New Insights into Testicular Germ Cell Tumorigenesis from Gene Expression Profiling

Rolf I. Skotheim, Outi Monni, Spyro Mousses, Sophie D. Fosså, Olli-P. Kallioniemi, Ragnhild A. Lothe, and Anne Kallioniemi

Department of Genetics, Institute for Cancer Research [R. I. S., R. A. L.] and Department for Oncology and Radiotherapy [S. D. F.], The Norwegian Radium Hospital, N-0310 Oslo, Norway; Biomedical Biochip Center, Biomedical Helsinki, FI-00290 Helsinki, Finland; and Laboratory of Cancer Genetics, Institute of Medical Technology, University of Tampere and Tampere University Hospital, FIN-33520 Tampere, Finland [A. K.]

ABSTRACT

We have shown recently that about half of the human TGCTs reveal DNA copy number increases affecting two distinct regions on chromosome arm 17q. To identify potential target genes with elevated expressions attributable to the extra copies, we constructed a cDNA microarray containing 636 genes and expressed sequence tags from chromosome 17. The expression patterns of 14 TGCTs, 1 carcinoma in situ, and 3 normal testis samples were examined, all with known chromosome 17 copy numbers. The growth factor receptor-bound protein 7 (GRB7) and junction plakoglobin (JUP) were the two most highly overexpressed genes in the TGCTs. GRB7 is tightly linked to ERBB2 and is coexpressed with this gene in several cancer types. Interestingly, the expression levels of ERBB2 were not elevated in the TGCTs, suggesting that GRB7 might be the target for the increased DNA copy number in TGCTs. Because of the limited knowledge of altered gene expression in the development of TGCTs, we also examined the expression levels of 512 additional genes located throughout the genome. Several genes novel to testicular tumorigenesis were consistently up- or down-regulated, including POVI, MYCL1, MYBL2, MXI1, and DNMT2. Additionally, overexpression of the proto-oncogenes CCND2 and MYCN were confirmed from the literature. The overexpressions were for some of the target genes closely associated to either seminoma or nonseminoma TGCTs, and hierarchical cluster analysis of the gene expression data effectively distinguished among the known histological subtypes. In summary, this focused functional characterization of TGCTs has lead to the identification of new gene targets associated with a common genomic rearrangement as well as other genes with potential importance to testicular tumorigenesis.

INTRODUCTION

TGCT is the most common malignancy among adolescent and young adult Caucasian males, and the incidence has been steadily increasing over the past 50 years (1, 2). TGCTs are classified into two main histological subtypes, seminomas and nonseminomas, and there are two models describing their development from carcinomas in situ (3). Either both subtypes develop independently from carcinomas in situ, or they develop as a continuum where seminomas may progress further into nonseminomas.

TGCTs are generally in the triploid range, and isochromosome 12p or gain of DNA sequences from chromosome arm 12p is a characteristic of virtually all TGCTs (4, 5). In addition, specific gains and losses from several other chromosomal regions have been described. Although molecular studies have shown some genes to be altered at the DNA and/or expression levels in a limited number of TGCTs, the target genes reflecting the nonrandom chromosomal aberrations remain unknown (for a review of TGCT genetics, see Refs. 6 and 7).

We have demonstrated recently by a genome-wide copy number analysis using CGH that sequences on chromosome arm 17q are frequently overrepresented in TGCTs, and two common regions of copy number increase were identified (8). Gain of the proximal region, 17q11–q21, is preferentially observed in nonseminomas, whereas gain of the distal region, 17q24–qter, is common to all TGCTs (8). Nonrandom gain at 17q has also been reported in several other cancer types (9–13).

To identify differentially expressed genes on chromosome 17 in TGCTs, we analyzed a series of TGCTs and normal testicular samples using a custom-made cDNA microarray with a comprehensive chromosome 17 coverage (14). All analyzed tumors had been studied previously by CGH, and thus, the expression profiles could be related to the DNA copy numbers. The expression levels of 512 additional genes mapping elsewhere in the genome, including many cancer-related genes, were also analyzed in the same set of TGCT samples.

