Gene expression in thyroid and thyroid cancer
Pauws, E.
Summary

Thyroid cancer is relatively rare and the prognosis of patients is one of the best of all human cancers. In spite of this, two main problems in the management of differentiated thyroid cancer remain:

- Lack of powerful diagnostic markers to distinguish between benign and malignant thyroid disease cause many patients to undergo surgery unnecessarily.
- The identification of patients with a risk for recurrent disease and/or metastasis is difficult. As a result, all patients undergo intensive initial therapy, which may not be necessary in all patients when suitable markers are available. Many clinical, histological and genetical studies have been performed to identify suitable markers. Until now, none of these studies have identified comprehensive markers able to solve the above-mentioned problems (Chapter 1). This thesis attempts to identify novel markers by gene expression profiling using the novel technique Serial Analysis of Gene Expression (SAGE). Using the SAGE technique, gene expression profiles are generated from normal and tumour thyroid tissues. Subsequently, putative genetic markers are identified and studied in more detail.

Mutations in G-proteins are frequent in human cancer. In thyroid cancer mutations in oncogenes ras and gsp are found but with a low frequency. Screening these genes for mutations in a cohort of sporadic juvenile thyroid carcinoma did not reveal any activating mutations (Chapter 2). Therefore, ras/gsp activation is not a frequent event in sporadic juvenile thyroid carcinoma. From other studies can be concluded that these oncogenes are unsuitable to distinguish high-risk from low-risk differentiated thyroid carcinoma.

The gene expression profile from normal thyroid tissue is analysed as described in Chapter 3. The presence and abundance of tags representing thyroid-specific gene transcripts show the expected expression pattern of a normal thyroid cell. Novel tags identified in this SAGE library have been used to elucidate novel proteins involved in thyroid physiology. Furthermore, this SAGE expression profile is the basis to identify differentially expressed genes in pathological thyroid tissue.

The problems of TAG-to-GENE identification in SAGE analysis are illustrated by a thyroid-specific example in Chapter 4. The thyroglobulin gene is represented by three different SAGE tags that are generated by alternative polyA cleavage sites in the pre-mRNA. Whole genome analysis shows this is a frequent phenomenon in human, rat and mouse genes. PolyA cleavage site heterogeneity has to be taken into account when analysing SAGE data.

The application of the SAGE technique on an aggressive variant of a thyroid carcinoma resulted in a gene expression profile that was compared to that of the normal thyroid. Chapter 5 describes a list of upregulated and downregulated gene transcripts that are identified by this comparison. Subsequently, the in silico Tissue Preferential Expression (TPE) analysis is performed. The TPE algorithm calculates a value for a SAGE tag of interest indicating its preferential expression in a reference tissue as compared to other tissues and pinpoints disease specific markers. Wet lab studies of gene expression in a panel of differentiated thyroid neoplasms and controls show that the transcript for Extracellular Matrix protein 1 (ECM1) is overexpressed in 50% of thyroid carcinoma, while it is completely absent in normal thyroid and thyroid adenoma.

SAGE expression profiles generate expression data from tags corresponding to known human transcript as well as from unidentified novel gene transcripts. These NoMatch tags can be used to characterise genes and their corresponding transcripts and proteins. In Chapter 5, one of these NoMatch tags is identified as
overexpressed in thyroid carcinoma. Its corresponding gene, named SMAP31, is expressed in thyroid, placenta, lung and heart. Chapter 6 describes the characterisation of the SMAP31 gene and the corresponding transcripts encoding two proteins. The main SMAP31 protein has recently been described as a homeobox transcription co-repressor named HOP, able to inhibit cardiac-specific gene expression. A similar role is hypothesized for SMAP31 in thyroid physiology.

SAGE analysis identifies ECM1 as a candidate transcript to distinguish malignant thyroid cancer from benign thyroid cancer. Artificial expression of this protein in a cell line made it possible to study the downstream sequela of ECM1. SAGE was used to generate expression profiles from cell lines expressing ECM1 and controls. Up- and downregulated genes were subjected to a specifically designed TPE analysis, enabling the identification of transcripts that are co-expressed with ECM1. The data in Chapter 7 indicate that ECM1 induces the expression of genes involved in metabolism of the extracellular matrix.

In conclusion, the analysis of expression profiles of normal thyroid tissue and thyroid carcinoma tissue using the SAGE and TPE techniques resulted in the identification of a number of candidate genes. SMAP31 and ECM1 are the most promising markers and further studies have to decide if any of these genes can help solve the diagnostic dilemmas in thyroid cancer in the future.