiros-Neto G, Kopp P (2001). A novel mutation (Q40P) in PAX8 associated with congenital hypothyroidism and thyroid hypoplasia: evidence for phenotypic variability in mother and child. *The Journal of clinical endocrinology and metabolism* **86**(8)**:** 3962-3967.

Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, *et al.* (2009). Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* **19**(11)**:** 1167-1214.

Cooper DS, Specker B, Ho M, Sperling M, Ladenson PW, Ross DS, *et al.* (1998). Thyrotropin suppression and disease progression in patients with differentiated thyroid cancer: results from the National Thyroid Cancer Treatment Cooperative Registry. *Thyroid* **8**(9)**:** 737-744.

Couch RM, Dean HJ, Winter JS (1985). Congenital hypothyroidism caused by defective iodide transport. *The Journal of pediatrics* **106**(6)**:** 950-953.

Curcio-Morelli C, Gereben B, Zavacki AM, Kim BW, Huang S, Harney JW, *et al.* (2003). In vivo dimerization of types 1, 2, and 3 iodothyronine selenodeiodinases. *Endocrinology* **144**(3)**:** 937-946.

Curtis AJ, Scrimshaw BJ, Topliss DJ, Stockigt JR, George PM, Barlow JW (1994). Thyroxine binding by human transthyretin variants: mutations at position 119, but not position 54, increase thyroxine binding affinity. *J Clin Endocrinol Metab* **78**(2)**:** 459-462.

Darras VM, Van Herck SL (2012). Iodothyronine deiodinase structure and function: from ascidians to humans. *J Endocrinol* **215**(2)**:** 189-206.

Davidsen ML, Dalhoff K, Schmiegelow K (2008). Pharmacogenetics influence treatment efficacy in childhood acute lymphoblastic leukemia. *J Pediatr Hematol Oncol* **30**(11)**:** 831-849.

Davies L, Welch HG (2006). Increasing incidence of thyroid cancer in the United States, 1973-2002. *Jama* **295**(18)**:** 2164-2167.

De Benedictis G, Falcone E, Ruffolo R, Spadafora P, Carotenuto L (1996). Linkage disequilibrium studies in the thyroid peroxidase and apolipoprotein B genes. *Hum Biol* **68**(1)**:** 147-154.

de Jong FJ, Peeters RP, den Heijer T, van der Deure WM, Hofman A, Uitterlinden AG, *et al.* (2007). The association of polymorphisms in the type 1 and 2 deiodinase genes with circulating thyroid hormone parameters and atrophy of the medial temporal lobe. *J Clin Endocrinol Metab* **92**(2)**:** 636-640.

de la Chapelle A (2013). Unraveling the genetic predisposition to differentiated thyroid carcinoma. *J Clin Endocrinol Metab* **98**(10)**:** 3974-3976.

Dechairo BM, Zabaneh D, Collins J, Brand O, Dawson GJ, Green AP, *et al.* (2005). Association of the TSHR gene with Graves' disease: the first disease specific locus. *Eur J Hum Genet* **13**(11)**:** 1223-1230.

Desai MM, Walczak AM, Fisher DS (2013). Genetic diversity and the structure of genealogies in rapidly adapting populations. *Genetics* **193**(2)**:** 565-585.

Di Palma T, Filippone MG, Pierantoni GM, Fusco A, Soddu S, Zannini M (2013). Pax8 has a critical role in epithelial cell survival and proliferation. *Cell death & disease* **4:** e729.

Diallo-Krou E, Yu J, Colby LA, Inoki K, Wilkinson JE, Thomas DG, *et al.* (2009). Paired box gene 8-peroxisome proliferator-activated receptor-gamma fusion protein and loss of phosphatase and tensin homolog synergistically cause thyroid hyperplasia in transgenic mice. *Endocrinology* **150**(11)**:** 5181-5190.

Dong LM, Potter JD, White E, Ulrich CM, Cardon LR, Peters U (2008). Genetic susceptibility to cancer: the role of polymorphisms in candidate genes. *Jama* **299**(20)**:** 2423-2436.

Dumitrescu AM, Liao XH, Abdullah MS, Lado-Abeal J, Majed FA, Moeller LC, *et al.* (2005). Mutations in SECISBP2 result in abnormal thyroid hormone metabolism. *Nat Genet* **37**(11)**:** 1247-1252.

Durante C, Puxeddu E, Ferretti E, Morisi R, Moretti S, Bruno R, *et al.* (2007). BRAF mutations in papillary thyroid carcinomas inhibit genes involved in iodine metabolism. *J Clin Endocrinol Metab* **92**(7)**:** 2840-2843.

Dwight T, Thoppe SR, Foukakis T, Lui WO, Wallin G, Hoog A, *et al.* (2003). Involvement of the PAX8/peroxisome proliferator-activated receptor gamma rearrangement in follicular thyroid tumors. *J Clin Endocrinol Metab* **88**(9)**:** 4440-4445.

Eberhardt NL, Grebe SK, McIver B, Reddi HV (2010). The role of the PAX8/PPARgamma fusion oncogene in the pathogenesis of follicular thyroid cancer. *Mol Cell Endocrinol* **321**(1)**:** 50-56.

Ehrenborg E, Skogsberg J, Ruotolo G, Large V, Eriksson P, Arner P, *et al.* (2000). The Q/E27 polymorphism in the beta2-adrenoceptor gene is associated with increased body weight and dyslipoproteinaemia involving triglyceride-rich lipoproteins. *J Intern Med* **247**(6)**:** 651-656.

Episkopou V, Maeda S, Nishiguchi S, Shimada K, Gaitanaris GA, Gottesman ME, *et al.* (1993). Disruption of the transthyretin gene results in mice with depressed levels of plasma retinol and thyroid hormone. *Proc Natl Acad Sci U S A* **90**(6)**:** 2375-2379.

Fagin JA, Tang SH, Zeki K, Di Lauro R, Fusco A, Gonsky R (1996). Reexpression of thyroid peroxidase in a derivative of an undifferentiated thyroid carcinoma cell line by introduction of wild-type p53. *Cancer Res* **56**(4)**:** 765-771.

Farid NR, Shi Y, Zou M (1994). Molecular basis of thyroid cancer. *Endocr Rev* **15**(2)**:** 202-232.

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, *et al.* (2013). Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* **49**(6)**:** 1374-1403.

Filetti S, Bidart JM, Arturi F, Caillou B, Russo D, Schlumberger M (1999). Sodium/iodide symporter: a key transport system in thyroid cancer cell metabolism. *Eur J Endocrinol* **141**(5)**:** 443-457.

Flaxman SM, Feder JL, Nosil P (2012). Spatially explicit models of divergence and genome hitchhiking. *J Evol Biol* **25**(12)**:** 2633-2650.

Fliess J (1981). Statistical Methods for rates and proportions. 17.

Franco AT, Malaguarnera R, Refetoff S, Liao XH, Lundsmith E, Kimura S, *et al.* (2011). Thyrotrophin receptor signaling dependence of Braf-induced thyroid tumor initiation in mice. *Proc Natl Acad Sci U S A* **108**(4)**:** 1615-1620.

Friesema EC, Jansen J, Milici C, Visser TJ (2005). Thyroid hormone transporters. *Vitam Horm* **70:** 137-167.

Gabriel EM, Bergert ER, Grant CS, van Heerden JA, Thompson GB, Morris JC (1999). Germline polymorphism of codon 727 of human thyroid-stimulating hormone receptor is associated with toxic multinodular goiter. *J Clin Endocrinol Metab* **84**(9)**:** 3328-3335.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, *et al.* (2002). The structure of haplotype blocks in the human genome. *Science* **296**(5576)**:** 2225-2229.

Gao L, Nieters A, Brenner H (2008). Meta-analysis: tumour invasion-related genetic polymorphisms and gastric cancer susceptibility. *Aliment Pharmacol Ther* **28**(5)**:** 565-573.

Geraldo MV, Yamashita AS, Kimura ET (2012). MicroRNA miR-146b-5p regulates signal transduction of TGF-beta by repressing SMAD4 in thyroid cancer. *Oncogene* **31**(15)**:** 1910-1922.

Gereben B, Zeold A, Dentice M, Salvatore D, Bianco AC (2008). Activation and inactivation of thyroid hormone by deiodinases: local action with general consequences. *Cell Mol Life Sci* **65**(4)**:** 570-590.

Gilliland FD, Hunt WC, Morris DM, Key CR (1997). Prognostic factors for thyroid carcinoma. A population-based study of 15,698 cases from the Surveillance, Epidemiology and End Results (SEER) program 1973-1991. *Cancer* **79**(3)**:** 564-573.

Goffredo P, Roman SA, Sosa JA (2012). Hurthle cell carcinoma: A population-level analysis of 3311 patients. *Cancer*.

Gomez Saez JM (2011). Diagnostic and prognostic markers in differentiated thyroid cancer. *Curr Genomics* **12**(8)**:** 597-608.

Grarup N, Andersen MK, Andreasen CH, Albrechtsen A, Borch-Johnsen K, Jorgensen T, *et al.* (2007). Studies of the common DIO2 Thr92Ala polymorphism and metabolic phenotypes in 7342 Danish white subjects. *The Journal of clinical endocrinology and metabolism* **92**(1)**:** 363-366.

Grasberger H, Ringkananont U, Lefrancois P, Abramowicz M, Vassart G, Refetoff S (2005). Thyroid transcription factor 1 rescues PAX8/p300 synergism impaired by a natural PAX8 paired domain mutation with dominant negative activity. *Molecular endocrinology* **19**(7)**:** 1779-1791.

Greco A, Borrello MG, Miranda C, Degl'Innocenti D, Pierotti MA (2009a). Molecular pathology of differentiated thyroid cancer. *Q J Nucl Med Mol Imaging* **53**(5)**:** 440-453.

Greco A, Borrello MG, Miranda C, Degl'Innocenti D, Pierotti MA (2009b). Molecular pathology of differentiated thyroid cancer. *The quarterly journal of nuclear medicine and molecular imaging : official publication of the Italian Association of Nuclear Medicine* **53**(5)**:** 440-453.

Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Masson G, He H, *et al.* (2012). Discovery of common variants associated with low TSH levels and thyroid cancer risk. *Nat Genet* **44**(3)**:** 319-322.

Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Sigurdsson A, Bergthorsson JT, *et al.* (2009). Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nat Genet* **41**(4)**:** 460-464.

Gudnason V, Kakko S, Nicaud V, Savolainen MJ, Kesaniemi YA, Tahvanainen E, *et al.* (1999). Cholesteryl ester transfer protein gene effect on CETP activity and plasma high-density lipoprotein in European populations. The EARS Group. *Eur J Clin Invest* **29**(2)**:** 116-128.

Guerrieri F, Piconese S, Lacoste C, Schinzari V, Testoni B, Valogne Y, *et al.* (2013). The sodium/iodide symporter NIS is a transcriptional target of the p53-family members in liver cancer cells. *Cell death & disease* **4:** e807.

Guo TW, Zhang FC, Yang MS, Gao XC, Bian L, Duan SW, *et al.* (2004). Positive association of the DIO2 (deiodinase type 2) gene with mental retardation in the iodine-deficient areas of China. *J Med Genet* **41**(8)**:** 585-590.

Hansen PS, Brix TH, Sorensen TI, Kyvik KO, Hegedus L (2004). Major genetic influence on the regulation of the pituitary-thyroid axis: a study of healthy Danish twins. *The Journal of clinical endocrinology and metabolism* **89**(3)**:** 1181-1187.

Hansen PS, van der Deure WM, Peeters RP, Iachine I, Fenger M, Sorensen TI, *et al.* (2007). The impact of a TSH receptor gene polymorphism on thyroid-related phenotypes in a healthy Danish twin population. *Clin Endocrinol (Oxf)* **66**(6)**:** 827-832.

Haymart MR, Repplinger DJ, Leverson GE, Elson DF, Sippel RS, Jaume JC, *et al.* (2008). Higher serum thyroid stimulating hormone level in thyroid nodule patients is associated with greater risks of differentiated thyroid cancer and advanced tumor stage. *J Clin Endocrinol Metab* **93**(3)**:** 809-814.

He H, Bronisz A, Liyanarachchi S, Nagy R, Li W, Huang Y, *et al.* (2013a). SRGAP1 is a candidate gene for papillary thyroid carcinoma susceptibility. *J Clin Endocrinol Metab* **98**(5)**:** E973-980.

He H, Jazdzewski K, Li W, Liyanarachchi S, Nagy R, Volinia S, *et al.* (2005). The role of microRNA genes in papillary thyroid carcinoma. *Proc Natl Acad Sci U S A* **102**(52)**:** 19075-19080.

He H, Li W, Wu D, Nagy R, Liyanarachchi S, Akagi K, *et al.* (2013b). Ultra-rare mutation in long-range enhancer predisposes to thyroid carcinoma with high penetrance. *PLoS One* **8**(5)**:** e61920.