MATERIALS AND METHODS

Tumor and Cell Line Samples. Eighteen testicular tissue samples were analyzed, including 8 pure seminomas, 6 nonseminomas, 1 carcinoma in situ (from the vicinity of a nonseminoma), and 3 normal samples. The tumors were selected from a series of primary TGCTs analyzed previously by CGH (8). Four of the 8 seminomas had gains at distal 17q, including the 17q24–qter region (Fig. 1A). The 6 nonseminomas (4 embryonal carcinomas and 2 immature teratomas) had large gains at chromosome 17, all including the 17q11–q21 region. The use of these samples in cDNA microarray experiments was approved by the Regional Committee for Medical Research Ethics in Norway and the NIH Office of Human Subjects Research.

A pool of two breast cancer cell lines, HBL100 and MDA-436 (American Type Culture Collection, Manassas, VA), was used as a common reference in the cDNA microarray experiments. These cell lines were selected based on the facts that they show no increase in copy number at 17q and express most genes on the cDNA microarray to some extent (10, 14).

cDNA Microarray Experiments. The construction of the cDNA microarray with comprehensive chromosome 17 coverage has been described previously by Monni et al. (14). The microarray consisted of printed PCR products from 636 sequence-verified IMAGE cDNA clones (Research Genetics, Huntsville, AL), including 201 known genes from the entire chromosome 17 and 435 ESTs from the 17q arm. An additional 512 sequence-verified IMAGE cDNA clones were placed on the array, representing transcribed sequences located elsewhere in the genome. Eighty-eight of these were housekeeping genes and were used for calibration among the different experiments (15, 16). 162 were a selection of known or putative cancer-related genes, and 262 were a collection of genes and ESTs from chromosome 10.

Preparation and printing of the cDNA clones on glass slides, probe preparations, hybridizations, and image generation and analyses were performed as described (16). Briefly, mRNA was isolated from the test samples by using the Trizol reagent (Life Technologies, Inc., Rockville, MD) and oligo(dT)25 dyna-
beads (Dynal Biotech, Oslo, Norway) according to the manufacturers’ specifications. From the reference cell lines, mRNA was isolated directly by using FastTrack 2.0 mRNA isolation kit (Invitrogen, Carlsbad, CA). Labeled cDNA was synthesized from 1–30 ng mRNA (test or reference, respectively) in an oligo(dT)-primed polymerization with SuperScript II reverse transcriptase (Life Technologies) in the presence of either Cy3 (test) or Cy5 (reference) labeled dUTP (Amersham Pharmacia, Piscataway, NJ). The Cy3-labeled test cDNA from the various testicular samples and Cy5-labeled reference cDNA were mixed and simultaneously hybridized to the cDNA microarray. The fluorescence intensities at the targets were detected by a laser confocal scanner (Agilent Technologies, Palo Alto, CA). For each array element, a ratio between the relative fluorescence intensities of the test and reference was calculated. This ratio was divided by the average expressions of the 88 housekeeping genes, giving a calibrated ratio. The calibrated ratio was then normalized by dividing it by the average calibrated ratio of the three normal testicular samples. Thus, this normalized ratio reflects relative up- or down-regulated gene expression from normal to neoplastic testicular tissues.

Statistics. A two-tailed t test for equality of means was used to calculate the statistical significance of differences in expression levels between different groups of samples. The hierarchical cluster analysis was done on successful gene elements (i.e., clones where all experiments had spot sizes >100 area units and fluorescence intensities stronger than 200 fluorescence units) with more than 4-fold differential expression within the sample set. The resulting data were hierarchically clustered by both gene and sample sides (501 clones and 18 experiments). The average-linkage clustering method was used with Pearson’s correlation similarity measure. Before calculation of the correlation between two genes or samples, the original ratio was log transformed, followed by subtracting the mean from the ratio. The sample tree (dendrogram) is drawn with “real” instead of fixed distances (in-house cluster analysis software at the National Human Genome Research Institute, NIH).

Validation by Real-Time RT-PCR. We used real time RT-PCR (TaqMan system; Applied Biosystems, Foster City, CA) to validate the mRNA expression levels of three genes (GRB7, JUP, and POV1) in 10 testicular samples (3 normal testicular tissues and 7 TGCTs). In this quantitative RT-PCR, a dual-labeled fluorescent probe is degraded concomitant with PCR amplification. Input target mRNA levels are calculated from the time (measured in PCR cycles) at which the reporter fluorescent emission increases beyond a threshold level, as measured by an ABI PRISM 7700 Sequence Detector (Applied Biosystems).

