Pediatric Malignant Germ Cell Tumors Show Characteristic Transcriptome Profiles

Roger D. Palmer,1,2 Nuno L. Barbosa-Morais,3 Emma L. Gooding,1 Balaji Muralidhar,1 Claire M. Thornton,4 Mark R. Pett,1 Ian Roberts,1 Dominik T. Schneider,3 Natalie Thorne,5 Simon Tavare,6 James C. Nicholson,2 and Nicholas Coleman1

1MRC Cancer Cell Unit, Hutchison/MRC Research Center; 2Department of Paediatric Oncology, Addenbrooke's Hospital; 3Computational Biology Group, Department of Oncology, University of Cambridge, Cambridge Research Institute, Cambridge, United Kingdom; 4Department of Pathology, Institute of Clinical Science, Royal Group of Hospitals, Belfast, United Kingdom; and 5Clinic of Pediatrics, Dortmund, Germany

Abstract

Malignant germ cell tumors (GCT) of childhood are rare and heterogeneous neoplasms thought to arise from primordial germ cells. They vary substantially in their natural history and show important clinical differences from their adult counterparts. To address the biological basis for these observations, we have undertaken a comprehensive analysis of global gene expression patterns in pediatric malignant GCTs and compared these findings with published data on adult testicular GCTs (TGCT). Our study included 27 primary tumors and assessed the principal malignant histologic types of pediatric GCT, yolk sac tumor (YST; n = 18), and seminoma (n = 9). Analysis of Affymetrix U133A GeneChip data was performed using the statistical software environment R, including gene set enrichment analysis, with cross-validation at the RNA and protein level. Unsupervised analysis showed complete separation of YSTs and seminomas by global gene expression profiles and identified a robust set of 657 discriminatory transcripts. There was no segregation of tumors of the same histology arising at different sites or at different ages within the pediatric range. In contrast, there was segregation of pediatric malignant GCTs and adult malignant TGCTs, most notably for the YSTs. The pediatric seminomas were significantly enriched for genes associated with the self-renewing pluripotent phenotype, whereas the pediatric YSTs were significantly enriched for genes associated with a differentiation and proliferation phenotype. We conclude that histologic type is the key discriminator in pediatric malignant GCTs and that the observed clinical differences between malignant GCTs of children and adults are mirrored by significant differences in global gene expression. [Cancer Res 2008;68(11):4239–47]
pediatric malignant GCTs. More recently, limited expression data from cancer-focused arrays of 588 genes were published on GCTs that included five malignant cases from children with ages ≤16 years (13). This suggested that ovarian YSTs of adolescence (and adulthood) differed from ovarian dysgerminomas based on overexpression of genes in the Wnt/β-catenin pathway, a determinant of differentiation in embryonal carcinoma cell development in vitro and in adult malignant TGCTs (14, 15).

We have performed global gene expression analysis on 27 pediatric malignant GCTs. Such large tumor sets could be studied as samples were made available from the entire U.K. bank of pediatric GCTs (as well as cases from Germany). Our primary aim was to compare the two principal histologic types, seminomatous tumors (hereafter termed seminoma, irrespective of site of origin) and YSTs. All samples had previously been analyzed by metaphase comparative genomic hybridization (CGH) to identify genomic copy number imbalances (5). We interrogated our results, including a comparison with published adult TGCT transcriptome data (8) obtained using the same microarray platform, to identify differences in gene expression signature related to clinicopathologic variables, such as histologic type, genomic imbalances, site of origin, and patient age. Here, we present the first comprehensive analysis of the transcriptome of malignant pediatric GCTs, which informs on the pathogenesis of these unique tumors.

### Materials and Methods

**Tumor samples and nomenclature.** Banked frozen tissue was provided by the Children's Cancer and Leukaemia Group (CCLG) from patients managed on U.K. pediatric treatment schedules (i.e., children of ages ≤16 y). Twenty-two samples were from U.K. patients treated on the extracranial GC8901 study, with three additional specimens from the intracranial study. We also obtained two unselected specimens from the German GPOH-MAKEI archive. Our study received Multicenter Research Ethics Committee approval (ref 02/4/071).

