Cloning of the genes for non-medullary thyroid cancer: Methods and advances

Karmen Stankov¹, Giovanni Romeo²

ABSTRACT

In last ten years, significant advances have occurred in thyroid endocrinology, as a consequence of the generalized use of molecular biology techniques. New genes involved in the development of thyroid cancer have been identified, which had a great impact on our understanding of thyroid cancer predisposition. All cancers are genetic in origin because they arise from mutations in a single somatic cell, but the genetic changes in sporadic cancers are confined to a particular tissue. In inherited cancers, a predisposing mutation is present in all somatic cells and in the germ line, which enables the transmission of risk to the next generation. Cancer genetics offers a model of how information on the genetics of inherited cancers could affect identification of individuals at increased genetic risk.

KEY WORDS: Thyroid Neoplasms; Cloning, Molecular; Oncogenes; Genes, Tumor Suppressor; Genetic Predisposition to Disease

IDENTIFICATION OF THE SUSCEPTIBILITY GENE(S) THROUGH POSITIONAL CLONING

As shown in numerous studies, there is a high likelihood that non-medullary thyroid cancer (NMTC) has a familial component (1,2). However, the nature of the gene(s) that cause this predisposition is as yet unknown.

There are two main ways such susceptibility genes may be identified, the first is the identification of a biological or biochemical anomaly associated with the disease due to the identification of the biochemical change the causative gene may then be identified. The second technique is reverse genetics. Reverse genetics allows the identification of the causative gene without prior knowledge of its biological function. Therefore as the majority of genetic diseases identified cannot be associated with an obvious biochemical change, reverse genetics has been the technique of choice for identification of susceptibility genes. Reverse genetics uses the fact that the disease in question is a familial and hence members of an affected family will share a common inherited factor that is causing the disease. As reverse genetics is reliant on this fact, the first step is to identify that the trait of interest is a genetic trait.

Some inherited traits can be the result of somatic translocations of chromosomes which are segregated through the affected family, for example the FHIT: TRC8 translocation which was shown to cause susceptibility to renal cancer and to a less extent PTC (papillary thyroid carcinoma) (3) and 5q21 was identified as the candidate area for FAP (familial adenomatous polyposis) through an interstitial deletion in a patient with Gardner syndrome (4).

By examination of the families that are linked to this susceptibility locus, recombinant events will allow the identification of an area in which all the linked families share a region of DNA. In the genetic instability in cancer, loss of heterozygosity (LOH) has greatly aided the identification of tumour suppressors and gains, through techniques such as comparative staining techniques.

If cytogenetics is unable to identify the region of the genome in which the susceptibility gene is located, the next step in positional cloning is to use the technique of linkage analysis to identify the region of the genome in which the susceptibility gene is positioned. Linkage analysis works on the premise that each affected patient from each family shares a susceptibility gene. As the susceptibility gene segregates through the family the process of independent segregation of the chromosomes and recombination along the chromosome will mean that in addition to the susceptibility gene, the affected patients from each family will also share the genomic area surrounding the gene. Using one large or large sets of families, linkage analysis allows the identification of this shared genomic region and hence the approximate position of the susceptibility gene. When such a region identified and results confirmed by subsequent studies, the position of the susceptibility gene in the genome has been determined, which is named a susceptibility locus. Linkage analysis has been very successful in localizing a large number of simple Mendelian traits. However, as the complexity of the genetic trait increases the effectiveness of linkage analysis to detect susceptibility loci decreases. Familial prostate cancer is the example of how a complex genetic trait can significantly impact on the power of linkage analysis to identify susceptibility. Using large sample sets a number of prospective loci have been identified, however, due to the high phenocopy rate, late age of onset, lack of a means to stratify patients into more homogeneous groups and considerable genetic heterogeneity, confirmation of the linkage results has proven arduous.

www.onk.ns.ac.yu/Archive June 10, 2006
genomic hybridisation (CGH), offer the same potential to identify protooncogenes. An novel approach was used by Hemminki et al. (5) to map a locus on chromosome 19 in Peutz-Jeghers syndrome, this approach used CGH and LOH to identify the chromosome 19 region and then used targeted linkage analysis in affected families to confirm the result, as mentioned above the susceptibility gene was subsequently identified as STK11 (6,7), a similar approach was applied in the search for prospective BRCA3 loci, and a possible locus on chromosome 13 using CGH and targeted linkage analysis (8), however to date this linkage result has not been able to be confirmed.