Heemstra KA, Hoftijzer HC, van der Deure WM, Peeters RP, Fliers E, Appelhof BC, *et al.* (2008a). The Thr92Ala polymorphism in the type 2 deiodinase is not associated with thyroxine dose in athyroid patients or patients with Hashimoto thyroiditis. *Clin Endocrinol (Oxf)*.

Heemstra KA, van der Deure WM, Peeters RP, Hamdy NA, Stokkel MP, Corssmit EP, *et al.* (2008b). Thyroid hormone independent associations between serum TSH levels and indicators of bone turnover in cured patients with differentiated thyroid carcinoma. *Eur J Endocrinol* **159**(1)**:** 69-76.

Hemminki K, Granstrom C, Sundquist J, Bermejo JL (2006). The updated Swedish family-cancer database used to assess familial risks of prostate cancer during rapidly increasing incidence. *Hered Cancer Clin Pract* **4**(4)**:** 186-192.

Hermanns P, Grasberger H, Refetoff S, Pohlenz J (2011). Mutations in the NKX2.5 gene and the PAX8 promoter in a girl with thyroid dysgenesis. *The Journal of clinical endocrinology and metabolism* **96**(6)**:** E977-981.

Hershman JM, Due DT, Sharp B, My L, Kent JR, Binh LN, *et al.* (1983). Endemic goiter in Vietnam. *J Clin Endocrinol Metab* **57**(2)**:** 243-249.

Hesch RD, Brunner G, Soling HD (1975). Conversion of thyroxine (T4) and triiodothyronine (T3) and the subcellular localisation of the converting enzyme. *Clin Chim Acta* **59**(2)**:** 209-213.

Hodge WG, Whitcher JP, Satariano W (1995). Risk factors for age-related cataracts. *Epidemiol Rev* **17**(2)**:** 336-346.

Hu W (2009). The role of p53 gene family in reproduction. *Cold Spring Harbor perspectives in biology* **1**(6)**:** a001073.

Jazdzewski K, Boguslawska J, Jendrzejewski J, Liyanarachchi S, Pachucki J, Wardyn KA, *et al.* (2011). Thyroid hormone receptor beta (THRB) is a major target gene for microRNAs deregulated in papillary thyroid carcinoma (PTC). *J Clin Endocrinol Metab* **96**(3)**:** E546-553.

Jemal A, Center MM, DeSantis C, Ward EM (2010a). Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol Biomarkers Prev* **19**(8)**:** 1893-1907.

Jemal A, Clegg LX, Ward E, Ries LA, Wu X, Jamison PM, *et al.* (2004). Annual report to the nation on the status of cancer, 1975-2001, with a special feature regarding survival. *Cancer* **101**(1)**:** 3-27.

Jemal A, Siegel R, Xu J, Ward E (2010b). Cancer statistics, 2010. *CA Cancer J Clin* **60**(5)**:** 277-300.

Joba W, Spitzweg C, Schriever K, Heufelder AE (1999). Analysis of human sodium/iodide symporter, thyroid transcription factor-1, and paired-box-protein-8 gene expression in benign thyroid diseases. *Thyroid* **9**(5)**:** 455-466.

Jones AM, Howarth KM, Martin L, Gorman M, Mihai R, Moss L, *et al.* (2012). Thyroid cancer susceptibility polymorphisms: confirmation of loci on chromosomes 9q22 and 14q13, validation of a recessive 8q24 locus and failure to replicate a locus on 5q24. *J Med Genet* **49**(3)**:** 158-163.

Kakko S, Tamminen M, Paivansalo M, Kauma H, Rantala AO, Lilja M, *et al.* (2001). Variation at the cholesteryl ester transfer protein gene in relation to plasma high density lipoproteins cholesterol levels and carotid intima-media thickness. *Eur J Clin Invest* **31**(7)**:** 593-602.

Kilfoy BA, Zheng T, Holford TR, Han X, Ward MH, Sjodin A, *et al.* (2009). International patterns and trends in thyroid cancer incidence, 1973-2002. *Cancer Causes Control* **20**(5)**:** 525-531.

Kim Y (2013). Stochastic patterns of polymorphism after a selective sweep over a subdivided population. *Genet Res (Camb)***:** 1-11.

Kloos RT, Eng C, Evans DB, Francis GL, Gagel RF, Gharib H, *et al.* (2009). Medullary thyroid cancer: management guidelines of the American Thyroid Association. *Thyroid* **19**(6)**:** 565-612.

Kogai T, Brent GA (2012). The sodium iodide symporter (NIS): regulation and approaches to targeting for cancer therapeutics. *Pharmacology & therapeutics* **135**(3)**:** 355-370.

Kohrle J (2000). The deiodinase family: selenoenzymes regulating thyroid hormone availability and action. *Cell Mol Life Sci* **57**(13-14)**:** 1853-1863.

Kohrle J (1999). Local activation and inactivation of thyroid hormones: the deiodinase family. *Mol Cell Endocrinol* **151**(1-2)**:** 103-119.

Kohrle J (2007). Thyroid hormone transporters in health and disease: advances in thyroid hormone deiodination. *Best Pract Res Clin Endocrinol Metab* **21**(2)**:** 173-191.

Kohrle J, Jakob F, Contempre B, Dumont JE (2005). Selenium, the thyroid, and the endocrine system. *Endocr Rev* **26**(7)**:** 944-984.

Kohrle J, Spanka M, Irmscher K, Hesch RD (1988). Flavonoid effects on transport, metabolism and action of thyroid hormones. *Prog Clin Biol Res* **280:** 323-340.

Kojic KL, Kojic SL, Wiseman SM (2012). Differentiated thyroid cancers: a comprehensive review of novel targeted therapies. *Expert Rev Anticancer Ther* **12**(3)**:** 345-357.

Komatsu M, Takahashi T, Takahashi I, Nakamura M, Takada G (2001). Thyroid dysgenesis caused by PAX8 mutation: the hypermutability with CpG dinucleotides at codon 31. *The Journal of pediatrics* **139**(4)**:** 597-599.

Kondo T, Ezzat S, Asa SL (2006). Pathogenetic mechanisms in thyroid follicular-cell neoplasia. *Nat Rev Cancer* **6**(4)**:** 292-306.

Kosugi S, Bhayana S, Dean HJ (1999). A novel mutation in the sodium/iodide symporter gene in the largest family with iodide transport defect. *The Journal of clinical endocrinology and metabolism* **84**(9)**:** 3248-3253.

Kosugi S, Inoue S, Matsuda A, Jhiang SM (1998). Novel, missense and loss-of-function mutations in the sodium/iodide symporter gene causing iodide transport defect in three Japanese patients. *The Journal of clinical endocrinology and metabolism* **83**(9)**:** 3373-3376.

Krohn K, Paschke R (2002). Somatic mutations in thyroid nodular disease. *Mol Genet Metab* **75**(3)**:** 202-208.

Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, *et al.* (2000). PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. *Science* **289**(5483)**:** 1357-1360.

Kruglyak L (1997). The use of a genetic map of biallelic markers in linkage studies. *Nat Genet* **17**(1)**:** 21-24.

Kudzi W, Dodoo AN, Mills JJ (2009). Characterisation of CYP2C8, CYP2C9 and CYP2C19 polymorphisms in a Ghanaian population. *BMC Med Genet* **10:** 124.

Kudzi W, Dodoo AN, Mills JJ (2010). Genetic polymorphisms in MDR1, CYP3A4 and CYP3A5 genes in a Ghanaian population: a plausible explanation for altered metabolism of ivermectin in humans? *BMC Med Genet* **11:** 111.

Kuiper GG, Kester MH, Peeters RP, Visser TJ (2005). Biochemical mechanisms of thyroid hormone deiodination. *Thyroid* **15**(8)**:** 787-798.

Kuivenhoven JA, de Knijff P, Boer JM, Smalheer HA, Botma GJ, Seidell JC, *et al.* (1997). Heterogeneity at the CETP gene locus. Influence on plasma CETP concentrations and HDL cholesterol levels. *Arterioscler Thromb Vasc Biol* **17**(3)**:** 560-568.

Kumar D (2008). Disorders of the genome architecture: a review. *Genomic Med* **2**(3-4)**:** 69-76.

Kurose K, Sugiyama E, Saito Y (2012). Population differences in major functional polymorphisms of pharmacokinetics/pharmacodynamics-related genes in Eastern Asians and Europeans: implications in the clinical trials for novel drug development. *Drug Metab Pharmacokinet* **27**(1)**:** 9-54.

Lado-Abeal J, Celestino R, Bravo SB, Garcia-Rendueles ME, de la Calzada J, Castro I, *et al.* (2010). Identification of a paired box gene 8-peroxisome proliferator-activated receptor gamma (PAX8-PPARgamma) rearrangement mosaicism in a patient with an autonomous functioning follicular thyroid carcinoma bearing an activating mutation in the TSH receptor. *Endocr Relat Cancer* **17**(3)**:** 599-610.

Lake SL, Lyon H, Tantisira K, Silverman EK, Weiss ST, Laird NM, *et al.* (2003). Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. *Hum Hered* **55**(1)**:** 56-65.

Lander ES (1996). The new genomics: global views of biology. *Science* **274**(5287)**:** 536-539.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**(6822)**:** 860-921.

Landers KA, Mortimer RH, Richard K (2013). Transthyretin and the human placenta. *Placenta* **34**(7)**:** 513-517.

Lazar V, Bidart JM, Caillou B, Mahe C, Lacroix L, Filetti S, *et al.* (1999). Expression of the Na+/I- symporter gene in human thyroid tumors: a comparison study with other thyroid-specific genes. *The Journal of clinical endocrinology and metabolism* **84**(9)**:** 3228-3234.

Leonard JL, Visser TJ, Leonard DM (2001). Characterization of the subunit structure of the catalytically active type I iodothyronine deiodinase. *J Biol Chem* **276**(4)**:** 2600-2607.

Li H, Hao X, Zhang W, Wei Q, Chen K (2008a). The hOGG1 Ser326Cys polymorphism and lung cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* **17**(7)**:** 1739-1745.

Li P, Li Q, Chen L, Sun Y, Zhang J (2008b). No association between Fas A/G polymorphism and therapeutic effects induced by methimazole treatment for Graves' disease in Northern Chinese. *Pharmazie* **63**(10)**:** 748-750.

Liang JA, Chen CP, Huang SJ, Ho TY, Hsiang CY, Ding HJ, *et al.* (2005). A novel loss-of-function deletion in sodium/iodide symporter gene in follicular thyroid adenoma. *Cancer letters* **230**(1)**:** 65-71.

Lifton RP (1996). Molecular genetics of human blood pressure variation. *Science* **272**(5262)**:** 676-680.

Lin JD, Hsueh C, Chao TC, Weng HF (2001). Expression of sodium iodide symporter in benign and malignant human thyroid tissues. *Endocrine pathology* **12**(1)**:** 15-21.

Liyanarachchi S, Wojcicka A, Li W, Czetwertynska M, Stachlewska E, Nagy R, *et al.* (2013). Cumulative risk impact of five genetic variants associated with papillary thyroid carcinoma. *Thyroid* **23**(12)**:** 1532-1540.

Lu C, Zhao L, Ying H, Willingham MC, Cheng SY (2010). Growth activation alone is not sufficient to cause metastatic thyroid cancer in a mouse model of follicular thyroid carcinoma. *Endocrinology* **151**(4)**:** 1929-1939.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, *et al.* (2005). MicroRNA expression profiles classify human cancers. *Nature* **435**(7043)**:** 834-838.

Lunceford JK, Liu N (2008). Statistical performance of cladistic strategies for haplotype grouping in pharmacogenetics. *Stat Med* **27**(28)**:** 5816-5833.

Magner JA (1990). Thyroid-stimulating hormone: biosynthesis, cell biology, and bioactivity. *Endocr Rev* **11**(2)**:** 354-385.

Maia AL, Dupuis J, Manning A, Liu C, Meigs JB, Cupples LA, *et al.* (2007). The type 2 deiodinase (DIO2) A/G polymorphism is not associated with glycemic traits: the Framingham Heart Study. *Thyroid* **17**(3)**:** 199-202.

Maia AL, Goemann IM, Meyer EL, Magaginin SW (2011a). Type 1 iodothyronine deiodinase in human physiology and disease. *J Endocrinol*.

Maia AL, Goemann IM, Meyer EL, Wajner SM (2011b). Deiodinases: the balance of thyroid hormone: type 1 iodothyronine deiodinase in human physiology and disease. *J Endocrinol* **209**(3)**:** 283-297.

Mäkinen HS, Shikano T, Cano JM, Merilä J (2008). Hitchhiking Mapping Reveals a Candidate Genomic Region for Natural Selection in Three-Spined Stickleback Chromosome VIII. *Genetics* **178**(1)**:** 453-465.

Mallick UK (2010). The revised American Thyroid Association management guidelines 2009 for patients with differentiated thyroid cancer: an evidence-based risk-adapted approach. *Clin Oncol (R Coll Radiol)* **22**(6)**:** 472-474.

