Editorial

Unraveling the Genetic Predisposition to Differentiated Thyroid Carcinoma

Albert de la Chapelle

Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio 43210

Case-control studies have established with little doubt that the etiology of papillary thyroid carcinoma (PTC) and follicular thyroid carcinoma (FTC) is heavily influenced by hereditary factors, ie, predisposing genes. In classic studies from Utah (1) and Sweden (2), the risk to first-degree relatives of probands was 8- and 12-fold, respectively, being among the highest of all cancers (3) and sharing the top rankings with multiple myeloma and testicular cancer, among others. One interesting observation in this regard is that none of the 3 above-mentioned cancers is known to occur often in large families. Typical pedigrees often do not evoke Mendelian inheritance but instead are small with irregular transmission of the cancer phenotype. Perhaps most importantly, in these cancers, no predisposing genes have yet been detected that account for anything but a tiny fraction of all familial or sporadic cases. In contrast, some other cancers such as colorectal, breast, and prostate cancers have a substantially lower degree of heritability/familiality as shown in case-control studies (2- to 4-fold enrichment in first-degree relatives), yet gene mutations such as the mismatch repair genes and the *BRCA* genes account for sizeable fractions of the disease phenotypes.

The above facts speak a clear language; the mechanisms of inheritance are different in the two groups of cancers. In PTC/FTC, this is obvious in that among the few genes known to or believed to strongly predispose to PTC so far, none accounts for more than a handful of cases. In contrast, genes displaying weak predisposition may be common. In other words, high-penetrance genes are relatively rarely involved (but there may be many of them), whereas low-penetrance genes may be a more important basis for predisposition. We examine these two classes of genes here.

The classic method to look for high-penetrance genes is by linkage analysis in affected families followed by positional cloning of the culpable gene, followed by the demonstration of disease-causing mutations therein. Several early studies based on numerous PTC families yielded positive evidence of linkage, a crucial first step. Among the much-publicized loci detected were 14q31 (4), 1q21 (5), 2q21 (6), and 19p13 (7). Somewhat surprisingly, attempts at positional cloning of the putative genes in these loci were not successful (8). Among the possible explanations, we note that some of the linkage analyses included multiple nodules as the affected phenotype alongside cancer cases. Because multiple nodules are by no means an obligatory precursor to cancer, the linkage results in question may have been misleading.

A few loci found by the linkage approach have led to some insight. In a linkage locus in chromosome 8q24, a short haplotype that segregated with PTC located in 3 introns of the thyroglobulin (*TG*) gene was found (9). This region apparently contains a long noncoding RNA gene; however, the gene has not been well defined, and its function is unknown. A locus in 4q32 showing strong linkage was identified in a large PTC family. Being located in a gene-poor region, it turned out to comprise a deleterious single-nucleotide mutation in an enhancer, but so far, the relevant target genes of the enhancer have not been identified. Interestingly, this mutation was not found in over 2600 sporadic PTC cases or in over 2400 controls, suggesting that the mutation is very rare (10). By linkage analysis, a locus in chromosome 12q14 turned out to harbor the *SRGAP1* gene that displayed missense mutations in 3 PTC families and some sporadic cases. Although the functional effects of the mutations were highly suggestive of being deleterious, population data indicated low penetrance. Thus, the overall role of *SRGAP1* mutations in

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

Received September 13, 2013. Accepted September 16, 2013.

For article see page E1674

Abbreviations: FTC, follicular thyroid carcinoma; GWAS, genome-wide association study; NGS, next-generation sequencing; OR, odds ratio; PTC, papillary thyroid carcinoma; SNP, single-nucleotide polymorphism.

the etiology of PTC is likely small (11). If it turns out that many mutations predisposing to thyroid cancer are rare (but some with high penetrance), this certainly explains why linkage is not likely to disclose many of them.

Today, high-penetrance genes are preferentially sought by next-generation sequencing (NGS) in families with multiple affected members. Note that a family with, say, 5 PTC cases available for study will at best yield only weak or suggestive evidence of linkage, whereas a successful search for mutations by NGS may produce a decisive result.

The present time is one of rapid development both in the sequencing technology and in the bioinformatics and statistics needed to handle and interpret the data. Already whole-exome sequencing and whole-genome sequencing are used in hundreds of laboratories around the world to search for genes predisposing to cancer. In addition, The Cancer Genome Atlas initiative (http://cancergenome.nih.gov) and the International Cancer Genome Consortium (http://www.icgc.org) are acquiring massive data on the DNA and RNA sequence and methylation status of germline and tumor tissue from numerous cancers. These data are in part available to all researchers subject to specific application procedures. Numerous opportunities are already available (12).