Through the examination of LOH or recombiant events, if this region is sufficiently small in size, if no obvious candidate gene exists, a physical map of the area is then constructed. Previously this involved the tedious task of assembling a physical map of the region, using southern blot, to identify the correct area followed by construction of the physical map by sequencing overlaying Bacterial Artificial Chromosome (BAC) clones. Once the physical map has been constructed the genes contained in the region can be identified. Techniques such as exon trapping allow the identification of genes in the area, and once identified these genes can be screened for mutations that segregate with the disease. However, the completion of the Human Genome Project (9), and in addition the availability of the privately funded Celera raw sequence (10), has greatly facilitated the construction of such physical maps, making the construction of overlapping BAC clones and sequencing redundant. The availability of the raw sequence additionally offers a larger number of markers to allow further and finer restriction of an area of interest through the exploitation of the informative recombinations and LOH or CGH. Subsequently to the identification of the genes in that area are able to be identified, either by identification of known genes that map to that area or by prediction through computer programs, by identification of Expressed Sequence Tags (ESTs) which match the sequence, gene prediction programs and more recently comparative sequencing between mouse and human genomes allows areas that are highly conserved to be identified (highly conserved areas are under greater evolutionary pressure and therefore are more likely to be functional significant i.e. genes). Once the genes in that area are identified, screening the coding region in patients identifies mutations that segregate in affected families and allow the identification of the causative gene. Positional cloning has been successful in identifying genes implicated in human disease, notably in the field of cancer research, PTEN, APC, BRCA1, BRCA2, MENW, PAK, and STK11 (11).

CONTRIBUTION OF MOLECULAR BIOLOGY IN THYROID CANCER DIAGNOSTICS AND TREATMENT

The existence of numerous difficulties in thyroid cancer diagnostic, justifies the research of new diagnostic tools in the domain of immunohistochemistry and molecular biology. The results of immunohistochemical studies proposed a significant number of "candidate" genes, which may be screened in order to improve the thyroid cancer diagnostics, therapy or prognosis. Such "candidate" genes for diagnostics are: the loss of immunoactivity of thyroid peroxidase (TPO), expression of galectine-3, HBME-1, Ret, CK 19, PPARγ, etc. The important determinants of tumour aggressiveness are: increased expression of proliferation indicators (Ki-67 immunostaining), loss of bcl-2 expression, p53 expression, etc. Potential therapeutic efficiency may be estimated by screening the expression of iodine transporter, loss of expression of retinoic acid receptors, etc. (12).

The significant number of such genetic markers of thyroid cancer has been tested with disappointment, although many of published studies seem interesting, however, the majority of these immunohistochemistry studies are of small size, nonhomogeneous, with the results sufficiently diverse to require additional histopathological support. The similar phenomenon is occurring in molecular biology with the detection of RET rearrangements, which are involved in the majority of radiation-induced papillary thyroid carcinoma. Recent discovery of BRAF mutation exclusively in papillary thyroid carcinoma and some poorly differentiated forms, without occurrence in follicular adenoma and carcinoma, seems to open new horizons in genetic testing of thyroid cancer (13). The same hope of improved diagnostics and therapy brings the discovery of PAX8/PPARγ rearrangement in follicular cancer, as well as certain Ras mutations in papillary, follicular and poorly differentiated cancer (Table 1) (14).