Master A, Wojcicka A, Piekielko-Witkowska A, Boguslawska J, Poplawski P, Tanski Z, *et al.* (2010). Untranslated regions of thyroid hormone receptor beta 1 mRNA are impaired in human clear cell renal cell carcinoma. *Biochim Biophys Acta* **1802**(11)**:** 995-1005.

McCullagh P, Nedler JA (1989). *Generalized Linear Models*. Second edn. Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

McGraw J, Waller D (2012). Cytochrome P450 variations in different ethnic populations. *Expert Opin Drug Metab Toxicol* **8**(3)**:** 371-382.

Mentuccia D, Thomas MJ, Coppotelli G, Reinhart LJ, Mitchell BD, Shuldiner AR, *et al.* (2005). The Thr92Ala deiodinase type 2 (DIO2) variant is not associated with type 2 diabetes or indices of insulin resistance in the old order of Amish. *Thyroid* **15**(11)**:** 1223-1227.

Michalatos-Beloin S, Tishkoff SA, Bentley KL, Kidd KK, Ruano G (1996). Molecular haplotyping of genetic markers 10 kb apart by allele-specific long-range PCR. *Nucleic Acids Res* **24**(23)**:** 4841-4843.

Michaud P, Tellez R, Rodriguez J, Correa L, Risopatron V (1986). [Bioavailability of a thyroxine preparation and its use in primary hypothyroidism]. *Rev Med Chil* **114**(3)**:** 206-210.

Miot F, Dupuy C, Dumont J, Rosset B (2012). Thyroid hormone synthesis and secretion. *Thyroid Disease Manager*.

Monzani A, Moia S, Prodam F, Walker GE, Bellone S, Petri A, *et al.* (2012). A novel familial variation of the thyroid hormone receptor beta gene (I276N) associated with resistance to thyroid hormone. *Thyroid*.

Moses AC, Rosen HN, Moller DE, Tsuzaki S, Haddow JE, Lawlor J, *et al.* (1990). A point mutation in transthyretin increases affinity for thyroxine and produces euthyroid hyperthyroxinemia. *J Clin Invest* **86**(6)**:** 2025-2033.

Muthukrishnan J, Harikumar KV, Verma A, Modi K (2010). Central hypothyroidism. *Indian journal of pediatrics* **77**(1)**:** 94-96.

Namo Cury A, Longui CA, Kochi C, Calliari LE, Scalissi N, Salles JE, *et al.* (2008). Graves' disease in Brazilian children and adults: lack of genetic association with CTLA-4 +49A>G polymorphism. *Horm Res* **70**(1)**:** 36-41.

Nebert DW, Zhang G, Vesell ES (2008). From human genetics and genomics to pharmacogenetics and pharmacogenomics: past lessons, future directions. *Drug Metab Rev* **40**(2)**:** 187-224.

Neta G, Brenner AV, Sturgis EM, Pfeiffer RM, Hutchinson AA, Aschebrook-Kilfoy B, *et al.* (2011). Common genetic variants related to genomic integrity and risk of papillary thyroid cancer. *Carcinogenesis* **32**(8)**:** 1231-1237.

Neumann S, Schuchardt K, Reske A, Reske A, Emmrich P, Paschke R (2004). Lack of correlation for sodium iodide symporter mRNA and protein expression and analysis of sodium iodide symporter promoter methylation in benign cold thyroid nodules. *Thyroid* **14**(2)**:** 99-111.

Nikiforov YE (2011). Molecular diagnostics of thyroid tumors. *Archives of pathology & laboratory medicine* **135**(5)**:** 569-577.

Nikiforov YE, Steward DL, Robinson-Smith TM, Haugen BR, Klopper JP, Zhu Z, *et al.* (2009). Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. *J Clin Endocrinol Metab* **94**(6)**:** 2092-2098.

Nikiforova MN, Chiosea SI, Nikiforov YE (2009). MicroRNA expression profiles in thyroid tumors. *Endocr Pathol* **20**(2)**:** 85-91.

Nikiforova MN, Nikiforov YE (2008a). Molecular genetics of thyroid cancer: implications for diagnosis, treatment and prognosis. *Expert Rev Mol Diagn* **8**(1)**:** 83-95.

Nikiforova MN, Tseng GC, Steward D, Diorio D, Nikiforov YE (2008b). MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. *J Clin Endocrinol Metab* **93**(5)**:** 1600-1608.

Ogawa Y, Sugawara T, Seki H, Sakuma T (2006). Thyroid follicular carcinoma metastasized to the lung, skull, and brain 12 years after initial treatment for thyroid gland--case report. *Neurol Med Chir (Tokyo)* **46**(6)**:** 302-305.

Okamura S, Arakawa H, Tanaka T, Nakanishi H, Ng CC, Taya Y, *et al.* (2001). p53DINP1, a p53-inducible gene, regulates p53-dependent apoptosis. *Mol Cell* **8**(1)**:** 85-94.

Pacanowski MA, Hopley CW, Aquilante CL (2008). Interindividual variability in oral antidiabetic drug disposition and response: the role of drug transporter polymorphisms. *Expert Opin Drug Metab Toxicol* **4**(5)**:** 529-544.

Pacini F, Castagna MG, Brilli L, Pentheroudakis G (2012). Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **21 Suppl 5:** v214-219.

Pallante P, Visone R, Ferracin M, Ferraro A, Berlingieri MT, Troncone G, *et al.* (2006). MicroRNA deregulation in human thyroid papillary carcinomas. *Endocr Relat Cancer* **13**(2)**:** 497-508.

Panicker V, Cluett C, Shields B, Murray A, Parnell KS, Perry JR, *et al.* (2008). A common variation in deiodinase 1 gene DIO1 is associated with the relative levels of free thyroxine and triiodothyronine. *J Clin Endocrinol Metab* **93**(8)**:** 3075-3081.

Panicker V, Saravanan P, Vaidya B, Evans J, Hattersley AT, Frayling TM, *et al.* (2009a). Common variation in the DIO2 gene predicts baseline psychological well-being and response to combination thyroxine plus triiodothyronine therapy in hypothyroid patients. *The Journal of clinical endocrinology and metabolism* **94**(5)**:** 1623-1629.

Panicker V, Saravanan P, Vaidya B, Evans J, Hattersley AT, Frayling TM, *et al.* (2009b). Common variation in the DIO2 gene predicts baseline psychological well-being and response to combination thyroxine plus triiodothyronine therapy in hypothyroid patients. *J Clin Endocrinol Metab* **94**(5)**:** 1623-1629.

Paragliola RM, Concolino P, De Rosa A, Mello E, Zuppi C, Pontecorvi A, *et al.* (2013). The first case of association between postpartum thyroiditis and thyroid hormone resistance in an Italian patient showing a novel p.V283A THRB mutation. *Thyroid* **23**(4)**:** 506-510.

Partsch CJ, Riepe FG, Krone N, Sippell WG, Pohlenz J (2006). Initially elevated TSH and congenital central hypothyroidism due to a homozygous mutation of the TSH beta subunit gene: case report and review of the literature. *Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association* **114**(5)**:** 227-234.

Patel A, Jhiang S, Dogra S, Terrell R, Powers PA, Fenton C, *et al.* (2002). Differentiated thyroid carcinoma that express sodium-iodide symporter have a lower risk of recurrence for children and adolescents. *Pediatr Res* **52**(5)**:** 737-744.

Peeters RP, van den Beld AW, van Toor H, Uitterlinden AG, Janssen JA, Lamberts SW, *et al.* (2005). A polymorphism in type I deiodinase is associated with circulating free insulin-like growth factor I levels and body composition in humans. *J Clin Endocrinol Metab* **90**(1)**:** 256-263.

Peeters RP, van der Deure WM, van den Beld AW, van Toor H, Lamberts SW, Janssen JA, *et al.* (2007). The Asp727Glu polymorphism in the TSH receptor is associated with insulin resistance in healthy elderly men. *Clin Endocrinol (Oxf)* **66**(6)**:** 808-815.

Peeters RP, van der Deure WM, Visser TJ (2006). Genetic variation in thyroid hormone pathway genes; polymorphisms in the TSH receptor and the iodothyronine deiodinases. *European journal of endocrinology / European Federation of Endocrine Societies* **155**(5)**:** 655-662.

Peeters RP, van Toor H, Klootwijk W, de Rijke YB, Kuiper GG, Uitterlinden AG, *et al.* (2003a). Polymorphisms in thyroid hormone pathway genes are associated with plasma TSH and iodothyronine levels in healthy subjects. *J Clin Endocrinol Metab* **88**(6)**:** 2880-2888.

Peeters RP, van Toor H, Klootwijk W, de Rijke YB, Kuiper GG, Uitterlinden AG, *et al.* (2003b). Polymorphisms in thyroid hormone pathway genes are associated with plasma TSH and iodothyronine levels in healthy subjects. *The Journal of clinical endocrinology and metabolism* **88**(6)**:** 2880-2888.

Peeters RP, Wouters PJ, Kaptein E, van Toor H, Visser TJ, Van den Berghe G (2003c). Reduced activation and increased inactivation of thyroid hormone in tissues of critically ill patients. *J Clin Endocrinol Metab* **88**(7)**:** 3202-3211.

Peltsverger MY, Butler PW, Alberobello AT, Smith S, Guevara Y, Dubaz OM, *et al.* (2012). The -258A/G (SNP rs12885300) polymorphism of the human type 2 deiodinase gene is associated with a shift in the pattern of secretion of thyroid hormones following a TRH-induced acute rise in TSH. *Eur J Endocrinol* **166**(5)**:** 839-845.

Penna-Martinez M, Epp F, Kahles H, Ramos-Lopez E, Hinsch N, Hansmann ML, *et al.* (2014). FOXE1 Association with Differentiated Thyroid Cancer and Its Progression. *Thyroid*.

Piekielko-Witkowska A, Nauman A (2011). Iodothyronine deiodinases and cancer. *J Endocrinol Invest* **34**(9)**:** 716-728.

Pierce JG, Parsons TF (1981). Glycoprotein hormones: structure and function. *Annual review of biochemistry* **50:** 465-495.

Pitt-Rivers R, Stanbury JB, Rapp B (1955). Conversion of thyroxine to 3-5-3'-triiodothyronine in vivo. *J Clin Endocrinol Metab* **15**(5)**:** 616-620.

Pohlenz J, Medeiros-Neto G, Gross JL, Silveiro SP, Knobel M, Refetoff S (1997). Hypothyroidism in a Brazilian kindred due to iodide trapping defect caused by a homozygous mutation in the sodium/iodide symporter gene. *Biochem Biophys Res Commun* **240**(2)**:** 488-491.

Polimanti R, Piacentini S, Manfellotto D, Fuciarelli M (2012). Human genetic variation of CYP450 superfamily: analysis of functional diversity in worldwide populations. *Pharmacogenomics* **13**(16)**:** 1951-1960.

Qiu L, Wang Z, Shi X, Wang Z (2008). Associations between XPC polymorphisms and risk of cancers: A meta-analysis. *Eur J Cancer* **44**(15)**:** 2241-2253.

Rajput-Williams J, Knott TJ, Wallis SC, Sweetnam P, Yarnell J, Cox N, *et al.* (1988). Variation of apolipoprotein-B gene is associated with obesity, high blood cholesterol levels, and increased risk of coronary heart disease. *Lancet* **2**(8626-8627)**:** 1442-1446.

Rastelli F, Crispino S (2008). Factors predictive of response to hormone therapy in breast cancer. *Tumori* **94**(3)**:** 370-383.

Reinwein D, Durrer HA, Meinhold H (1981). Iodine, thyroxine (T4), triiodothyronine (T3), 3,3',5'-triiodothyronine (rT3), 3,3'-diiodothyronine (T2) in normal human thyroids. Effect of excessive iodine exposure. *Horm Metab Res* **13**(8)**:** 456-459.

Riesco-Eizaguirre G, Gutierrez-Martinez P, Garcia-Cabezas MA, Nistal M, Santisteban P (2006). The oncogene BRAF V600E is associated with a high risk of recurrence and less differentiated papillary thyroid carcinoma due to the impairment of Na+/I- targeting to the membrane. *Endocr Relat Cancer* **13**(1)**:** 257-269.

Rigoli L, Raimondo G, Di Benedetto A, Romano G, Porcellini A, Campo S, *et al.* (1995). Apolipoprotein AI-CIII-AIV genetic polymorphisms and coronary heart disease in type 2 diabetes mellitus. *Acta Diabetol* **32**(4)**:** 251-256.

Risch A, Plass C (2008). Lung cancer epigenetics and genetics. *Int J Cancer* **123**(1)**:** 1-7.

Risch N, Merikangas K (1996). The future of genetic studies of complex human diseases. *Science* **273**(5281)**:** 1516-1517.