By the tumor-node metastasis staging system (1), 18 tumors were stage 1, four tumors were stage 2, two tumors were stage 3, and three tumors were stage 4. Two children died (one from disease, one unrelated to disease) and four relapsed (three of whom had stage 1 disease and received no primary chemotherapy), with follow-up ranging from 12 to 144 mo in disease-free cases. In each case, a histopathologic diagnosis had been provided after CCLG expert panel review. All specimens underwent frozen section analysis, with independent histologic review by three separate histopathologists, which confirmed the CCLG diagnosis. All cases used were completely or predominantly (>90%) composed of a single malignant element.

### Table 1. Clinicopathologic data for the 27 pediatric malignant GCTs analyzed

<table>
<thead>
<tr>
<th>Sample</th>
<th>Histology</th>
<th>Tumor composition</th>
<th>Site</th>
<th>Age (y)</th>
<th>Gender</th>
<th>Stage</th>
<th>Genomic 12p gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Seminoma (Dysgerminoma)</td>
<td>Pure</td>
<td>Ovary</td>
<td>10</td>
<td>Female</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Seminoma (Dysgerminoma)</td>
<td>Pure</td>
<td>Ovary</td>
<td>12</td>
<td>Female</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Seminoma (Dysgerminoma)</td>
<td>Pure</td>
<td>Ovary</td>
<td>12</td>
<td>Female</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Seminoma (Dysgerminoma)</td>
<td>Pure</td>
<td>Ovary</td>
<td>12</td>
<td>Female</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Seminoma (Dysgerminoma)</td>
<td>Pure</td>
<td>Ovary</td>
<td>13</td>
<td>Female</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>Seminoma (Dysgerminoma)</td>
<td>Pure</td>
<td>Ovary</td>
<td>13</td>
<td>Female</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>Seminoma (Germinoma)</td>
<td>Pure</td>
<td>Brain</td>
<td>10</td>
<td>Female</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>Seminoma (Germinoma)</td>
<td>Pure</td>
<td>Brain</td>
<td>16</td>
<td>Male</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>Seminoma (Dysgerminoma)</td>
<td>Within teratoma</td>
<td>Ovary</td>
<td>12</td>
<td>Female</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>0</td>
<td>Male</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>11</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>1</td>
<td>Male</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>12</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>1</td>
<td>Male</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>13</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>2</td>
<td>Male</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>4</td>
<td>Male</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>0</td>
<td>Female</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>16</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>9</td>
<td>Female</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>17</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>12</td>
<td>Female</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>18</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>12</td>
<td>Female</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>19</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>13</td>
<td>Female</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>20</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>13</td>
<td>Female</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>21</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>14</td>
<td>Female</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>22</td>
<td>YST</td>
<td>Pure</td>
<td>Testis</td>
<td>14</td>
<td>Female</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>23</td>
<td>YST</td>
<td>Within teratoma</td>
<td>SCT</td>
<td>2</td>
<td>Female</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>24</td>
<td>YST</td>
<td>Within teratoma</td>
<td>SCT</td>
<td>3</td>
<td>Female</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>25</td>
<td>YST</td>
<td>Within teratoma</td>
<td>Brain</td>
<td>12</td>
<td>Male</td>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>

NOTE: The malignant GCTs called seminomas in this study were ovarian dysgerminomas and extragonadal germinomas. No testicular seminomas were available, as these are extremely rare in the pediatric age range. 12p gain was determined by metaphase CGH, as reported previously (6). For all tumors arising within a teratoma, the sample was confirmed to be completely or predominantly (>90%) composed of the single malignant element on initial frozen section analysis.

Abbreviation: SCT, sacrococcygeal teratoma.

*Contralateral streak ovary noted perioperatively.
The histopathologic nomenclature used in this paper is based on the WHO criteria (16, 17). In particular, the term seminoma refers to all tumors with seminomatous histology, regardless of site (i.e., testicular seminoma, ovarian dysgerminoma, and extragonadal germinoma).