**Table 1. Genes implicated in thyroid cancer (14)**

<table>
<thead>
<tr>
<th>Genes</th>
<th>Type of abnormality</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>RET/PTC</td>
<td>Rearrangement</td>
<td>Adults 20-30%, children 50-80%</td>
</tr>
<tr>
<td>TRK</td>
<td>Rearrangement</td>
<td>10% (TRK/PTM, TPR, TGF)</td>
</tr>
<tr>
<td>RAS</td>
<td>Point mutation</td>
<td>10-16%</td>
</tr>
<tr>
<td>BRAF</td>
<td>Point mutation</td>
<td>36%</td>
</tr>
<tr>
<td>CTNNB1 (gene coding for beta-catenine)</td>
<td>Mutation</td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>Mutation</td>
<td></td>
</tr>
<tr>
<td>PAX-B/PPAR gamma</td>
<td>Rearrangement</td>
<td>25-50%</td>
</tr>
<tr>
<td>RAS (N-RAS and H-RAS)</td>
<td>Mutations</td>
<td>25-50%</td>
</tr>
<tr>
<td>(codon 61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p53, PTEN, beta-catenine</td>
<td>Mutation</td>
<td></td>
</tr>
<tr>
<td>Chromosomes 2, 3, 6, 7, 8, 9, 10q, 11, 17q, 22</td>
<td>Chromosomal aberrations</td>
<td></td>
</tr>
</tbody>
</table>

FOLLICULAR CANCER

Recently, it has been reported that the translocation t(2;3)(q13;p25) is associated with the development of follicular thyroid cancer (15). The most frequently reported translocation comprises the fusion of the DNA binding domain of Pax8 transcription factor with A to F domains of PPARα factor (peroxisome proliferator activated receptor α). The expression of protein product of this translocation leads to under-expression of wild-type PPARα protein. However, the under-expression of wild-type PPARα protein seems to be a frequent event, independent from the presence of PPARα-Pax8 rearrangement (16). This decrease of expression could also be considered as an initiating event of thyroid tumorigenesis. Rearrangement of PPARα-Pax8 is detected in 8%-31% of follicular adenosomas and 35%-78% of follicular cancers (17). It is not detected in papillary thyroid carcinoma (including follicular variant), nor in Hurthle cell carcinoma or anaplastic carcinoma. In follicular cancer samples, it is rarely associated with mutations in Ras oncogene (3% of cases). The recent studies showed that two different pathways of follicular tumorigenesis exist: the pathway of Ras oncogene, leading to development of non-invasive, HBME-1 positive follicular carcinoma, and the pathway of PPARα-Pax8 leading to invasive Galectin-3 positive cancers, occurring at young age (18). The PPARα-Pax8 rearrangement is found with a high prevalence in secondary follicular cancers, following the exposition to ionising irradiation (19).

PAPILLARY THYROID CANCER

Raf proteins are the serine-threonine kinases localised downstream from the Ras in the mitotic signalization pathway of thyroid cells. In human, three Raf genes are described as ARAF, BRAF et CRAF-1. Point mutations in BRAF gene are associated with papillary thyroid cancer development, in parallel with Ret/PTC rearrangement (20). Mutations of this gene are also implicated in development of more than 66% of melanomas, as well as 2%-18% of colon, ovary and lung cancers.

Malchoff et al. (21) characterized a distinct inherited tumour syndrome as the familial association of papillary thyroid cancer, nodular thyroid disease, and papillary renal neoplasia. To characterize more fully the clinical phenotype of familial papillary thyroid carcinoma, Malchoff et al. (21) investigated the clinical and pathologic characteristics of large 3-generation PTC kindred. They performed linkage analysis to determine the chromosomal location of a susceptibility gene. In addition to the known association of PTC with nodular thy-
roid, they observed the otherwise rare entity of papillary renal neoplasia (PRN) in 2 kinds of members, one affected with PTC and the other an obligate carrier. These authors considered the multifocality of PRN in one subject to add weight to the likelihood of a true genetic predisposition to PRN. Both genetic linkage and sequence analysis excluded MET, the protooncogene of isolated familial PRN, as the cause of the PTC/PRN phenotype. A genome-wide screening and an investigation of specific candidate genes demonstrated that the PTC/PRN phenotype was linked to 1q21. A maximum 3-point log of likelihood ratio score of 3.58 was observed for markers D1S2343 and D1S2345 and for markers D1S2343 and D1S305. Critical recombination events limited the region of linkage to approximately 20 cM (21).