Rivera M, Ricarte-Filho J, Knauf J, Shaha A, Tuttle M, Fagin JA, *et al.* (2010). Molecular genotyping of papillary thyroid carcinoma follicular variant according to its histological subtypes (encapsulated vs infiltrative) reveals distinct BRAF and RAS mutation patterns. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **23**(9)**:** 1191-1200.

Rivolta CM, Moya CM, Esperante SA, Gutnisky VJ, Varela V, Targovnik HM (2005). [The thyroid as a model for molecular mechanisms in genetic diseases]. *Medicina (B Aires)* **65**(3)**:** 257-267.

Robbins J (1981). Factors altering thyroid hormone metabolism. *Environ Health Perspect* **38:** 65-70.

Robbins J (2002). Transthyretin from discovery to now. *Clin Chem Lab Med* **40**(12)**:** 1183-1190.

Rocha AS, Marques R, Bento I, Soares R, Magalhaes J, de Castro IV, *et al.* (2007). Thyroid hormone receptor beta mutations in the 'hot-spot region' are rare events in thyroid carcinomas. *J Endocrinol* **192**(1)**:** 83-86.

Ruiz-Llorente S, Carrillo Santa de Pau E, Sastre-Perona A, Montero-Conde C, Gomez-Lopez G, Fagin JA, *et al.* (2012). Genome-wide analysis of Pax8 binding provides new insights into thyroid functions. *BMC Genomics* **13:** 147.

Russo D, Arturi F, Schlumberger M, Caillou B, Monier R, Filetti S, *et al.* (1995). Activating mutations of the TSH receptor in differentiated thyroid carcinomas. *Oncogene* **11**(9)**:** 1907-1911.

Russo D, Wong MG, Costante G, Chiefari E, Treseler PA, Arturi F, *et al.* (1999). A Val 677 activating mutation of the thyrotropin receptor in a Hurthle cell thyroid carcinoma associated with thyrotoxicosis. *Thyroid* **9**(1)**:** 13-17.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, *et al.* (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**(6822)**:** 928-933.

Samollow PB, Perez G, Kammerer CM, Finegold D, Zwartjes PW, Havill LM, *et al.* (2004). Genetic and environmental influences on thyroid hormone variation in Mexican Americans. *J Clin Endocrinol Metab* **89**(7)**:** 3276-3284.

SantaLucia J, Jr. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A* **95**(4)**:** 1460-1465.

Sato H, Sakai H (2006). A family showing resistance to thyroid hormone associated with chronic thyroiditis and its clinical features: A case report. *Endocrine journal* **53**(3)**:** 421-425.

Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA (2002). Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* **70**(2)**:** 425-434.

Schmitt AC, Cohen C, Siddiqui MT (2010). Paired box gene 8, HBME-1, and cytokeratin 19 expression in preoperative fine-needle aspiration of papillary thyroid carcinoma: diagnostic utility. *Cancer cytopathology* **118**(4)**:** 196-202.

Schneider KA, Kim Y (2013). Genetic Hitchhiking under Heterogeneous Spatial Selection Pressures. *PLoS One* **8**(4)**:** e61742.

Schwertheim S, Sheu SY, Worm K, Grabellus F, Schmid KW (2009). Analysis of deregulated miRNAs is helpful to distinguish poorly differentiated thyroid carcinoma from papillary thyroid carcinoma. *Horm Metab Res* **41**(6)**:** 475-481.

Scipioni A, Ferretti E, Soda G, Tosi E, Bruno R, Costante G, *et al.* (2007). hNIS protein in thyroid: the iodine supply influences its expression and localization. *Thyroid* **17**(7)**:** 613-618.

Shah SH (2007). Gene polymorphisms and susceptibility to coronary artery disease. *Pediatr Blood Cancer* **48**(7)**:** 738-741.

Shargorodsky M, Serov S, Gavish D, Leibovitz E, Harpaz D, Zimlichman R (2006). Long-term thyrotropin-suppressive therapy with levothyroxine impairs small and large artery elasticity and increases left ventricular mass in patients with thyroid carcinoma. *Thyroid : official journal of the American Thyroid Association* **16**(4)**:** 381-386.

Shen S, Berry W, Jaques S, Pillai S, Zhu J (2004). Differential expression of iodothyronine deiodinase type 2 in growth plates of chickens divergently selected for incidence of tibial dyschondroplasia. *Anim Genet* **35**(2)**:** 114-118.

Shupnik MA, Ridgway EC, Chin WW (1989). Molecular biology of thyrotropin. *Endocr Rev* **10**(4)**:** 459-475.

Siraj AK, Al-Rasheed M, Ibrahim M, Siddiqui K, Al-Dayel F, Al-Sanea O, *et al.* (2008a). RAD52 polymorphisms contribute to the development of papillary thyroid cancer susceptibility in Middle Eastern population. *Journal of endocrinological investigation* **31**(10)**:** 893-899.

Siraj AK, Ibrahim M, Al-Rasheed M, Abubaker J, Bu R, Siddiqui SU, *et al.* (2008b). Polymorphisms of selected xenobiotic genes contribute to the development of papillary thyroid cancer susceptibility in Middle Eastern population. *BMC medical genetics* **9:** 61.

Smanik PA, Fithian LJ, Jhiang SM (1994). Thyroid peroxidase expression and DNA polymorphisms in thyroid cancer. *Biochem Biophys Res Commun* **198**(3)**:** 948-954.

Smanik PA, Liu Q, Furminger TL, Ryu K, Xing S, Mazzaferri EL, *et al.* (1996). Cloning of the human sodium lodide symporter. *Biochem Biophys Res Commun* **226**(2)**:** 339-345.

Smithies O, Maeda N (1995). Gene targeting approaches to complex genetic diseases: atherosclerosis and essential hypertension. *Proc Natl Acad Sci U S A* **92**(12)**:** 5266-5272.

Sodre AK, Rubio IG, Galrao AL, Knobel M, Tomimori EK, Alves VA, *et al.* (2008). Association of low sodium-iodide symporter messenger ribonucleic acid expression in malignant thyroid nodules with increased intracellular protein staining. *J Clin Endocrinol Metab* **93**(10)**:** 4141-4145.

Sorensen HG, van der Deure WM, Hansen PS, Peeters RP, Breteler MM, Kyvik KO, *et al.* (2008). Identification and consequences of polymorphisms in the thyroid hormone receptor alpha and beta genes. *Thyroid* **18**(10)**:** 1087-1094.

Srichomthong C, Ittiwut R, Siriwan P, Suphapeetiporn K, Shotelersuk V (2013). FOXE1 mutations in Thai patients with oral clefts. *Genet Res (Camb)* **95**(5)**:** 133-137.

St Germain DL, Galton VA (1997). The deiodinase family of selenoproteins. *Thyroid* **7**(4)**:** 655-668.

Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, *et al.* (2001). Haplotype variation and linkage disequilibrium in 313 human genes. *Science* **293**(5529)**:** 489-493.

Streckfuss F, Hamann I, Schomburg L, Michaelis M, Sapin R, Klein MO, *et al.* (2005). Hepatic deiodinase activity is dispensable for the maintenance of normal circulating thyroid hormone levels in mice. *Biochem Biophys Res Commun* **337**(2)**:** 739-745.

Sunthornthepvarakui T, Gottschalk ME, Hayashi Y, Refetoff S (1995). Brief report: resistance to thyrotropin caused by mutations in the thyrotropin-receptor gene. *N Engl J Med* **332**(3)**:** 155-160.

Swierniak M, Wojcicka A, Czetwertynska M, Stachlewska E, Maciag M, Wiechno W, *et al.* (2013). In-depth characterization of the microRNA transcriptome in normal thyroid and papillary thyroid carcinoma. *J Clin Endocrinol Metab* **98**(8)**:** E1401-1409.

Sykiotis GP, Neumann S, Georgopoulos NA, Sgourou A, Papachatzopoulou A, Markou KB, *et al.* (2003). Functional significance of the thyrotropin receptor germline polymorphism D727E. *Biochem Biophys Res Commun* **301**(4)**:** 1051-1056.

Tan SH, Lee SC, Goh BC, Wong J (2008). Pharmacogenetics in breast cancer therapy. *Clin Cancer Res* **14**(24)**:** 8027-8041.

Teofoli F, Camilot M, Tato L (2007). Lack of association between thyrotropin receptor gene polymorphisms and subclinical hypothyroidism in children. *J Endocrinol Invest* **30**(2)**:** 163-166.

Tomasini R, Samir AA, Pebusque MJ, Calvo EL, Totaro S, Dagorn JC, *et al.* (2002). P53-dependent expression of the stress-induced protein (SIP). *Eur J Cell Biol* **81**(5)**:** 294-301.

Tonacchera M, Viacava P, Agretti P, de Marco G, Perri A, di Cosmo C, *et al.* (2002a). Benign nonfunctioning thyroid adenomas are characterized by a defective targeting to cell membrane or a reduced expression of the sodium iodide symporter protein. *J Clin Endocrinol Metab* **87**(1)**:** 352-357.

Tonacchera M, Viacava P, Agretti P, de Marco G, Perri A, di Cosmo C, *et al.* (2002b). Benign nonfunctioning thyroid adenomas are characterized by a defective targeting to cell membrane or a reduced expression of the sodium iodide symporter protein. *The Journal of clinical endocrinology and metabolism* **87**(1)**:** 352-357.

Tonacchera M, Viacava P, Fanelli G, Agretti P, De Marco G, De Servi M, *et al.* (2004). The sodium-iodide symporter protein is always present at a low expression and confined to the cell membrane in nonfunctioning nonadenomatous nodules of toxic nodular goitre. *Clin Endocrinol (Oxf)* **61**(1)**:** 40-45.

Tong Z, Li H, Goljer I, McConnell O, Chandrasekaran A (2007). In vitro glucuronidation of thyroxine and triiodothyronine by liver microsomes and recombinant human UDP-glucuronosyltransferases. *Drug Metab Dispos* **35**(12)**:** 2203-2210.

Torlontano M, Durante C, Torrente I, Crocetti U, Augello G, Ronga G, *et al.* (2008). Type 2 deiodinase polymorphism (threonine 92 alanine) predicts L-thyroxine dose to achieve target thyrotropin levels in thyroidectomized patients. *J Clin Endocrinol Metab* **93**(3)**:** 910-913.

Trouttet-Masson S, Selmi-Ruby S, Bernier-Valentin F, Porra V, Berger-Dutrieux N, Decaussin M, *et al.* (2004). Evidence for transcriptional and posttranscriptional alterations of the sodium/iodide symporter expression in hypofunctioning benign and malignant thyroid tumors. *Am J Pathol* **165**(1)**:** 25-34.

Tzvetkov MV, Brockmoller J, Roots I, Kirchheiner J (2008). Common genetic variations in human brain-specific tryptophan hydroxylase-2 and response to antidepressant treatment. *Pharmacogenet Genomics* **18**(6)**:** 495-506.

Ukkola O, Savolainen MJ, Salmela PI, von Dickhoff K, Kesaniemi YA (1994). DNA polymorphisms at the locus for human cholesteryl ester transfer protein (CETP) are associated with macro- and microangiopathy in non-insulin-dependent diabetes mellitus. *Clin Genet* **46**(3)**:** 217-227.

Utiger RD (1980). Decreased extrathyroidal triiodothyronine production in nonthyroidal illness: benefit or harm? *Am J Med* **69**(6)**:** 807-810.

van der Deure WM, Appelhof BC, Peeters RP, Wiersinga WM, Wekking EM, Huyser J, *et al.* (2008a). Polymorphisms in the brain-specific thyroid hormone transporter OATP1C1 are associated with fatigue and depression in hypothyroid patients. *Clin Endocrinol (Oxf)* **69**(5)**:** 804-811.

van der Deure WM, Hansen PS, Peeters RP, Kyvik KO, Friesema EC, Hegedus L, *et al.* (2008b). Thyroid hormone transport and metabolism by organic anion transporter 1C1 and consequences of genetic variation. *Endocrinology* **149**(10)**:** 5307-5314.

van der Deure WM, Hansen PS, Peeters RP, Uitterlinden AG, Fenger M, Kyvik KO, *et al.* (2009). The effect of genetic variation in the type 1 deiodinase gene on the interindividual variation in serum thyroid hormone levels: an investigation in healthy Danish twins. *Clin Endocrinol (Oxf)* **70**(6)**:** 954-960.

van Schaik RH (2005). Cancer treatment and pharmacogenetics of cytochrome P450 enzymes. *Invest New Drugs* **23**(6)**:** 513-522.

Vasko V, Hu S, Wu G, Xing JC, Larin A, Savchenko V, *et al.* (2005). High prevalence and possible de novo formation of BRAF mutation in metastasized papillary thyroid cancer in lymph nodes. *J Clin Endocrinol Metab* **90**(9)**:** 5265-5269.

Veiga MI, Asimus S, Ferreira PE, Martins JP, Cavaco I, Ribeiro V, *et al.* (2009). Pharmacogenomics of CYP2A6, CYP2B6, CYP2C19, CYP2D6, CYP3A4, CYP3A5 and MDR1 in Vietnam. *Eur J Clin Pharmacol* **65**(4)**:** 355-363.