RNA isolation and quality control. Total RNA was extracted from tumor sections using the PolyTron 2100 homogenizer and TRIzol (Invitrogen Life Technologies). Glycoblue (Ambion, Inc.) was added to facilitate RNA visualization. Resultant total RNA was column cleaned using the RNeasy kit (Qiagen), with RNA quantity determined by spectrophotometry ($A_{260}$) and RNA quality using the Bioanalyzer 2100 system (Agilent Technologies). The presence of mRNA was confirmed by gel electrophoresis of reverse transcription–PCR (RT-PCR) products for housekeeping genes TBP and ACTB.

RNA amplification and microarray hybridization. Double-stranded cDNA was synthesized from column cleaned total RNA using SuperScript II (Invitrogen Life Technologies), using the (dT)$_{24}$-T7 promoter primer. Biotin-labeled cRNA was generated by Bioarray in vitro transcription (Enzo) and fragmented by metal-induced hydrolysis (Affymetrix). Fifteen micrograms of each probe was hybridized, washed, stained, and scanned by MRC GeneService using standard Affymetrix procedures with the GeneChip 3000 Scanner (Affymetrix). We used the U133A GeneChip (Affymetrix), composed of 22,283 probe sets corresponding to 13,042 gene names and 13,390 LocusLinks.

Analysis and validation of microarray data. Preprocessing of microarray data (including background correction and normalization) was performed applying the robust multiarray (RMA) method (18), included in the Bioconductor package affy for the statistical software environment R (19). Following the MIAME guidelines (20), raw (.CEL) data files have been deposited in the GEO repository (ref. 21; GEO accession number GSE 10615). The $B$ statistic (22) and prediction analysis of microarrays (PAM; ref. 23) methods were used for identification of significantly differentially expressed genes and gene ranking, using Bioconductor packages limma and pamr, respectively. Gene set enrichment analysis (GSEA; ref. 24) was performed with the Bioconductor package Category.

Real-time quantitative RT-PCR validation. Real-time single-step RT-PCR (Qiagen) was performed on 2.5 ng DNASe–treated (TURBO DNase-free, Ambion, Inc.), column cleaned total RNA from nine of the tumors. QuantiteeT validated primer sets (Qiagen) for nine differentially expressed genes (HERC5, APOA1, POU5F1, FOXA1, CDK2AP1, GPC3, CTSL2, MAN1C1, and FLJ22662) were selected and normalized to three housekeeping genes [glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ACTB, and TBP] using the Opticon DNA Engine (MJ Research, Inc.) and QuantiTeeT SYBR Green RT-PCR kit (Qiagen). Expression relative to 2.5 ng pooled normal human RNA (FirstChoice Human Total RNA Survey Panel, Ambion, Inc.) was calculated, using the Pfaff equation, once primer efficiencies had been determined (25).

Immunohistochemistry. Validation of expression changes at the protein level was performed using immunohistochemistry, as previously described (26). We stained 5 µm sections from a series of 10 unrelated paraffin-embedded cases of pediatric malignant GCTs, provided by the Addenbrooke’s Hospital Tissue Bank. The primary antibodies were mouse monoclonal anti-AP2$\gamma$ (dilution, 1:400), anti-TNNC1 (dilution, 1:250), anti-KRT19 (dilution, 1:500), and rabbit monoclonal anti-MAP4K1 (dilution, 1:50; all Abcam plc).

Results

Characteristics of pediatric GCTs examined. In total, 27 primary malignant GCTs from 27 different patients were analyzed (Table 1). There were 18 female and 9 male patients. Nine tumors...
were seminomas (six dysgerminomas, one dysgerminoma arising within an ovarian teratoma, and two intracranial germinomas), and 18 were YSTs (eight ovarian, seven testicular, two sacrococcygeal, and one intracranial). Twenty-two of the primary tumors were gonadal (15 ovarian and 7 testicular), and five were extragonadal. There was one potential case of gonadal dysgenesis (case 18), wherein a contralateral streak ovary was noted intraoperatively. Seventeen tumors occurred in children with ages \( \geq 5 \) years (14 ovarian, 3 intracranial) and included all nine seminoma cases. Ten tumors occurred in children with ages <5 years at diagnosis (seven testicular, two sacrococcygeal, one ovarian), and all of these were YSTs. This is not surprising, as malignant GCTs of other histologic types are exceedingly rare in this younger age group.