**HÜRTHLE THYROID CANCER**

Hürthle cell carcinomas (HCC) are distinct diagnostic category, which represent less than 5% of all differentiated thyroid malignancies. However, HCC are aggressive tumours with a poor prognosis and further elucidation of their etiopathogenesis and molecular characteristics is essential for adequate diagnostic and therapeutic approach (22). Histologically HCC are characterized by the presence of an abundant granular acidophlic cytoplasm, containing abnormally large number of mitochondria, as revealed by special cytochemical stainings and ultrastructural studies. The precise role of the mitochondrial hyperplasia in the pathogenesis of these tumours is unknown. Many studies on HCC showed changes of genomic and mitochondrial DNA, represented by the loss of the genetic material, mutations or altered expression of key genes coding the proteins involved in energy production, proliferation and apoptosis (22-26).

Both Hürthle cell carcinomas and adenomas frequently show chromosome copy number changes (23). Frequent abnormalities included a variety of gains either of chromosomes 7 and 12, or 5 and 7, or gains of all three chromosomes, suggesting a variety of pathways to Hürthle cell tumorigenesis. The study using the In Situ End-Labelling (ISEL) technique (24) showed a very high occurrence of nuclear DNA fragmentation in Hürthle cell tumours with a parallel absence of immunoreactivity for activated caspases. Peritumoral thyroid and oxyphilic non-malignant lesions as well as non-oxyphilic benign and malignant tumours, showed a focal pattern of DNA fragmentation in a lower percentage of cases. This peculiar genomic DNA fragmentation pattern in Hürthle cell tumours may be the consequence of or even the key to a rigid response of oxyphilic tumours to ischemic stimuli, leading to a necrotic rather than an apoptotic response to stress conditions (24).

Somatic mtDNA mutations have also been described in Hürthle cell tumours (25). The compensatory increase of abnormal mitochondria that occurs in Hürthle cells is greatly facilitated by the increased replication rate of mitochondria with deletions and/or mutations over that of normal mitochondria. A high percentage (nearly 100%) of Hürthle cell tumours display the mitochondrial common deletion and/or somatic mitochondrial point mutations, probably because of the high susceptibility of mtDNA to damage by ROS and mutagens (26).

Recent results of Volante et al. (27), showed that Hürthle cell tumours have a low proliferative index, as detected by Ki-67 immunostaining. In the same tumour samples they have found a surprisingly high mean expression of E2F-1 (26), the member of the transcription factor family involved in developmental, tumorigenic and apoptotic processes (29). The study of Hoos and Stojadinovic et al. (26) revealed a good correlation between molecular phenotype and clinical outcome, showing that the Ki-67(+), Bcl-2(-) phenotype was significantly associated with the diagnosis of widely invasive carcinoma as compared to normal tissue or other diagnoses. The finding of Bcl-2 down-regulation is in accordance with the results of Müller- Höcker et al. (31) and Maximo et al. (26).

**NEW FRONTIERS IN NON-MEDULLARY THYROID CANCER GENETICS**

In recent years we have witnessed an unprecedented advance in our understanding of the aetiology of inherited disease susceptibility. This progress has been made possible through the rapid development of molecular genetics and genome research and their application to human genetics. The success has been most evident in the area of the less frequent monogenic diseases, i.e. those caused by the action of a mutation at mostly a single gene locus.

In contrast multifactorial diseases such as diabetes, hypertension and some allergic diseases, to name a few, are much more frequent in the general population and reach prevalence rates of up to several percent points. These diseases are caused on the one hand by the interaction of exogenic factors like environment, life style and nutrition with variants at a number of genetic loci (oligogenic or polygenic component) on the other hand. The role of genetic factors is now well documented through twin and family studies. Since we usually see familial clustering but no simple Mendelian inheritance, the term genetic complex diseases or just complex diseases has been coined.

Cancer is another form of complex genetic disease. Most forms of cancer are characterised by the accumulation of different genetic alterations affecting genes from a set of genes with pathogenic potential, which is specific for each tumour entity. While in the majority of malignant tumours these changes are somatically acquired, some mutations are transmitted through the germ line and account for an inherited tumour predisposition ("cancer families").