Venkataraman GM, Yatin M, Marcinek R, Ain KB (1999a). Restoration of iodide uptake in dedifferentiated thyroid carcinoma: relationship to human Na+/I-symporter gene methylation status. *J Clin Endocrinol Metab* **84**(7)**:** 2449-2457.

Venkataraman GM, Yatin M, Marcinek R, Ain KB (1999b). Restoration of iodide uptake in dedifferentiated thyroid carcinoma: relationship to human Na+/I-symporter gene methylation status. *The Journal of clinical endocrinology and metabolism* **84**(7)**:** 2449-2457.

Villanueva R, Inzerillo AM, Tomer Y, Barbesino G, Meltzer M, Concepcion ES, *et al.* (2000). Limited genetic susceptibility to severe Graves' ophthalmopathy: no role for CTLA-4 but evidence for an environmental etiology. *Thyroid* **10**(9)**:** 791-798.

Visser TJ (1996). Pathways of thyroid hormone metabolism. *Acta Med Austriaca* **23**(1-2)**:** 10-16.

Visser TJ, Leonard JL, Kaplan MM, Larsen PR (1981). Different pathways of iodothyronine 5'-deiodination in rat cerebral cortex. *Biochem Biophys Res Commun* **101**(4)**:** 1297-1304.

Visser TJvdD-T, I.; Docter, R.,; Hennemann, G. (1975). Conversion of thyroxine into tri-iodothyronine by rat liver homogenate. *Biochem J* **150**(3)**:** 489-493.

Visser WE, Friesema EC, Jansen J, Visser TJ (2008). Thyroid hormone transport in and out of cells. *Trends Endocrinol Metab* **19**(2)**:** 50-56.

Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, *et al.* (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A* **103**(7)**:** 2257-2261.

Vormfelde SV, Brockmoller J (2007). On the value of haplotype-based genotype-phenotype analysis and on data transformation in pharmacogenetics and -genomics. *Nat Rev Genet* **8**(12).

Vuissoz JM, Deladoey J, Buyukgebiz A, Cemeroglu P, Gex G, Gallati S, *et al.* (2001). New autosomal recessive mutation of the TSH-beta subunit gene causing central isolated hypothyroidism. *The Journal of clinical endocrinology and metabolism* **86**(9)**:** 4468-4471.

Wang D, Stapleton HM (2010). Analysis of thyroid hormones in serum by liquid chromatography-tandem mass spectrometry. *Anal Bioanal Chem* **397**(5)**:** 1831-1839.

Wang S, Wuu J, Savas L, Patwardhan N, Khan A (1998). The role of cell cycle regulatory proteins, cyclin D1, cyclin E, and p27 in thyroid carcinogenesis. *Hum Pathol* **29**(11)**:** 1304-1309.

Wang Y, Hou P, Yu H, Wang W, Ji M, Zhao S, *et al.* (2007). High prevalence and mutual exclusivity of genetic alterations in the phosphatidylinositol-3-kinase/akt pathway in thyroid tumors. *J Clin Endocrinol Metab* **92**(6)**:** 2387-2390.

Wang YL, Feng SH, Guo SC, Wei WJ, Li DS, Wang Y, *et al.* (2013a). Confirmation of papillary thyroid cancer susceptibility loci identified by genome-wide association studies of chromosomes 14q13, 9q22, 2q35 and 8p12 in a Chinese population. *J Med Genet* **50**(10)**:** 689-695.

Wang Z, Zhang H, He L, Dong W, Li J, Shan Z, *et al.* (2013b). Association between the expression of four upregulated miRNAs and extrathyroidal invasion in papillary thyroid carcinoma. *Onco Targets Ther* **6:** 281-287.

Ward EM, Jemal A, Chen A (2010). Increasing incidence of thyroid cancer: is diagnostic scrutiny the sole explanation? *Future Oncol* **6**(2)**:** 185-188.

Watson JD, Witkowski J, Gilman M, M Z (1996). Elucidation of the genetic code. *Recombinant DNA*(Second Edition)**:** 10-25.

Welker MJ, Orlov D (2003). Thyroid nodules. *Am Fam Physician* **67**(3)**:** 559-566.

Witt RL, Ferris RL, Pribitkin EA, Sherman SI, Steward DL, Nikiforov YE (2013). Diagnosis and management of differentiated thyroid cancer using molecular biology. *Laryngoscope* **123**(4)**:** 1059-1064.

Wojcicka A, de la Chapelle A, Jazdzewski K (2013). MicroRNA-related sequence variations in human cancers. *Hum Genet*.

Wood AJ, Zhou HH (1991). Ethnic differences in drug disposition and responsiveness. *Clin Pharmacokinet* **20**(5)**:** 350-373.

Xing M, Haugen BR, Schlumberger M (2013). Progress in molecular-based management of differentiated thyroid cancer. *Lancet* **381**(9871)**:** 1058-1069.

Xing M, Westra WH, Tufano RP, Cohen Y, Rosenbaum E, Rhoden KJ, *et al.* (2005). BRAF mutation predicts a poorer clinical prognosis for papillary thyroid cancer. *J Clin Endocrinol Metab* **90**(12)**:** 6373-6379.

Yeaman S (2013). Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc Natl Acad Sci U S A* **110(19)::** E1743-1751.

Yoder Graber AL, Ramirez J, Innocenti F, Ratain MJ (2007). UGT1A1*28 genotype affects the in-vitro glucuronidation of thyroxine in human livers. *Pharmacogenet Genomics* **17**(8)**:** 619-627.

Zarebczan B, Chen H (2010). Multi-targeted approach in the treatment of thyroid cancer. *Minerva Chir* **65**(1)**:** 59-69.

Zaridze DG (2008). Molecular epidemiology of cancer. *Biochemistry (Mosc)* **73**(5)**:** 532-542.

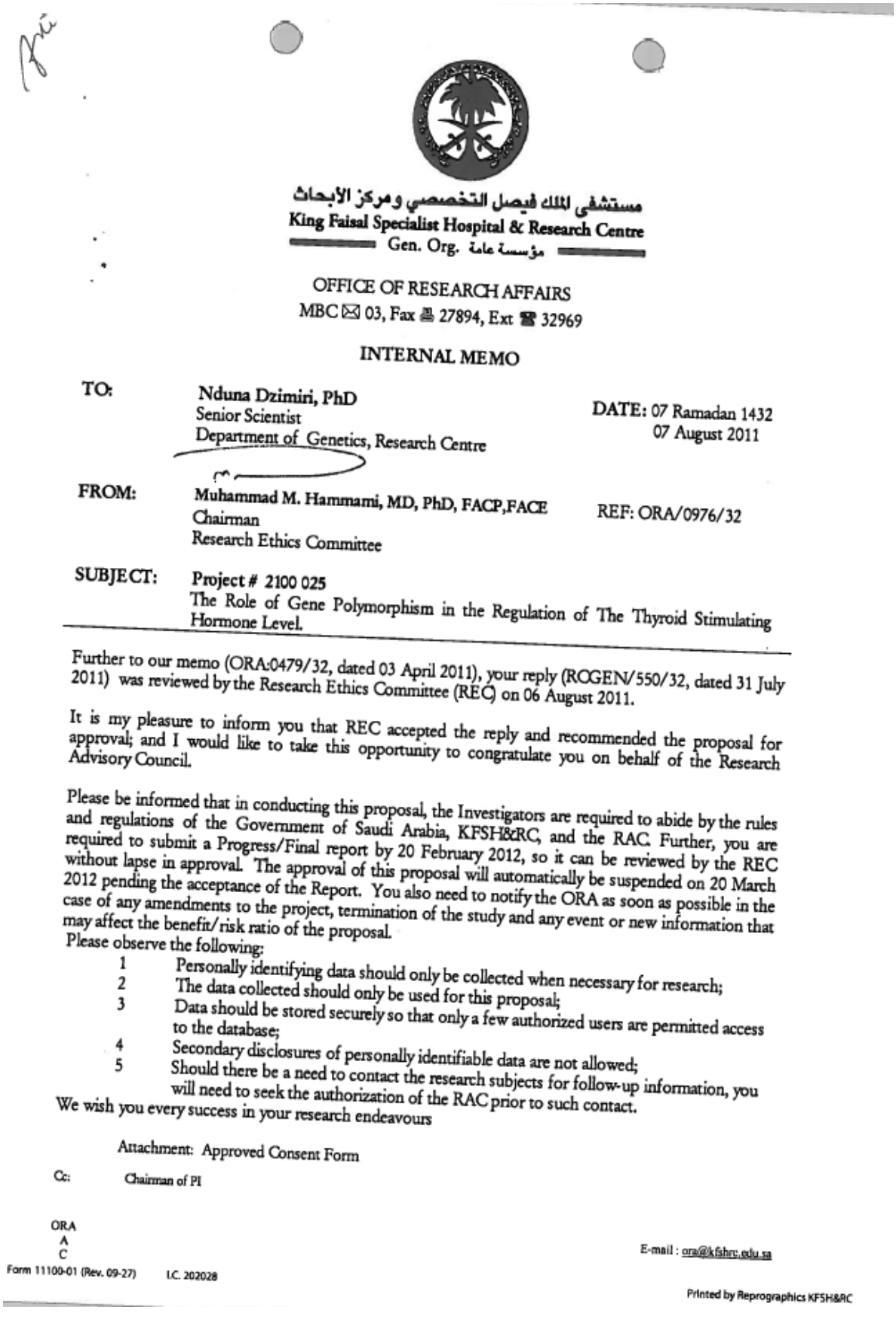
Zhou YL, Liu C, Dai XX, Zhang XH, Wang OC (2012). Overexpression of miR-221 is associated with aggressive clinicopathologic characteristics and the BRAF mutation in papillary thyroid carcinomas. *Med Oncol* **29**(5)**:** 3360-3366.

Zhu H, Xi Q, Liu L, Wang J, Gu M (2014). Quantitative Assessment of Common Genetic Variants on FOXE1 and Differentiated Thyroid Cancer Risk. *PLoS One* **9**(1)**:** e87332.

Zuo LJ, Guo T, Xia DY, Jia LH (2012). Allele and genotype frequencies of CYP3A4, CYP2C19, and CYP2D6 in Han, Uighur, Hui, and Mongolian Chinese populations. *Genet Test Mol Biomarkers* **16**(2)**:** 102-108.

Appendices

Appendix A: Study Approval letter



مستشفى الملك فيصل التخصصي و مركز الأبحاث
King Faisal Specialist Hospital & Research Centre
Gen. Org. مؤسسة عامة

OFFICE OF RESEARCH AFFAIRS
MBC ⊠ 03, Fax 🖷 27894, Ext ☎ 32969

INTERNAL MEMO

TO: **Nduna Dzimiri, PhD**
Senior Scientist
Department of Genetics, Research Centre

DATE: 07 Ramadan 1432
07 August 2011

FROM: **Muhammad M. Hammami, MD, PhD, FACP,FACE**
Chairman
Research Ethics Committee

REF: ORA/0976/32

SUBJECT: **Project # 2100 025**
The Role of Gene Polymorphism in the Regulation of The Thyroid Stimulating Hormone Level.

Further to our memo (ORA:0479/32, dated 03 April 2011), your reply (ROGEN/550/32, dated 31 July 2011) was reviewed by the Research Ethics Committee (REC) on 06 August 2011.

It is my pleasure to inform you that REC accepted the reply and recommended the proposal for approval; and I would like to take this opportunity to congratulate you on behalf of the Research Advisory Council.

Please be informed that in conducting this proposal, the Investigators are required to abide by the rules and regulations of the Government of Saudi Arabia, KFSH&RC, and the RAC. Further, you are required to submit a Progress/Final report by 20 February 2012, so it can be reviewed by the REC without lapse in approval. The approval of this proposal will automatically be suspended on 20 March 2012 pending the acceptance of the Report. You also need to notify the ORA as soon as possible in the case of any amendments to the project, termination of the study and any event or new information that may affect the benefit/risk ratio of the proposal.
Please observe the following:

1. Personally identifying data should only be collected when necessary for research;
2. The data collected should only be used for this proposal;
3. Data should be stored securely so that only a few authorized users are permitted access to the database;
4. Secondary disclosures of personally identifiable data are not allowed;
5. Should there be a need to contact the research subjects for follow-up information, you will need to seek the authorization of the RAC prior to such contact.

We wish you every success in your research endeavours

Attachment: Approved Consent Form

Cc: Chairman of PI

ORA
A
C

E-mail : ora@kfshrc.edu.sa

Form 11100-01 (Rev. 09-27) I.C. 202028

Printed by Reprographics KFSH&RC

Appendix B: Informed Consent

INFORMED WRITTEN CONSENT

موافقة متنورة

Part I: RESEARCH PARTICIPANT INFORMATION SHEET

الجزء : معلومات عن الدراسة للمتطوع في البحث

After receiving full explanation of the intended research project from Dr. **Nduna Dzimiri** and having all my inquiries about this study answered I, ---------- the undersigned, give my consent that I am (or my child is) participating in a research project.