**Gene expression signatures.** Although seminomas and YSTs are presumed to arise from the same progenitor cell, the histologic differences between them were mirrored in their gene expression profiles. Unsupervised analysis of the entire probe set data showed complete separation of the histologies in the cluster dendrogram (Fig. 1A). Principal component analysis (PCA) showed that component 2 was responsible for the segregation by histologic type. We analyzed the data using two techniques (Bayesian linear model and PAM) to derive a final list of transcripts that robustly discriminated between the two tumor types. This list incorporated only transcripts shown to be significantly differentially expressed by both analysis techniques and, for which, there was no overlap in expression levels between the two histologic categories, irrespective of site of origin. The minimum number of probe sets that were defined, using these strict criteria, as discriminating between seminomas and YSTs, was 657 (538 genes), with 266 overrepresented in seminomas relative to YSTs and 391 overrepresented in YSTs relative to seminomas (Fig. 1B). The top 25 transcripts are shown in Fig. 1C. The top-ranked gene ontology processes (27) for the 657 discriminatory transcripts included FZD signaling, Wnt receptor signaling, lipid metabolism, phosphatidylinositol binding, and protease/enzyme inhibitor activity (data not shown).

Clustering algorithms and PCA showed that there was no segregation of tumors of the same histology arising at different sites or at different ages within the pediatric range. When we compared our findings to published data obtained using the same platform from pure histology, malignant TGCTs of adults (10 YST, 12 seminoma; ref. 8; GEO accession number GSE 3218), we observed segregation of pediatric and adult tumors, most conspicuously for the YSTs (Fig. 2A–C). Principal component 2 again separated the samples on the basis of histology, although we note that one adult YST from the previous study (our reference K14) was technically poor and should be removed (see Supplementary Materials and Methods). Another adult YST (K18) did not clearly separate with the other adult YSTs on the basis of principal component 2, instead taking an intermediate position that we have also seen in preliminary work to be occupied by control tissues and teratomas (data not shown). This raises the possibility that the histologic diagnosis for case K18 was incorrect or that inadequate
amounts of the malignant element were present. Principal component 3 seemed to separate the tumors (especially the YSTs) on the basis of age group (pediatric versus adult), with the exception of one adult YST (K20), which intermingled with the pediatric cases (Fig. 2D).

For seminomas, 380 transcripts discriminated between pediatric and adult tumors with a Bayesian linear model B statistic of >0 (Fig. 2B), whereas 1,198 transcripts discriminated YSTs after removal of the two uncertain cases (K18 and K20; Fig. 2C). Sample expression plots for genes significantly overexpressed in pediatric seminomas (CDCA2BPA) or adult seminomas (RAC1, RPS6YA1) and in pediatric YSTs (LARP1, BMP4) or adult YSTs (CAPN7) are shown in Supplementary Fig. S1. Of particular note, childhood testicular YSTs segregated away from the adult testicular YSTs (Fig. 2D), suggesting that site of origin, at least for the YSTs, was not responsible for the differences between the pediatric and adult cases. Similar direct comparison was not possible for the seminomas, as none of the childhood cases were testicular.

To confirm the separation of pediatric YSTs and adult testicular YSTs, all comparable data for the pediatric (n = 18) and adult (n = 8) YSTs were selected. This YST data were RMA normalized independently, and unsupervised clustering again showed segregation into two principal groups (Supplementary Fig. S2A). There were slight changes in the relationships between the YSTs compared with those observed in the overall analysis, with discrimination between pediatric and adult YSTs being improved by this approach (Supplementary Fig. S2A and B). The one adult YST (K20) still intermingled with the pediatric YSTs. Furthermore, there was clustering of all pediatric YSTs, irrespective of site (i.e., including testicular, ovarian, and extragonadal cases) and age (i.e., including cases from patients of ages <5 and ≥5 years; Supplementary Fig. S2B).

Validation of microarray data. Array data were confirmed by real-time quantitative RT-PCR analysis of nine genes (five overexpressed in YSTs and four overexpressed in seminomas), using total RNA from nine samples (five YSTs and four seminomas; Fig. 3). Additionally, overexpression of the proteins encoded by four of these genes (two overexpressed in each histologic grouping) was confirmed by immunohistochemistry using an unrelated series of formalin-fixed, paraffin-embedded pediatric GCTs (five seminomas, five YSTs; Fig. 4).