The next frontier in cancer genetics is to find genes with high prevalence alleles conferring a low increase or decrease of cancer risk. The traditional kindred-based method for identifying cancer susceptibility genes is not ideal for this task. One alternative is to use very large association studies, where the number of subjects under investigation provides the power to distinguish slight variations in cancer risk. Examples of this type of studies are based on the collection of biological samples like in the Estonian Genome Project.

Following the same type of approach our group has focused on the study of the genetic predisposition to papillary thyroid carcinoma, a tumour derived from the epithelial cells of the thyroid which accounts for ~90% of all thyroid cancers. Papillary thyroid carcinoma and follicular carcinoma represent the two main variants of non-medullary thyroid carcinoma (NMTC). Epidemiological studies demonstrate that there is familial clustering of NMTC (Table 2) (14).

Table 2. Familial syndromes contributing to susceptibility of thyroid cancer (14)

<table>
<thead>
<tr>
<th>Familial syndrome</th>
<th>Manifestations</th>
<th>Thyroid tumour</th>
<th>Gene/chromosomal localisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familial polyposis of colon</td>
<td>Polypos of colon</td>
<td>Papillary thyroid cancer</td>
<td>APC/5q21</td>
</tr>
<tr>
<td>Gardner Sy</td>
<td>Polypos, osteomas, fibromas, lipomas</td>
<td>Papillary thyroid cancer</td>
<td>APC/5q21, others?</td>
</tr>
<tr>
<td>Cowden Sy</td>
<td>Hamartomas, breast Ca</td>
<td>Adenoma, follicular cancer, goitre</td>
<td>PTEN/10q22-23</td>
</tr>
<tr>
<td>Carney complex</td>
<td>Skin tumours, myxomas, Schwannomas, hypercorticism, phylloides adenomas, testicular tumours</td>
<td>Adenoma, follicular cancer, goitre</td>
<td>?, 2q16, 17q23</td>
</tr>
<tr>
<td>Multi nodular goitre</td>
<td>Thyroid tumour</td>
<td>Papillary thyroid cancer</td>
<td>14q, Xp22</td>
</tr>
<tr>
<td>Familial non-medullary thyroid cancer</td>
<td>Papillary thyroid cancer, oxyphilic tumours</td>
<td>Papillary thyroid cancer</td>
<td>2q21, 18p13</td>
</tr>
</tbody>
</table>
Familial NMTC (non medullary thyroid carcinoma) is characterized by a more aggressive behaviour than the sporadic cases. It follows an incompletely penetrant autosomal dominant mode of inheritance, and it is thought to represent ~5% of all cases of thyroid cancer. Very little is known about genetic predisposition to NMTC. To this end, since 1996, through an international clinical consortium, we are collecting blood samples from families with recurrence of NMTC and other thyroid disease (goiter, adenoma, thyroiditis). So far, we collected samples from 261 families (mostly of small size), for a total of about 845 individuals, 422 of whom affected with NMTC. DNA is available from all individuals and lymphoblastoid cell lines have been established for 305 of the NMTC patients. Tumour tissue and its normal counterparts from 150 samples of sporadic cases of NMTC have also been obtained from the patients of the members of the Consortium. This represents the largest collection of material from NMTC families in the world.

The approach used for localization of predisposing genes has been so far the linkage analysis, carried out using a genotyping panel of ~400 microsatellite markers covering the entire genome. We focused on the largest pedigrees, which always show recurrence of NMTC and of benign thyroid disease (most frequently multinodular goiter, MNG). In this way we contributed to the localization of MNG1, a gene predisposing to MNG and papillary thyroid cancer. We subsequently localized in a large French pedigree another gene, TCO, which predisposes to thyroid tumours with cell oxyphilia. Other large pedigrees are under investigation. The series of small families has been analyzed so far only at selected candidate regions (MNG1, TCO and RET), which did not show any significant linkage (32).

The candidate region for TCO, identified on chromosome 19, remains very broad after initial genome screening (about 7 Mb) (33). In order to achieve the positional cloning of TCO, our methodological approach includes utilization of new polymorphic markers and further examination of public databases for selection of known genes situated in this region.