أنا.....

أوقع على هذه الموافقة بَعد أن شَرح لي الدِكتورندونا دزميري أنني سِأشارك (أو أن طفلي سيشارك) في بَحث علمي وأجاب عن كُل تساؤلاتي بِخُصوص هَذا البَحث.

The project I am asked to participate in is entitled:

إن عنوان البَحث هو :

Association of Gene Polymorphisms in Thyroid function-related Genes with Differentiated Thyroid Cancer Risk and Post thyroidectomy L-Thyroxine Suppressive Dose Requirements

دراسة دور التحورات الجينية في الجينات المتعلقة بوظائف الغدة الدرقية بسرطان الغدة الدرقية و مقدار الجرعة الكافية من هرمون الثيروكسين

The purpose of this research is to give understanding of the patients' different responses to therapy so proper drugs and doses are tailored for individual patients.

إن الهدف من البَحث هو محاولة تحديد استجابات المرضى المختلفة لعلاجات الاضرابات النفسية لنتمكن من تحديد العلاجات والجرعات المناسبة لكل مريض بشكل انفرادي.

I am participating in this research because:

My (my child) participation requires that:

إن مشاركتي (مشاركة طفلي) تتطلُب:

1) Provide saliva sample (once).
2) The DNA from collected samples will be isolated and analyzed.

1) إعطاء عينة واحدة من اللعاب بالطرق الاعتيادية.
2) استخراج المادة الوراثية من العينة وتحليلها.

The risks and discomfort involved in participation are:

الخطر والأنزعاج المحتمل من المشاركة في هذه الدراسة هو:

There will be no pain or discomfort from collecting saliva samples. However, the genetic studies may reveal some information that some people may not like to know, such as a carrier state of a disease or having the disease without symptoms. People usually consider such information confidential and don't like that they leak out. Participating in research, if adequate precautions are not undertaken, may increase the chance of such a leakage.

لن يكون هناك اي ألم أو انزعاج ناتج عن تجميع عينة اللعاب. ولكن قد تكشف الدراسات الوراثية عن معلومات قد لايرغب بعض الناس بمعرفتها (مثل معرفة أنه حامل لمورثة المرض أو أنه مصاب بمرض لم تظهر أعراضه بعد).عادة ما يعتبر الناس هذه المعلومات خاصة وسرية ولايرغبون أن يعلمها الاخرون. إن المشاركة في البحث قد تزيد من تسرب هده المعلومات إذا لم تتخذ الاجراءات المناسبة.

Potential benefits:

الفوائد المرجوة:

This study may help indentify the genetic defect underlying the disease that my family suffers from. Such discovery may help my doctor in early diagnosis, deciding the proper drugs and dosage for the condition and in directing clinical care, and in family counseling. The study may also add to scientific knowledge if it identifies a new genetic abnormality.

قد تساعد هذه الدراسة على معرفة الخلل الوراثي المسبب للمرض الذي تشكو منه عائلتي. قد يساعد هذا الاكتشاف على الوصول للجرعات المناسبة للعلاج وعلى توجيه الطبيب لتقديم العناية الطبية المناسبة وعلى اعطاء النصائح العائلية. كما أن الدراسة قد تفيد في اغناء المعرفة العلمية إذا تم اكتشاف خلل وراثي جديد.

My (my child) participation is completely voluntary and my decision will not affect the medical care we receive. I can refuse participating without prejudice from my treating physician.

إن مشاركتي (مشاركة طفلي) هَذه **هي تطَوع مِني بِمَحض إرادتي** ولايؤثر قراري بالمشاركة أو عدم المشاركة على العناية الطبية المقدمة لنا. ان باستطاعتي رَفض المُشاركة دون أن يؤثر هذا على معاملتنا من قبل الطبيب المُعالج.

If I (my child) do not enroll in this study the available alternative is:

وفي حال رَفضي المشاركة يكون البَديل هو:

Not to perform the genetic analysis. I (my child) will then continue to receive the required clinical care.

أن لاتجرى الدراسات الوراثية. في هذا الحال سأستمر (سيستمر طفلي) بتلقي العناية الطبية اللازمة.

I can withdraw from this study whenever I wish, and I can request disposing of any information or samples taken from me without affecting the medical care I am entitled to receive.

كما أن **بإمكاني سَحب مشارَكتي (مشاركة طفلي)** من الدِراسة متى شئت و أن أطلُب إتلاف أي عينات أو مَعلومات مُتَعلِقَة بي دون أن يؤثر ذلك على تلقي الرّعاية الطِبّيَة اللازمة.

I will bear no extra cost as a result of my participation in this study. In case of any injury resulting from my participation, the King Faisal Specialist Hospital and Research Center will provide me with the necessary medical care.

I will not (my child will not) receive any **reimbursement** for participation in this study.

I understand that the information generated from my participation in this project will be **kept confidential** and no person or entity will have access to it besides those directly involved in the study. There will be no reference to my identity or my tribe in any published article about this study. I will be (not) informed about my results from this study. My treating physician will be (not) informed about my results from this study.

Contact persons:

I may call the Section of Assurance & Compliance, Office of Research Affairs telephone 4647272, ext. 32934 for general questions concerning research at KFSHRC or research subjects' rights. For specific questions on this study, or in the event of research-related adverse events, I may call DR Nduna Dzimiri, telephone 4647272, Ext. 27870, or telephone 4647272, pager 49319.

مُشاركتي في هذه الدراسة لَن تُحَملني أي نَفقات إضافية وفي حالة حدوث أي ضرر من مُشاركتي – لا قدر الله – فَسوفَ يُقَدم مستشفى الملك فيصل التخصصي ومركز الأبحاث العناية الطبية اللازمة.

لن اتلقى(لن يتلقى طفلي) أي **تعويض مالي** عن مشاركتي في هذه الدراسة.

إن المَعلومات النَاتِجَة عَن مُشاركتي في البَحث سوف تُعامَل **بِسرية تامَة** ولَن يَطَلِع عَليها أي شَخص أو جِهة عَدا المسؤولين عن إجراء الدِراسة إلا بموافقتي و لن تكون هناك إشارة إلى شخصي أو عشيرتي في أي بَحث مَنشور عَن هذه الدِراسة. سوف (لن) يخبرني الباحث بنتائجي المتعلقة بهذه الدراسة. سوف (لن) يخبر الباحث طبيبي بنتائجي المتعلقة بهذه الدراسة.

الأشخاص الذين يمكن الاتصال بهم:

يمكنني الاتصال بقسم رقابة وجودة الأبحاث التابع لمكتب شؤون الأبحاث على هاتف رقم 014647272 تحويلة32934 اذا كان لدي أي استفسارعن الأبحاث في المستشفى بشكل عام أو عن حقوق المشاركين بالبحث. كما يمكنني الاتصال بالدكتورندونا دزميري على هاتف رقم 4647272 تحويلة 27870 أو على هاتف رقم 4647272 جهاز نداء 49319 إذا كان لدي أسئلة خاصة بالبحث أو في حال حدوث اصابة أو انزعاج له علاقة بالمشاركة بالبحث.

PART II: CONSENT FOR STUDIES INVOLVING GENETIC MATERIAL (DNA/RNA)

I am asked to participate in a study that involves analysis of DNA/RNA (the genetic material in the cell), this type of study is subject to the rules of the Kingdom of Saudi Arabia in general and King Faisal Specialist Hospital and Research Centre in particular. There are certain options that I can choose from.

I have studied the following options carefully and checked the statements that I do accept.

1. I want ☐ / I Don't want ☐:

   To receive a general summary of the study results.

2. I want ☐ / I Don't want ☐:

   To receive the results of the study that belong to me

   (and/or my family).

3. I want ☐ / I Don't want ☐:

   To be asked to give consent before my (and/or my family's) leftover samples are used in other studies.

4. I want ☐ / I Don't want ☐:

   To be asked to give consent before my (and/or my family's) leftover samples are used in studies related to this study.

5. I want ☐ / I Don't want ☐:

   To remove all identifying information that links the sample to my (and/or my family's) identity once this study is completed.

6. I want ☐ / I Don't want ☐:

   My (and/or my family's) leftover samples to be destroyed once this study is completed.

7. I allow ☐ / I Don't allow☐:

   Other investigators outside KFSH&RC to access/use my (and/or my family's) samples.

8. I agree ☐ / I disagree☐:

   That my (and/or my family's) samples be sent outside the Kingdom for research purposes.

9. I agree ☐ / I disagree☐:

   That my (and/or my family's) samples be used for commercial purposes.

I have carefully studied the above-mentioned options and checked what I do agree with and I sign this consent with full understanding.

الجزء الثاني : الموافقة المتنورة على دراسة على المادة الوراثية

الَبحث الَذي طُلب مِني المشاركة فيه يتضَمَن تحليلاً للجينات وهي المواد الوراثية في الخلايا. هذا النَوع من الدِراسة يَخضع لِلنُظم المَرعية في المملكة العربية السعودية بصفة عامه وفي مستشفي الملك فيصل التخصصي ومَركز الأبحاث بِصفة خاصة. هُنالك خيارات مُعَينَة يمكن للمشارك أن يختار منها.

إنني قد درَست الخيارات التالية بعناية واخترت منها ما أوافق عَليه.

1. أريد ☐ / لا أريد ☐ :

   أن أعطى مُلخصاً عن نتائج الدِراسة بشكل عام.

2. أريد ☐ / لا أريد ☐ :

   أن أعطى النتائج الخاصة بي (و/ أو بأسرتي) المتعلقة بهذه الدراسة.

3. أريد ☐ / لا أريد ☐ :

   أن تؤخَذ موافقتي قبل استخدام المتبَقي من عيناتي (و/ أو عينات أسرتي) في دِراسات أخرى.

4. أريد ☐ / لا أريد ☐ :

   أن تؤخَذ موافقتي قبل استخدام المتبَقي من عيناتي (و/ أو عينات أسرتي) في أبحاث ذات صِلm بهذا البحث.

5. أريد ☐ / لا أريد ☐ :

   أن يُزال كُل مُؤشر يَربِط العَينات بِشخصي (و/ أو بأسرتي) عند استكمال هذا البَحث.

6. أريد ☐ / لا أريد ☐ :

   إتلاف المتبَقي من عيناتي (و/ أو عينات أسرتي) عند استكمال هذا البَحث.

7. أوافق ☐ / لا أوافق ☐ :

   على استخدام عيناتي (و/ أو عينات أسرتي) مِن قِبَل باحثين من خارج مستشفى الملك فيصل التخصصي ومركز الأبحاث .

8. أوافق ☐ / لا أوافق ☐ :

   على إرسال عيناتي (و/ أو عينات أسرتي) خارج المملكة من أجل البَحث العلمي.

9. أوافق ☐ / لا أوافق ☐ :

   على استخدام عيناتي (و/ أو عينات أسرتي) لأغراض تِجارية.

لقد درست الخيارات المذكورة أعلاه بِتَمَعن واخترت مِنها ما رأيته مناسِباً وعلى هذا أوقع.

Research Subject or Surrogate:

Print Name_____

Signature: _____Date: _____

Relationship: _____
(if signed by person other than the research subject)

المشارك بالبحث أو ولي الأمر:

الاسم : _____

التوقيع :_____ التاريخ :_____

صلة القرابة _____

(إذا كان الموقع غير المريض المشارك)

WITNESS

I confirm that I have accurately translated and/ or read the information to the subject:

Print name: _____

KFSH&RC ID#: _____

Signature: _____ Date: _____

الشاهد

أقر بأنني قد قرأت / أو ترجمت للمشارك هذه المعلومات بشكل صحيح.

الاسم : _____

رقم بطاقة المستشفى: _____

التوقيع : _____ التاريخ : _____

Investigator or Delegate

I have fully explained to the above volunteer/ relative/ surrogate the nature and purpose of the above-mentioned research project.

Print Name: _____

KFSH&RC ID#: _____

Signature: _____Date: _____

الباحث أو ممثله

أقر بأنني قد شرحت للمتطوع/ لقريبه/ أو ولي أمره المذكور أعلاه بصورة كاملة طبيعة وأهداف الدراسة المذكورة.

الاسم : _____

رقم بطاقة المستشفى : _____

التوقيع :_____التاريخ: _____

Appendix C: Patient compliance questionnaire

1. Medical Record Number ☐
2. Gender

   Male ☐ Female ☐
3. Age ☐
4. Height _____
5. Weight _____
6. What is the dose of L-thyroxine you are taking?

_____

7. How long have you been on this dose?

_____

8. Do you take it every day? Yes ☐ No ☐

   If not, how many days per week on average do you miss taking it?