Primers and probes targeting the mRNA sequences (Table 1) were designed using the Primer Express software (Applied Biosystems). cDNA synthesized from 50 ng of mRNA was used as PCR template in a total volume of 25 μl containing 200 nM of each oligonucleotide primer and probe, 0.2 mM of each deoxynucleotide triphosphate, 1 × TaqMan buffer, 6 mM MgCl2, 1.25 units of AmpliTaq Gold, 0.25 units of AmpErase UNG (all Applied Biosystems), and 0.8% glycerol. The PCR program was initiated by 2
min at 50°C and 10 min at 95°C before 40 thermal cycles, each of 15 s at 95°C and 1 min at 60°C.

Primers and probe targeting GAPDH were multiplexed together with primers and probes targeting each gene of interest. For both the test genes and GAPDH, standard curves were made from which relative expression values were calculated. The expression levels of the genes of interest were then calibrated by dividing by the expression of GAPDH. Again, division by the average values of the three normal testicular samples normalized all calibrated expression values, and thus, these values were comparable with the normalized ratios from the microarray experiments.

RESULTS

The expression levels of 636 chromosome 17-specific transcripts as well as 512 transcripts located elsewhere in the genome were determined in 18 testicular tissue samples by cDNA microarrays. Hierarchical cluster analysis with a set of 501 differentially expressed genes separated the TGCT samples according to their known histological subgroups (Fig. 1B). The single carcinoma in situ sample, representing a precursor stage, was most closely related to the normal testis specimens. The seminomas formed a single cluster, whereas within the nonseminomas, immature teratomas and embryonal carcinomas clustered into separate groups. A comprehensive gene expression map for the 51 genes that were differentially expressed at a 0.01 significance level and had on average >3-fold up- or down-regulation across all tumors, or within a histological subgroup, is shown in Fig. 1, C and D.

To identify up-regulated genes from the two regions with frequent copy number increase on chromosome 17, the normalized ratios of the genes were plotted as a function of their physical map positions (Fig. 2). This visualization indicated that not all transcripts located in a region with increased copy number show increased expression. At the proximal region (17q11–q21), GRB7 and JUP were consistently the most overexpressed transcripts in the TGCTs (Fig. 1C). At the distal region (17q24–qter), LGGL2, PDE6G, and EST (IMAGE clone 124915) were the most up-regulated transcripts (Fig. 1C). Among the overexpressed genes on 17q, GRB7 was significantly more expressed in nonseminomas and JUP in seminomas (both P < 0.01).

For the clones mapping elsewhere in the genome, the most up-regulated transcribed sequences in TGCTs, i.e., the highest average normalized ratios across all tumor samples, were in decreasing order MYBL2, CCND2, MYCN, POV1, EST (272938), and MYCL1. The average expression levels of POV1 and MYCL1 were significantly higher in seminomas than in nonseminomas (P < 0.01). The expression data also revealed several genes, such as MXII and DNMT2, that were down-regulated in the TGCTs (on average >3-fold down-regulated and P < 0.01 for differential expression between normal testis and TGCTs; Fig. 1, C and D).

The expression levels of GRB7, JUP, and POV1 were validated in 10 samples by real time RT-PCR, and overexpression in tumors (compared with normal testicular samples) were seen by both methods (Fig. 3).

DISCUSSION

Increased DNA copy number is a common mechanism for overexpression of genes promoting neoplastic and malignant cell behavior. In TGCTs, frequent DNA copy number increase of several chromosomal regions has been observed (17, 18). However, not much is known with regard to the specific genes that are targeted for overexpression. Recently, we identified two novel regions on chromosome arm 17q with common copy number increase in TGCT (8). In the present study, we used gene expression analysis by cDNA microarrays as a high throughput method to identify potential target genes in these two regions. The comprehensive coverage of the microarray enabled us to determine which genes were overexpressed in TGCT, as compared with normal testicular tissue, and therefore most likely to be involved in driving the genomic alteration. Furthermore, the microarray was constructed to include several additional genes with known or
putative cancer-related functions, which makes the present study the most extensive expression analysis of potentially cancer-promoting genes in TGCT.