Chromosome locations of differentially expressed genes. GSEA (24) was performed for all 46 chromosomal arms using the entire gene list and reperformed for transcripts where the B statistic was >0 (3,670 probe sets) to identify any additional enrichment. Seminomas were significantly more likely to express transcripts on 15q, 12p, 8q, and 1p, whereas YSTs were significantly more likely to express transcripts on 11q, 11p, 3p, 5q, 10p, and 13q (Supplementary Fig. S3). Similarly, analysis of transcript location by age revealed that malignant GCTs from children of ages <5 years were enriched for transcripts on 11q, 3p, 11p, and 13q, whereas malignant GCTs from children of ages ≥5 years were enriched for transcripts on 19q, 2p, 1p, 15q, and 12p (Supplementary Fig. S4A). Because none of the cases in patients of ages <5 years were seminomas, the age-related analysis was repeated for the YSTs only. The entire gene list was used in this analysis, as only 61 probe sets that separated YSTs on the basis of age had a B statistic of >0. YSTs in children of ages 5 to 16 years were enriched for transcripts on 19, Xp, Xq, 2q, 7q, and 22q, whereas YSTs in children of ages <5 years were enriched for transcripts on 12q, 2p, 3p, and 18p (Supplementary Fig. S4B).

Of our final list of 538 discriminatory genes, 19 reside on chromosome 12p. Of these, 16 were overexpressed in seminomas, including the highly discriminatory genes FLJ22662, Nanog, and FOXA1.
C1orf35, whereas three were overexpressed in YSTs. The only chromosomallocations of overexpressed genesthatwereexclusive
to any particular histology were 10p (eight genes), 18p (eight genes),
and Y (one gene) for YSTs and 21p (three genes) for seminomas. In
tumors showing 12p copy number gain by metaphase CGH
(compared with tumors without 12p gain; Table 1; ref. 5), there
was enrichment for genes on 12p, as well as on 12q and 1p
(Supplementary Fig. S5).

**Genes overexpressed in pediatric seminomas.** The 266
transcripts overexpressed in pediatric seminomas relative to
pediatric YSTs included the known pluripotency genes
*Nanog* (rank 1), *POU5F1* (*OCT3/4*; ranks 3 and 4), and *UTF*
(*ERF1*; rank 2), an estrogen induced transcription factor
that may stimulate or repress gene transcription and have crucial,
but distinct, functions in embryologic development and malignant
transformation (28); *ASH1L* (*YY1AP1*; rank 5), a putative
transcription factor that maybe involved in chromatin remodeling
and cell-cell adhesion (29); and *PDPN* (*Aggrus*; rank 6), a trans-
membrane glycoprotein involved in cell-cell adhesion and
migration (30). We observed that an increased number of zinc
finger genes were overexpressed in seminomas, consistent with
previous findings (31).

By GSEA of the differentially expressed transcripts, wherein the
*B* statistic was >0, and of the entire gene set, the most significantly
enriched Kyoto Encyclopedia of Genes and Genomes (KEGG)
pathways in seminomas included genetic information processing
(notably ribosome genes), cellular processes (hematopoietic cell
lineage, natural killer cell mediated, antigen processing, T-cell and
B-cell receptor signaling), and signal transduction pathways (Fig. 5
and Supplementary Fig. S6). The latter included Toll-like receptor
and phosphatidylinositol signaling, as well as the Janus-activated
kinase (JAK)/STAT pathway (including *PIAS3* and *PIK3CD*) and the
canonical Wnt/β-catenin pathway (*FZD10, MAP4K1, WNT2B*
Supplementary Fig. S7).

**Genes overexpressed in pediatric YSTs.** The 391 transcripts
overexpressed in pediatric YSTs relative to pediatric seminomas
included the known tumor marker α-fetoprotein (AFP), as well as
differentiation genes, such as *KRT19* (rank 1) and *KRT8* (rank 8).
Other overexpressed transcripts included *APOA1* (ranks 2 and 4)
and *APOA2* (rank 5 and 6), which have a role in lipid transport (32);
*CCKBR* (rank 3), potentially involved in tumor growth (including
leukemia, pancreatic, and gastrointestinal tumors) in response to
gastrin via the phosphatidylinositol-calcium second messenger
system (33); and *PDZK1* (rank 7), involved in ion transport (notably
chloride via the CFTR), lipid and peptide handling, and potentially,
therefore, peptide-like drug resistance (34).