_____

9. Do you take it on empty stomach first thing in the morning?

   Yes ☐ No ☐
10. How long do you normally wait after taking thyroxine before you take anything per mouth? _____
11. Do you suffer from other illnesses? Yes ☐ No ☐

    If yes, list them: 1. ---------- 2. ---------- 3. ----------

    4. ---------- 5. ---------- 6. ----------
12. Do you take other medications? Yes No

    If yes, list them, 1. ---------- 2. ---------- 3. ----------

4. ---------- 5. ---------- 6. ----------

13 Do you have family history of thyroid diseases? Yes ☐ No ☐

If yes, who else have it? _____

What type of disease? _____

Appendix D: PCR Primers for gene amplification

All primers were tagged with M13 universal forward and reverse oligo sequences: Univ-M13-Fw 5'-GTA AAA CGA CGG CCA GT-3', Univ-M13-Rv 5'-CAG GAA ACA GCT ATG ACC-3'

Position	Strand	Primer Sequence	Annealing temp (^{o}C)
Promoter	F	AAAACCATGGATGTTTGTTGG	62.7
	R	CAGAGCCTCTTCAGCCACA	
Exon1A	F	TGTGAGAGAGCATCTAACAGGTTC	63.0
	R	GGATGTTCCGCTTGACTCTG	
Exon1B	F	GAGGTGGCTGTGCATGTG	61.0
	R	ATTTGAGTGCCTGGGAAGGT	
Exon2	F	CTCCAGCCCCTGCTAAGTC	61.0
	R	GTGAGGTTAGCCCAGTGCAG	
Exon3	F	AATTCCTGACGTGACCTTGC	61.0
	R	GTAAGTGGCTGGGTGGTGAT	
Exon4A	F	TGCTGATGGCTCCTACACAG	60.0
	R	GTCTCCAAAGTGGGAGCAGT	
Exon4B	F	TGGCTTTTACCCTTGACCTG	60.0
	R	TCCCTCCTCTTCCAACTCCT	
Exon4C	F	ATTCAAGCTGGCATTCCTTC	55.0
	R	GACTTGCGAAGAAATCAGAGAAA	
Exon4D	F	CCCTGAAATCTTCCACTAGCC	61.0
	R	CTGAATGTTCTTCCCCCAAA	
Exon4E	F	CACACGGCTGTGACTTGATT	60.0
	R	ATCTTGGCTCACTGCAACCT	

Table D1: DIO1 primers The table displays the forward (F) and reverse (R) stand primers used to identify SNPs in the DIO1 gene. The gene was divided into 10 segments.

Locus	Strand	Primer Sequence	Annealing temp (^{o}C)
Promoter	F	AGGGACGAGTACGCCAGCTT	62.7
	R	GCAAGTCTACGCTGAGGATGCCC	
Exon1	F	CCCTTCTCCCCCAACCCCCA	61.5
	R	ATGGCCTCTGGTCCCCAGCA	
Exon2	F	GGGTGCATAACCCAGCAAACCA	60.0
	R	AGGAGCCAGGGTCCCTTACTGA	
Exon3	F	GTTCATTGCAGGTCAGGGTT	61.0
	R	TCCGAGCAGACCTGTTGATCCT	
Exon4A	F	ACCAAATGCTCTGAAGGAAATACCACA	58.9
	R	GAGTCCCCCGGTATCGCCCA	
Exon4B	F	GAGCCAGCTGCCGCCTTCC	58.9
	R	TCTCCAGCCAATGCCGGACT	
Exon4C	F	GCCCCAGTGCCGAGTTGTGG	53.5
	R	TGGTTACTTACTCAGCCCAATGCCA	
Exon4D	F	ACAGAGCCCTATTGGCTTACTGA	55.5
	R	ACCACATGGCCTGGGTTCAAGGA	
Exon4E	F	TGCCTTGGCTCTATTTGGCATGGA	55.0
	R	TGGCTAACAAATGTGACCCACTCT	
Exon4F	F	TCTAGGCAGGTGGAGAGAGTGGT	58.9 (no Q solution/ increased Mgcl2)
	R	AGACCAGCACCACCAGACCCA	
Exon4G	F	TTGTGAGGGGTTTCTTTTGGAA	52.0
	R	TCCATGATATACCGGGGGTTGCCT	
Exon4H	F	CCAAAGTAGGTGAGCTATAGTGAAGA	52.5
	R	TGGTTGGTCCATTCTTTTGGGAGC	
Exon4I	F	TGGTCCCATTAGCTGAGGTTTAGT	60
	R	CCAGCTGGTGACACGACTGAT	
Exon4J	F	AAGGATCTTTTGTCCATTCCATT	60.0
	R	TGGGGGCAGAGATAAGCCTTGA	
Exon4K	F	TCATGGCAACATCAGCTTTCT	58.9
	R	TCCCCAGTTGACTAGCACTGCCT	
Exon4L	F	TGGTGTTGAGATTTACGTGGAACCC	59.0
	R	GCGTAGGATTCCTGAGCTTTGTCCA	
Exon4M	F	TCTCCCTGAGGTTAGCCCAA	62.5
	R	ACACATTTCCCTGTGGGACTAAGA	
Exon4N	F	AATCAATGCTTTATCTGATATGCTGAG	55.0
	R	TCCATTCATGCCCATTCAAGAAA	
Exon4P	F	ACATCTGGGTCTACTCAAACACT	58.9
	R	AAGCACTTCCTGGTTAACGG	
Exon4Q	F	ACTGACCTCCTTAGAGGCAGATT	58.9
	R	AGGCCATAGGGCATGCTGAGGA	
Exon4R	F	TGTTGGTGCTGCCTTGCTGC	58.9
	R	ACTCTGACTTTTGGCCATACCA	
Exon4S	F	CATCCCCACTGTCATGTGTCTCAGC	60.0
	R	TTGCAGCAGCCTTCAGCCTTCC	
Exon4T	F	TGCTAACTGGGAAGAACAAGGCCC	61.0
	R	GGAGGGGGAGAAGTGGAGGGT	
Exon4U	F	GCAAGTCGGGGCTGGCTACC	62.7
	R	TCTGTCTCCCTCTTCGGGCG	
Exon4V	F	GCCCATGTCACTGGTCAGCGT	60.0
	R	ACGCGTGCATTCTCACCACA	

Table D2: DIO2 Primers. The table displays the forward (F) and reverse (R) stand primers used to identify SNPs in the DIO2 gene. The gene was divided into 25 segments.

Locus	Strand	Primer Sequence	Annealing temp (^{o}C)
Promoter	F	GAGCCGGGGCCACCGAAG	63.0
	R	AATCTTGGGCTGAGGGCGGG	
Exon1A	F	CCAGGTCCCTGGTCGGGCTC	60.0
	R	GCAGGGAGCGGAGCATGGTG	
Exon1B	F	CCCTGCACTGCTGAAGCCCA	58.9
	R	CCCTCGTGCGCCTGCTTGAA	
Exon1C	F	CCGATGACCCGCCCATCTGC	62.7
	R	AGGCTCCGGTGCTGTGGGAT	
Exon1D	F	GCCTTCCAGCGCCTGGTCAC	62.7
	R	AGCGCCGTGCAGTTGCTCAT	
Exon1E	F	CCCCGACGGCTACCAGGTCT	61.0
	R	TCCTGGGGGCACACGGACAA	
Exon1F	F	ACACCTGCCTGGCTCACTGGAA	58.9
	R	GGGGCTCCTGTGCACCCAAA	
Exon1G	F	TTGCCTGGCACCCACCTGTC	61.0
	R	GGGACCTCAGGCTCTCCCCC	
Exon1H	F	AGGCATTGGCGAGGTTCGCA	61.0
	R	CATCCTCCTCCAGCGCCCCT	

Table D3: DIO3 Primers. The table gives the forward (F) and reverse (R) stand primers used to identify SNPs in the DIO3 gene. The gene was divided into 9 segments.

Locus	Strand	Primer Sequence	Annealing temp (^{o}C)
Promoter	F	CAGGTCAGCTTGACATCCTGT	57.0 (w/o Q solution)
	R	TGCTTTCTTATCTGAAAAGCATC	
Exon1	F	AGAGAGGAAAATGCATGCTTTAATA	60.0 (w/o Q solution)
	R	GATATTACAGCCTGTTGAAGCAAA	
Exon2	F	GGATCAGGGGGTTCCTAGAT	57.0
	R	GAATGGACTTCTTCAGGGCTTA	
Exon3A	F	TTGGCTCCTTAGAAGCAGAGT	57.0
	R	GCACTTGCCACACTTACAGC	
Exon3B	F	TTCTGTTCTTTCCCCAGGA	55.0
	R	CTTCAACTGAGCCCAAAAGG	

Table D4: TSHβ Primers. The table displays the forward (F) and reverse (R) stand primers used to identify SNP of the TSHβ gene. The gene was divided into 5 segments.

Locus	Strand	Primer Sequence	Annealing temp (^{o}C)
Prm, Exon1	F	TATTTCCAGGCCTGCTGAGT	57.0
	R	CGCTGAAGAAGCTCCAGACT	
Exon2	F	CTCCTACTCCTGGCAGACGA	57.0
	R	GGAAACCCCCTCTAACTCCA	
Exon3	F	ACCATTTGTGCCAGCATCTT	57.0
	R	CAGGACCAAAGCTGGACATT	
Exon4	F	GTTTGAGACCAGCCTTGGAG	57.0
	R	CTGCCTGATTGTTCAGCATC	
Exon5	F	AAATGCGGACAACTGAGGAG	55.0
	R	CAAGTGTTTGCACCTCTCCA	
Exon6	F	CCCTCTTCCCACCACCGCAG	59.0
	R	CACGTGGGCCCAGGAGCCTT	
Exon7A	F	TGTCCCTCAGTTGGCCAAGGCT	59.0
	R	AAGGCAGGTGCCCTGAGCCC	
Exon7B	F	CCCCCAGCCACACCAAAGGC	61.0
	R	AGCGAGGGTGGTGAGCTCGG	
Exon7C	F	TGCCATCAGCGTCTTTCCCAGA	61.0
	R	GGGGCACAATGGTAGGTGCTGG	
Exon8A	F	TGCTCCTGGGAAGGGGAGGC	61.0
	R	TGTCTGCTCTCAACCACTGCCCT	
Exon8B	F	AGGGCAGTGGTTGAGAGCAGACA	61.0
	R	TGGCAGCCCTCATCTCCCCA	
Exon8C	F	TCCAGGGATGCGGGAGGACA	64.0
	R	ACCCCTCGGCCCACCTTGAG	
Exon9	F	TTCGCTGCAGAAGCGCAGCA	56.0
	R	CTGGCGGTCTGCCCTGAGGA	
Exon10A	F	TGAGGAGTGGGGGCAGGAGG	61.0
	R	AGGCCGAGCTGGGGCAAGTT	
Exon10B	F	CCTGGAGACGATGTCCCAGTGC	59.0
	R	TGACTGGCAGGCAGGTTCCAGA	
Exon10C	F	TCTCTCAAGCCTGCATGGATAGGAC	62.0
	R	GCCCAGGAGATCCCAGCCCA	
Exon10D	F	CTGTGACCCCCATGCCAGCC	59.0
	R	GCAGGGGAAGGTGGAGGGGT	
Exon10E	F	CCAGTCTCCAGTCGGTTCCCCC	62.0
	R	ACTGGGAGATTGCAAGCAGTGG	
Exon11A	F	CCATACCCCTGCTGTGGGGC	60.0
	R	AGGGCTGGGGTGTGGGAAGG	
Exon11B	F	ACGGAAAAGGACACTGCCCCA	59.0
	R	ACACTGCCCTCCTGCTGCCT	
Exon12A	F	AAGGTGAGGCCAGAGAAGGAGG	59.0
	R	TCCCTGGGACTCCTGCCCCT	
Exon12B	F	GCGGCACACTCCCTGTGTGG	59.0
	R	GGGACTCGCTCCAGCCCTCA	
Exon12C	F	CCCACACGGAGGCTCTGGCT	61.0
	R	TCAGGCCTTGGGGGCAGGAG	
Exon12D	F	CTGCAGCTCACGCCCACCAG	61.0
	R	TGGGATGAGTGAGGTACCCAGCG	
Exon12E	F	CCGTGGGCTGGAAGGGACCA	61.0
	R	TCCCTGATGAGGACTCTGGGGT	

Locus	Strand	Primer Sequence	Annealing temp (oC)
Exon12F	F	GCCCTGTTTTACATACAGCTTCATTTT	61.0
	R	CCCTGCCCCCTTGTTCCAGC	
Exon12G	F	AGCTCGGGCTCTGCACTCCA	63.0
	R	TGGGTAGGGAGCCCTCGGGA	
Exon12H	F	AGGGGGAACCCTCCTGGAAGA	62.5
	R	CCTGGCAAGCCCAGCTCACC	

Table D5: PAX8 Primers. The table displays the forward (F) and reverse (R) stand primers used to identify SNP of the PAX8 gene. The gene was divided into 28 segments. Prm, promoter