The cDNA microarray-based expression survey in TGCTs and normal testis samples revealed several novel results:

(a) the hierarchical cluster analysis of the cDNA microarray data grouped the samples according to their correct histological subtypes. This is rather surprising, taken into account the limited number and highly selected nature of the transcripts included in this analysis, and might indicate that genes located on chromosome 17 are fundamental for the biological characteristics of these tumors.

(b) Comparison of the microarray expression data and the DNA copy number increases along chromosome 17 as determined by CGH showed that most genes were not transcriptionally up-regulated because of extra DNA copies. These results are in line with our previous data on breast cancer (14) and indicate that increased gene copy number does not always lead to increased gene expression.

In the present study, we have identified overexpressed genes located in the two common regions of copy number increase on chromosome 17 in TGCTs. At the proximal region (17q11–q21), the cDNA microarray survey showed consistent overexpression of the GRB7 and JUP genes. GRB7 is closely linked to the ERBB2 oncogene (20 kb apart), and has been shown frequently coamplified and coexpressed with ERBB2 in breast, esophageal, and gastric cancers (19–22). Interestingly, the expression of ERBB2 was not elevated in any of the TGCT samples. To our knowledge, this represents the first example where increased copy number at the ERBB2 locus does not lead to transcriptional activation of ERBB2. Furthermore, this indicates that other genes at this locus, such as GRB7, are critical for the development of TGCTs and possibly to other tumor types with ERBB2 amplification.

GRB7 encodes an adaptor protein that through its Src homology 2 domain interacts with the cytoplasmic domain of the growth factor receptor ERBB2 (19). Thus, increased expression of one of these proteins may be sufficient to promote tumor development. GRB7 also binds to several other tyrosine kinase growth factor receptors, including KIT, platelet-derived growth factor receptor, RET, and INSR (23–26), as well as to cytoplasmic tyrosine kinases (27, 28). The KIT proto-oncogene has previously been suggested to play a role in TGCT development, both attributable to increased expression (29), and by its involvement in survival, proliferation, and migration of primordial germ cells (30). Additional knowledge about GRB7 that strengthens its potential importance in TGCT development is the RAS-associating-like domain (31) and its role in cell migration (32, 33). Interestingly, the expression of GRB7 in esophageal carcinomas is related to metastatic progression (34).

JUP is also located within the proximal 17q region gained in TGCTs. It was up-regulated in tumors, with and without genomic gain by CGH. JUP belongs to the catenin family and encodes a submembranous junctional plaque protein in both desmosomes and intermediate junctions. It may have oncogenic potential through its function in the Wnt signaling pathway (35, 36), although the importance of this pathway in TGCT remains to be elucidated.

At the distal 17q region frequently gained in TGCTs, 17q24–qter, the cDNA microarray analyses implicated the LLGL2 and PDE6G genes, as well as an uncharacterized EST (124915), as consistently up-regulated in the TGCTs. The Drosophila orthologue to LLGL2, 1(2)gl, functions as a tumor suppressor (37). However, the function may be different in germ cells, because 1(2)gl is required for survival of germ-line cells in Drosophila (38), and our results show that LLGL2 is up-regulated in human TGCT. Furthermore, LLGL2 are abundantly represented in some cDNA libraries derived from human lung and prostate tumors. PDE6G encodes an effector protein involved in phototransduction in the eye (39), and to our knowledge, no cancer-related function has been linked to this gene.

From the genes located elsewhere in the genome, three human homologues of avian retroviral oncogenes, MYCN at 2p24.1, MYCL1 at 1p34.3, and MYBL2 at 20q13.1, were among the most overexpressed genes in the TGCTs. The chromosomal locations of MYCN and MYCL1 are both within regions that are gained in approximately one-third of all TGCTs, whereas the locus of MYBL2 is rarely involved in copy number changes (8). All three gene products are localized to the nucleus and function as transcriptional transactivators. The MYCN overexpression has been detected previously in TGCT (40). Remarkably, studies of neuroblastomas give evidence for both statistical and structural associations between MYCN overexpression and gain of 17q21–ter (12, 41). Furthermore, several E-boxes (the common DNA binding site of the MYC family proteins) are found in the promoter region of CCND2, and MYC overexpression has been shown to induce chromosomal and extrachromosomal instability of the CCND2 gene at 12p13 (42).