As far as we are aware, no study has so far been published on NGS in the predisposition to differentiated thyroid cancer. In the author's laboratory, 7 PTC families were studied by whole-exome sequencing, and one serious candidate gene was found in one of the families. Functional studies are underway to determine whether predisposition to PTC is a likely consequence of the missense mutation found. It is anticipated that the large-scale application of whole-genome sequencing that is presently ongoing in PTC (both familial and sporadic) will soon show whether our hypothesis, that high-penetrance genes exist but are rare, is right.

The modern way of looking for low-penetrance genes in human disease is by genome-wide association studies (GWASs). An example is provided in this issue of the *Journal* (13). Briefly, in a GWAS, a discovery cohort of cases and controls is studied for the occurrence of single-nucleotide polymorphisms (SNPs). Whenever the frequency of an SNP is different in the cases and controls, a hit has been made. To consolidate the findings from the discovery set, the most promising SNPs are typed in further cohorts of cases and controls (validation sets), preferably from different populations. SNPs that fulfill rigorous statistical criteria in more than one population are said to show association. Today, typically, the cohorts of cases and controls are large, many hundreds to many thousands, giving power to detect minute differences in SNP frequency. The numbers of (mostly chip-based) SNPs being tested are equally huge, most often in excess of half a million. Large numbers of studies have been published, and the number of different human diseases and phenotypes tested is rapidly growing. According to the National Human Genome Research Institute web site (http://www.genome.gov/gwastudies/), by September 2013, 1699 GWASs have reported 11 490 SNPs that have been implicated in various diseases and phenotypes, notably in cancer.

What is the significance of disease-associated SNPs? They more often reside outside of than within genes, and locations in gene deserts are common. Understandably, a single nucleotide substitution in a noncoding DNA sequence is mostly not easy to interpret. It is commonly proposed, but has not been often proven, that SNPs showing association represent flags for the true disease-causing event (SNP or other) in the vicinity, being flagged through linkage disequilibrium. The potential importance of an associated SNP depends on several factors, namely 1) the odds ratio (OR) hinting about the penetrance, 2) the statistical significance (*P* value) hinting about the robustness of the association, 3) the population incidence of the risk allele indicating the proportion of cases that will be influenced, and 4) the location, eg, in or outside of genes that may or may not hint about the causative variant. Prototype SNPs detected by GWASs have ORs between 1.2 and 1.5, *P* values $< 10^{-7}$, frequencies of risk allele from 0.1 to 0.6, and location outside of coding genes or in introns.

Two early GWASs were initiated in the Icelandic population and validated in The Netherlands, Spain, and Ohio (14, 15). Together, they disclosed 5 independently associated SNPs of variable types. Obviously interesting among the SNPs were those located in introns of genes: rs2439302 in an intron of *NGR1* and rs966423 in the *DIRC3* gene. However, no further information, mechanistic or epidemiologic, has been published. In the author's laboratory, the SNP rs944289 located in chromosome 14 has been further clarified. It appears that the SNP is the actor rather than a flag because it destroys a transcription factor binding site regulating a so-called long intergenic noncoding RNA that we named PTC susceptibility candidate 3 (*PTCSC3*) (16). The most interesting of the remaining SNPs is rs965513 in 9q22 that has an OR of 1.6 to 2.0. Unpublished results suggest that it, too, affects a long intergenic noncoding RNA, but work to determine its putative target genes is underway. Interestingly, this locus was found to be clearly associated with PTC arising in children from Belarus exposed to the Chernobyl accident (17). The paper in this issue of the *Journal* is a large GWAS of several south and central European populations. The main finding, in addition to confirmation of the previously

detected strong 9q22 locus (rs965513), was the confirmation of association of the *DIRC3* locus (13).

The ultimate significance of GWAS-detected low-penetrance genetic loci remains to be discussed. Two major outcomes are anticipated. First, once the culpable gene in a locus has been identified and the effects of its mutations clarified, these findings have the potential of stimulating the development of preventive or therapeutic measures. The road to this type of development is arduous, and we are not aware of any examples in thyroid carcinoma. Second, these loci may be used to guide diagnostics and counseling. The most typical situation might be risk assessment in relatives of probands with the disease. It has recently been demonstrated in PTC that the risk from GWAS-derived loci is additive. In a large series of PTC and FTC patients, the 5 SNPs detected by GWAS (14, 15) were analyzed and the individual risks calculated. It could be shown that individuals having 7 or more risk alleles (the maximum from 5 variants being 10) had OR values as high as 13. Individuals with few risk alleles had ORs close to 1 (18). Although figures of this magnitude probably will not be used in diagnostic assessments at the present time, it is obvious that as more risk loci are detected in the future, combined risk figures may well be significant enough to be used in counseling and diagnostics. Thus, one can safely conclude that more GWASs in thyroid cancer will be done and will eventually produce markers that, used with prudence, will prove to be valuable clinical tools.