By GSEA, the most significantly enriched KEGG pathways (Fig. 5
and Supplementary Fig. S6) included cellular communication (cell
communication, focal adhesion, adherens junction, and tight
junction) and signaling pathways [e.g., extracellular matrix
receptor interaction and transforming growth factor-β (TGF-β)/
BMP signaling]. Other signaling pathways with overrepresentation
in the YSTs included calcium signaling, the canonical (*FZD2, FZD4,
FZD5, FZD7, SFRP1, ILK, CDH2, TCF7L1, TCF7L2*) and noncanonical
Wnt/β-catenin pathways including members of the planar cell
polarity (PCP) pathway, notably *WNT11, DAAM1, RH0A*, and *B0R2,*
and the Wnt/calcium release pathway, notably *CAMKII*; Supple-
mentalary Fig. S6.

---

**Figure 4.** Immunohistochemical validation of expression of four proteins. Consistent with transcript levels, AP-2γ and MAP4K1 were relatively overexpressed in seminomas, whereas KRT19 and TNNC1 were relatively overexpressed in YSTs. The positive control tissues were (from left to right) normal epidermis, normal hepatocytes, normal cervical squamous epithelium, and normal cardiac muscle.
Discussion

Our data indicate that the key discriminating factor between malignant GCTs arising across the pediatric age range, irrespective of the site of origin, is the histologic diagnosis. We have determined a set of genes responsible for the divergence of YSTs and seminomas that is highly robust, thereby permitting interpretation of function and downstream hypothesis testing with confidence. Adult TGCTs are similarly discriminated by histology, but their expression profiles are far from identical to those of pediatric GCTs (8–11).

Several top-ranked genes that discriminate on the basis of histology in adult malignant TGCTs were associated with the histologic equivalent in the dataset for pediatric malignant GCTs. For instance, top-ranked YST genes in adult TGCT series (8–11) included GPC3, BMP2, and VTN, which were respectively ranked 15, 28, and 36 by the Bayesian linear model in the pediatric YST dataset. Similarly, top-ranked seminoma genes in adult TGCT series included Nanog, POU5F1, and PDPN, which ranked 1, 3, and 6 in the pediatric seminomas (Fig. 1C and Supplementary Fig. S1). Such consistency did not apply to all genes, however. For example, C5 ranked in the top 10 adult YST genes (11), but was lowly ranked (rank 166) in pediatric YSTs. Moreover, highly ranked pediatric YST genes, such as ERBB4 (rank 10) and CD24 (rank 12), did not appear in adult YST discriminatory gene lists.

These observations are reflected in the unsupervised clustering dendrogram (Fig. 2A) and PCA plot (Fig. 2D), which show that malignant GCTs most resemble tumors of the same histologic type, regardless of age, but that there is also separation by age category (i.e., pediatric or adult) within the seminomas and, especially, the YSTs (Fig. 2D). Indeed, we were able to derive gene lists that discriminate between adult and pediatric tumors within each histologic group (Fig. 2B and C).

As expected, immunohistochemical markers currently used for the histopathologic diagnosis of GCTs were represented in our gene expression lists. For example, diagnosis of seminoma commonly involves immunohistochemical detection of OCT3/4 (35) and KIT (CD117; ref. 36), both of which are overexpressed in the seminomas in our dataset (ranks 6 and 1857, respectively, in the pediatric YST versus seminoma list). Diagnosis of YSTs frequently involves immunohistochemical detection of AFP and GPC3 (37), which respectively ranked 125 and 30 in the discriminatory list. These observations support our own immunohistochemical validation data, as presented in Fig. 4.