Gain of chromosome arm 12p, often through the presence of isochromosome 12p (4), is the most common genetic aberration in TGCTs. Two smallest regions of overlapping amplifications on 12p have been suggested, one at 12p13 (8) and one more proximal region (43), harboring the candidate genes CCND2 and KRAS2, respectively. We showed that both genes were transcriptionally up-regulated in TGCTs, but CCND2 significantly more than KRAS2. Activating mutations of KRAS2 have only rarely been detected in TGCTs (44, 45), further reducing the importance of this proto-oncogene in TGCTs. The observed overexpression of CCND2 is in keeping with a study by Houldsworth et al. (46), where CCND2 had the highest increased expression among a set of six candidate genes on 12p, and with a study by Bartkova et al. (47), finding CCND2 protein expression related to early stages of TGCTs.

The POU1 gene at 11q12 was highly expressed in all seminomas and in the carcinoma in situ, but neither in the normal samples nor in the nonseminomas. Thus, independently of developmental model, this gene may be an early-onset gene in the development of seminoma.

---

4 Internet address: http://genome.ucsc.edu/ (Aug. 6, 2001 freeze).
TGCTs. In analogy, precursor lesions of prostate cancer have shown increased expression of POVI (48). In addition to DNA copy number, other means of transcriptional control are obviously important to TGCT development. These may include regulation of transcription factors and disruption of the DNA methylation pattern. Hence, the gene products of MXI1 and DNMT2 are potential candidates because of their consistently down-regulated mRNA levels demonstrated by our microarray survey. The MXI1 protein is a transcriptional repressor through its binding to MAX, a MYC heterodimerization partner (49). Thus, by its competition for MAX, MXI1 antagonizes the MYC transcription factors, of which we have shown increased mRNA levels in TGCTs for both MYCN and MYCL1. Interestingly, the MXI1 mouse homologue is mapped within the region with highest score in a genome-wide linkage analysis targeting TGCT susceptibility loci in mice (50, 51). Furthermore, the MXI1 gene is commonly mutated or deleted in prostate carcinomas (52). The DNMT2 gene has strong sequence homology to the DNA- (cytosine-5)-methyltransferases, although its catalytic activity has yet to be demonstrated (53). Additionally, DNMT2 is transcriptionally down-regulated in colorectal, stomach, and hepatocellular cancers (54, 55).

In summary, the present study has identified altered expression of several genes that are novel to testicular tumorigenesis. The increased copy numbers observed at 17q11–q21 in TGCTs are associated with overexpression of GRB7, and in contrast to other tumor types, not with overexpression of the ERBB2 oncogene. In addition, JUP, MYCN, MYCL1, MYBL2, CCND2, and POVI are consistently overexpressed in TGCTs compared with their expressions in nonneoplastic testicular tissue. Furthermore, our data show clear gene expression differences between seminoma and nonseminoma TGCTs. The average expression level of GRB7 was significantly higher in nonseminomas than in seminomas, whereas the expressions of JUP, MYCN, and POVI were highest in seminomas. The putative cancer-related functions of all these genes, in addition to the previously reported overexpression of MYCN and CCND2 in TGCT (40, 46), suggest that the applied cDNA microarrays are sensitive and specific enough to discover oncogenic gene expression changes in TGCT. Thus, the consistently overexpressed ESTs may also be the region with highest score in a genome-wide linkage analysis targeting TGCT susceptibility loci in mice (50, 51). Furthermore, the MXI1 gene is commonly mutated or deleted in prostate carcinomas (52). The DNMT2 gene has strong sequence homology to the DNA- (cytosine-5)-methyltransferases, although its catalytic activity has yet to be demonstrated (53). Additionally, DNMT2 is transcriptionally down-regulated in colorectal, stomach, and hepatocellular cancers (54, 55).

REFERENCES


New Insights into Testicular Germ Cell Tumorigenesis from Gene Expression Profiling

Rolf I. Skotheim, Outi Monni, Spyro Mousses, et al.