Acknowledgments

Address all correspondence and requests for reprints to: Albert de la Chapelle, MD, PhD, Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, 804 Biomedical Research Tower, 410 West 12th Avenue, Columbus, Ohio 43210. E-mail: albert.delachapelle@osumc.edu.

This work was aided by National Institutes of Health Grants P30 CA16058 and P01 CA124570.

Disclosure Summary: The author has nothing to disclose.

References

1. **Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH.** Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst*. 1994;86:1600–1608.
2. **Dong C, Hemminki K.** Modification of cancer risks in offspring by sibling and parental cancers from 2,112,616 nuclear families. *Int J Cancer*. 2001;92:144–150.
3. **Risch N.** The genetic epidemiology of cancer: Interpreting family and twin studies and their implications for molecular genetic approaches. *Cancer Epidemiol Biomarkers Prev*. 2001;10:733–741.
4. **Bignell GR, Canzian F, Shayeghi M, et al.** Familial nontoxic multinodular thyroid goiter locus maps to chromosome 14q but does not account for familial nonmedullary thyroid cancer. *Am J Hum Genet*. 1997;61:1123–1130.
5. **Malchoff CD, Sarfarazi M, Tendler B, et al.** Papillary thyroid carcinoma associated with papillary renal neoplasia: genetic linkage analysis of a distinct heritable tumor syndrome. *J Clin Endocrinol Metab*. 2000;85:1758–1764.
6. **McKay JD, Lesueur F, Jonard L, et al.** Localization of a susceptibility gene for familial nonmedullary thyroid carcinoma to chromosome 2q21. *Am J Hum Genet*. 2001;69:440–446.
7. **Canzian F, Amati P, Harach HR, et al.** A gene predisposing to familial thyroid tumors with cell oxyphilia maps to chromosome 19p13.2. *Am J Hum Genet*. 1998;63:1743–1748.
8. **Bonora E, Tallini G, Romeo G.** Genetic predisposition to familial nonmedullary thyroid cancer: An update of molecular findings and state-of-the-art studies. *J Oncol*. 2010;385:206.
9. **He H, Nagy R, Liyanarachchi S, et al.** A susceptibility locus for papillary thyroid carcinoma on chromosome 8q24. *Cancer Res*. 2009;69:625–631.
10. **He H, Li W, Wu D, et al.** Ultra-rare mutation in long-range enhancer predisposes to thyroid carcinoma with high penetrance. *PLoS One*. 2013;8:e61920.
11. **He H, Bronisz A, Liyanarachchi S, et al.** SRGAP1 is a candidate gene for papillary thyroid carcinoma susceptibility. *J Clin Endocrinol Metab*. 2013;98:E973–E980.
12. **Lawrence MS, Stojanov P, Polak P, et al.** Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499:214–218.
13. **Köhler A, Chen B, Gemignani F et al.** Genome-wide association study on differentiated thyroid cancer. *J Clin Endocrinol Metab*. 2013;98:E1674–E1681.
14. **Gudmundsson J, Sulem P, Gudbjartsson DF, et al.** Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nat Genet*. 2009;41:460–464.
15. **Gudmundsson J, Sulem P, Gudbjartsson DF, et al.** Discovery of common variants associated with low TSH levels and thyroid cancer risk. *Nat Genet*. 2012;44:319–322.
16. **Jendrzejewski J, He H, Radomska HS, et al.** The polymorphism rs944289 predisposes to papillary thyroid carcinoma through a large intergenic noncoding RNA gene of tumor suppressor type. *Proc Natl Acad Sci U S A*. 2012;109:8646–8651.
17. **Takahashi M, Saenko VA, Rogounovitch, et al.** The FOXE1 locus is a major genetic determinant for radiation-related thyroid carcinoma in Chernobyl. *Hum Mol Genet*. 2010;19:2516–2523.
18. **Liyanarachchi S, Wojcicka A, Li W, et al.** Cumulative risk impact of five genetic variants associated with papillary thyroid carcinoma. *Thyroid*. In press.