Locus	Strand	Primer Sequence	Annealing temp (^{o}C)
Promoter	F	AGTGAACCCCAGTCCTGCAT	62.0
	R	CCACCTCCCCTCTCACTCT	
Exon1A	F	CTAGGTCTGGAGGCGGAGTC	62.7
	R	TAGTCCCAGGCTCCGAAGG	
Exon1B	F	TCGACAGCCCATAGATTCCT	58.9
	R	AGGACCGAGTTCAGAAGCTG	
Exon1C	F	CTGTCTGCCAGCTTCATGTC	58.9
	R	CCAGGCTCTCTCTTCCTCCT	
Exon2	F	ACCAGTCCTCCAGGTCCAC	62.7
	R	CAGAAGGGTGGTCCCAGAAT	
Exon3	F	CCATGCTCCCGTTCTGAC	58.9
	R	CCAGAGTCTCCCCACAAGAG	
Exon4	F	TCATCCTGAACCAAGGTGTG	58.9
	R	CCCAGGATGAAGCAGGAG	
Exon5	F	TGGGCAAAAGACAGAAACCT	55.5
	R	CCCAGGCTGGTCTTGAACT	
Exon6	F	TCCCAAAGTGCTGGGATTAC	62.7
	R	GAGAGGGTCGCAGTCAGTGT	
Exon7	F	GCTTGCCGCACAGAGAAG	58.9
	R	CTCCTGACCTCAGGTGATCC	
Exon8	F	CTGCAGTGAGCCACATTCAT	61.0
	R	CTCTCTGCCAGCGCCTAC	
Exon9	F	AGGGATGTCTCAGGTCTTCG	58.9
	R	ATGGGAAAGAGGGAAAGGTG	
Exon10	F	AGATGGTGTGGACGGTCTCT	62.0
	R	CCTGGGACAACACAGTGAGA	
Exon11	F	CCACCCCCAAGAGATAATCA	62.5
	R	AGGACCCTCATGGTCTGCT	
Exon12	F	CTACTCTCCCAGGGCTCCTT	60.0
	R	CCCCAGGAATGCATTTTACT	
Exon13	F	CACAGATCTGGGCAGACAGTAA	61.0 (w/o Q-solution)
	R	AGCCTGGGTGACAGAGTGAG	
Exon14	F	TGAGGCTCAGAGGCTTGC	57.0
	R	TTCAGTCCTTGTGAAGGGAGA	
Exon15A	F	TGACATGTGAGTGGGGAAGA	62.0
	R	TTTGCATACAAGGCAATCCA	
Exon15B	F	TGTTGGACATGATGGTGGTC	62.0
	R	GGCACCTACAATCCCAGCTA	
Exon15C	F	AGTGCAGTGGTGTGATCTCG	61.5
	R	ATGTGGGGTCTGAATCATCC	
Exon15D	F	GGCTGGTCTCAAACTCCTGA	60.0
	R	CTGGCAATCCTTGGGTTCT	
Exon15E	F	CCAGAACCTCAGAACATGAGC	58.5
	R	GGCATCAAACAACAGAAATTTATTC	

Table D6: NIS Primers. The table displays the forward (F) and reverse (R) stand primers used to identify SNP of the NIS gene. The gene was divided into 22 segments.

Appendix E: Probes for the real-time PCR assays

Gene	Assay ID	SNP ID	Chromosom-al locus (NCBI Build 37)	Context Sequence [VIC/FAM]
TSHβ	C_8690289_10	rs1321108	1:115572365	AGAAAGCAGT[A/G]AATCAAATGCA
	C_8690288_20	rs1321109	1:115571916	AGATGTGGCC[A/G]TGTTCCTCAGAG
	C_2616411_10	rs28566771	1:115576482	CAATTCTTTC[C/G]CAGTTGTATTTGT
	C_27540229_10	rs10776792	1:115576023	TTTTTGGCCTT[A/G]CATGTGGGCAA
DIO1	C_485944_10	rs12033572	1:54369674	AGCTGCCTTC[C/G]GTGCCTCCCTAC
	C_31601232_10	rs12084242	1:54372124	CTTCTTGGAG[A/T]TCCTTCAATTTCC
	C_16187176_10	rs2294510	1:54359922	TGAAGAGGCT[C/T]TGGGTGCTCTTG
	C_16187186_10	rs2294511	1:54370264	AGAATTGAAC[A/T]TAGGAGTGAAA
	C_31601225_10	rs12095080	1:54376730	CAGTAAACTA[A/G]TAACAGAATTA
	C_334342_20	rs11206244	1:54375701	TTCTAGGTGA[C/T]CAACGGGAGGG
	C_485945_10	rs4926616	1:54369730	CAACAGCTTA[C/T]ATTTATGAAGCA
DIO2	C_26929412_10	rs225010	14:80672279	ATTTGGGTGA[C/T]GGATAAGCTCA
	C_26929411_10	rs225011	14:80672208	GTTGTTATAAT[C/T]CTAGGTCAGTG
	C_15819951_10	rs225014	14:80669580	ACCTCCTTCTG[C/T]ACTGGAGACAT
	C_568127_10	rs225015	14:80667579	ATGGGACCAA[A/G]GGAAGATCCAA
	C_26929404_10	rs225017	14:80667226	CTCTTTTCTCA[A/T]TTCAGACAACT
	C_31755133_10	rs11159446	14:80601392	CTCATGATCTA[C/G]TCATCCACAAA
	C_29283499_10	rs7140952	14:80668272	TGACCCACTCT[C/T]TACAGCTTAAG
DIO3	C_7565113_10	rs945006	14 : 102029277	AGTGATGTCG[G/T]GGATGGGGAGG
NIS	C_25473586_10	rs4808708	19:18001686	ATTTCTCCCC[A/G]CCTCTCAGGCCC
	C_25473598_10	rs4808709	19:18001839	TCCCAGATCT[A/G]GAGGCAGCCAA
	C_478522_10	rs7250346	19:18004362	GGAGGCAGAG[C/G]GGGCAGGACTG
	C_86391723_10	rs45602038	19:17999239	GCTGACCACTGTGCTGTG[C/T]GGAGC CCTCATCAGCTGC
	C_32102077_20	rs12327843	19:18004912	TATTAGACGCTGCAGCCC[C/T]GACGG CTCCCCCCAAATA
PAX8	C_194000220	rs2241975	2:113984503	GGTAATGTAG[C/T]AGGCCTGGGAA
	C_1940001_10	rs11123172	2:113984303	ACAGAACGGG[C/T]AAACTGGGATT
	C_1151707_10	rs67776659	2:113999736	AGAGGGTCAG[G/T]GGTGGGAGTGA
	C_25607046_10	rs3738913	2:113984854	GGCACGTTTC[A/G]TTCATGCTCAT
	C_1940022_10	rs3748915	2:113983937	AATGGCTCAA[C/T]TATGTCTCATCC
	C_1939999_1	rs3748916	2:113984033	TCCAGACCTT[A/G]TCATCTTGAATC
	C_1939984_10	rs1049137	2:113975110	CATTTACAGT[A/G]TATACAGTCAGG
	C_1939985_10	rs4849176	2:113977936	TGGAACCCAT[C/T]GATCTGGAAAA
	C_1940020_10	rs4849186	2:113999534	ATCAAAGCCT[C/G]AGCAAACTGCT
	C_1940022_10	rs13007173	2:114001921	GCTCCACCCA[A/G]GCCAGGCCTTT
	C_8737927_10	rs1478	2:113975066	AGGCCAAAGTCTGGGGG[G/T]TAGAA AGAAGGAAGCTTGA
	C_8737926_10	rs1479	2:113975104	ACCCAAACAAAGTTTCATT[G/T]ACAG TATATACAGTCAGGC
	C_8737928_10	rs874898	2:113974196	AGGTGGTTGGCCTAAAGT[C/G]CTAA GTCTCTACAGCCTC

The table list the probes used for the real-time PCR for selected single nucleotide polymorphisms in the 6 genes

Appendix F: Description of the 39 studied SNPs in the association studies

Gene	SNP ID	Probe ID	Nt location	Chromosomal location	Chan-ge	A/Acid Change
TSHβ	rs1321108 (A>G)	C_8690289_10	nt-50	1:115572365	A/G	Promoter region
	rs1321109 A>G	C_8690288_20	nt-499	1:115571916	A/G	Promoter region
	rs28566771 C>G	C_61819701_10	nt4068	1:115576482	C/G	Intronic
	rs10776792 G>A	C_27540229_10	nt3,609	1:115576023	A/G	ACA>GCA (Threonine >Alanine)
NIS	rs4808708 G>A	C_25473586_10	nt18,905	19:18001686	A/G	Intronic
	rs4808709 A>G	C_25473598_10	nt19,058	19:18001839	A/G	Intronic
	rs7250346 C>G	C_478522_10	nt21,581	19:18004362	A/G	Intronic
	rs45602038 C>T	C_86391723_10	nt16,458	19:17999239	C>T	TGC>TGT (Cys542Cys) Synonymous variant
	rs12327843 T>C	C_32102077_20	nt22,131	19:18004912	T>C	3'-UTR
PAX8	rs2241975 C>T	C_1940002_20	nt52,080	2:113984503	C/T	Intronic
	rs11123172 T>C	C_1940001_10	nt52,225	2:113984303	C/T	Intronic
	rs13015478 G>T	C_11511707_10	nt36,792	2:113999736	G/T	Intronic
	rs67776659 T>C	C_25607046_10	nt51,934	2:113984594	C/T	Intronic
	rs13007173 A>G	C_1940022_10	nt34,607	2:114001921	A/G	Intronic
	rs3738913 G>A	C_1940013_20	nt42,067	2:113994461	A/G	Intronic
	rs3748915 T>C	C_27494530_10	nt52,591	2:113983937	C/T	Intronic
	rs3748916 A>G	C_1939999_1_	nt52,495	2:113984033	A/G	Intronic
	rs1049137 A>G	C_1939984_10	nt61,418	2:113975110	A/G	3'-UTR
	rs4849176 T>C	C_1939985_10	nt58,592	2:113977936	C/T	Intronic
	rs4849186 G>C	C_1940020_10	nt36,994	2:113999534	C/G	Intronic
	rs1478 T>G	C_8737927_10	nt27,989	2:113975066	T>G	3'-UTR
	rs1479 T>G	C_8737926_10	nt52,228	2:113975104	T>G	3'-UTR
	rs874898 G>C	C_7507823_10	nt62,303	2:113974196	G>C	3'-UTR

Gene	SNP ID	Probe ID	Nt location	Chromosomal location	Chan -ge	A/Acid Change
DIO2	rs225010 C>T	C_26929412_10	nt181,822	14:80672279	C/T	Intronic
	rs225011 C>T	C_26929411_10	nt181,893	14:80672208	A/G	Intronic
	rs225014 T>C	C_15819951_10	nt184,521	14:80669580	C/T	ACA>GCA (T>A)
	rs225015 G>A	C_568127_10	nt186,522	14:80667579	A/G	3'-UTR
	rs225017 T>A	C_26929404_10	nt186,875	14:80667226	T/A	3'-UTR
	rs11159446 G>C	C_31755133_10	Downstream	14:80601392	C/G	Intergenic regulatory region
	Rs7140952 C>T	C_29283499_10	nt185,829	14:80668272	C/T	3'-UTR
DIO1	rs12033572 G>C	C_485944_10	nt12,763	1:54369674	C/G	Intronic
	rs12084242 T>A	C_31601232_10	nt15,213	1:54372124	A/T	Intronic
	rs2294510 C>T	C_16187176_10	nt3,011	1:54359922	C/T	CTC>CTT Synonymous
	rs2294511 A>T	C_16187186_10	nt13,353	1:54370264	A/T	Intronic
	rs12095080 A>G	C_31601225_10	nt19,819	1:54376730	A/G	3'-UTR
	rs11206244 C>T	C_334342_20	nt18,790	1:54375701	C/T	3'-UTR
	rs4926616 C>T	C_485945_10	nt12,819	1:54369730	C/T	Intronic
	rs17109582 T>C	C_25651670_10	nt18,437	1:54375348	C/T	Intronic
DIO3	rs945006 G>T	C_7565113_10	nt1590	14 : 102029277	G/T	3'-UTR
The table shows the characteristic of the 39 SNPs involved in the association study						

Appendix G: Description of the transcripts used for the studied genes

Gene	Transcipt ID
Iodothyronine Deiodinase type I (DIO1)	ENST00000361921
Iodothyronine Deiodinase type II (DIO2)	ENST00000557010
Iodothyronine Deiodinase type III (DIO3)	ENST00000510508
Thyroid Stimulating Hormone- Beta (TSHβ)	ENST00000256592
Paired box gene 8 (PAX8)	ENST00000429538
Sodium Iodide Symporter (NIS)	ENST00000222248
The table shows the transcript used for the six studied genes	