The linkage to TCO is compatible with the segregation of chromosome 19 markers in 6 out of 8 families with oxyphilic tumours. In one family of Austrian descent the linkage to 19p13 region was confirmed and candidate region reduced to less than 2 Mb. In the oxyphilic carcinoma tumour tissue from one member of French family, the FISH analysis identified a deletion present in the tumour, but not in the normal tissue. From the haplotype analysis it appears that this patient transmits to his progeny the mutated TCO allele which is not the allele deleted in the tumour, thus suggesting that TCO might be a tumour suppressor gene. This cytogenetic result has been confirmed by LOH analysis (34). The typical oxyphilic appearance of these tumours is caused by an abundance of mitochondria in the cell and this feature was one of the criteria applied in selection of candidate genes. A systematic screening of the genes mapped to the region has been carried out in TCO patients contributing to the linkage signal. Genes with mitochondrial functions, considering the high copy number of mitochondria present in these tumour types, and/or with role in proliferation, differentiation and programmed cell death, were selected for analysis. Genomic DNA from the patients was amplified and direct sequencing was carried out for the coding and regulatory regions of the following genes: EDG5, CystC, LASS1, LASS4, TIMM4, CLC25, ELAV1. Five missense changes were identified in these genes; however, these variants were either known SNPs (by comparison with dbsNP database), or they did not cosegregate with the tumour phenotype. Further screening of other candidate genes in the 19p13.2 region is currently ongoing.

Since the mutations in BRAF have been reported with high prevalence in sporadic papillary thyroid carcinomas, we investigated whether germ-line mutations in BRAF might be predisposing events in familial cases. One relative per family was chosen for the analysis of exon 11 and 15 of BRAF, where the majority of mutations were previously found. Direct sequencing did not reveal any mutations in the exons. In parallel, screening of the group of sporadic cases identified the presence of BRAF T1796V (V599E) activating mutation in 11.6% of PTC cases, with a frequency corresponding to the results of previously reported studies. Interestingly, the same mutation was identified in 1 case of oxyphilic follicular carcinoma (20).

The relationship between the mitochondrial proliferation and pathogenesis of these tumours is still unknown and presents a lead for the functional studies we are performing. The data on activity of mitochondrial respiratory chain activities showed significantly increased activity of mitochondrial complex I and significantly decreased activity of complex III, which is further emphasized by the increased production of the reactive oxygen species (ROS) in XTC-1 tumour cell line derived from breast metastasis of oxyphilic thyroid tumour, in comparison with non-oxyphilic PTC cell line (35,36). One of the subunits of mitochondrial complex I, which is encoded by nuclear genome, has been mapped to 19p13 region, within the interval of recombination observed and this gene is considered as an interesting candidate gene in our study.

Other large pedigrees are under investigation. An extensive genome wide scan of a large Tasmanian family with recurrence of PTC revealed a common haplotype on 2q21, and the subsequent linkage analysis of other 80 families with at last two cases of PTC, confirmed a significant LOD score (37,38). This has increased after stratification of 17 pedigrees based on the presence of at least 1 case of the follicular variant of PTC (fPTC), phenotype observed in the Tasmanian family. Recombinations in these pedigrees with fPTC identified a critical region, named TMC1 that spans now few kb, which we are characterizing the genomic organization of the region is complicated by the vicinity of an ancient centromere, which has lost its function but has kept the characteristic structure and might therefore give origin to duplications, rearrangements, repetitions, pseudogenes. There are three genes in this region, with very peculiar genomic organization. We are performing mutation analysis in the most interesting transcribed regions of these genes and in other functional candidate genes in patient genomic DNA samples. In addition we are trying to characterize the molecular organization of a breakpoint observed in a cell line, derived from a benign PTC which presents a translocation in the same region of interest and might therefore give us important clues.

Acknowledgements

Karmen Stankov held a Postdoctoral Fellowship from the International Agency for Research on Cancer and American Cancer Society while working on the study reported in this paper.

REFERENCES