Our data provide support for clinical evidence, suggesting that pediatric and adult malignant GCTs differ biologically. The pediatric seminomas in our study were all from ovarian or extragonadal sites, and we could not exclude the possibility that differences from adult testicular seminomas may, at least in part, be site related. Our

![Figure 5](https://example.com/figure5.png)

**Figure 5.** KEGG pathways significantly enriched in genes differentially expressed in YSTs and seminomas. The pathways were determined from differentially expressed transcripts identified using the Bayesian linear model (B statistic > 0). A, the quantile-quantile (Q-Q) plot of the actual (sample) values and expected values (based on a theoretical normal distribution) of the statistical score for the 191 KEGG pathways. Enrichment toward seminomas is shown as positive values, whereas enrichment toward YSTs is shown as negative values. The green line passes through the first and third quartiles. The most deviant top score pathways are highlighted in red for seminomas and blue for YSTs. The statistical score and number of genes in each top-ranked pathway are given in B (YSTs) and C (seminomas).
comparison of YSTs alone shows that pediatric and adult cases cluster separately in unsupervised analysis and that all pediatric YSTs group together regardless of tumor site or patient age (Supplementary Fig. S2). In particular, testicular YSTs of adults and children were separated on clustering analysis and PCA. As expected for pediatric YSTs, the testicular cases occurred exclusively in children of ages ≤5 years. However, it is noteworthy that one adult testicular YST clustered with the pediatric cases in unsupervised analysis. The clinical and biological behavior of such adult cases will be of interest. In view of the lack of published global gene expression data for adult ovarian or extragonadal malignant GCTs, it is not possible to determine the timing of the potential biological shift between childhood and adult-type patterns, although this might occur during puberty and be hormone driven.

As CGH profiles in childhood ovarian GCT have been reported to resemble those in adults (38), further study of adolescent tumors is warranted. Moreover, despite having access to the entire United Kingdom archive of pediatric malignant GCTs, plus additional German cases, only five extragonadal tumors were available, and investigation of more examples of such tumors is required.

Differences between adult and pediatric malignant GCTs are supported by the genomic differences between the two groups. Isochromosome 12p is present in 80% of adult TGCTs, and gain on 12p is invariably present in the remaining cases (39). Adult malignant GCTs at other sites (e.g., ovarian and intracranial tumors) seem to show near-identical genetic aberrations (38, 40). However, 12p gain is found with a lower incidence in adolescence and is much less common in infancy (<5 years of age; refs. 4, 5). We observed enrichment for transcripts on 12p in GCTs of children of ages ≥5 years compared with children of ages <5 years (Supplementary Figs. S3–S5). However, this can be attributed to increased expression of 12p genes in seminomas compared with YSTs, as there was no enrichment for 12p genes in YSTs of children of ages ≥5 years compared with children of ages <5 years (Supplementary Fig. S4). Taken together, our findings do not support the claim that 12p gain has a ubiquitous role in the biology of malignant GCTs in children of ages ≥5 years. Indeed, our overall expression data does not clearly support the proposed separation of pediatric malignant GCTs into type I and type II tumors (7) and would also be consistent with an age-related cutoff, as used in previous studies (4).

When compared with YSTs, seminomatous tumors largely recapitulated the features of self-renewing, pluripotent human embryonic stem (hES) cells, which by microarray analysis are characterized by activation of signaling pathways, including the Wnt/β-catenin, fibroblast growth factor, and leukemia inhibitory factor (JAK/STAT) pathways (41), together with up-regulation of cell cycle control genes, DNA repair enzymes, zinc finger-specific transcription repressors, and ribosomal proteins and down-regulation of proapoptotic genes (42). Whereas such similarities have previously been described for adult seminomas (11, 15), we now show that this also applies to pediatric seminomas, irrespective of patient age (at least within the range of 10–16 years) and tumor site.

Relative to seminomas, YSTs overexpressed genes activated during differentiation, including early markers of hES cell differentiation, such as GATA3 and GATA6; keratin genes, such as KRT19, KRT18 and KRT8; and markers of extraembryonic endodermal differentiation, such as VTN, BMP2, and FOXA2. There was also evidence of overexpression of genes in the Wnt/β-catenin pathway, in concordance with previous studies of nonseminomatous malignant GCTs (13, 14). In comparison to the seminomas, we observed overexpression in YSTs of genes from both canonical and noncanonical (PCP and Wnt/calcium release) Wnt/β-catenin pathways, as well as overexpression of RYK, which could potentially cause Wnt activation independently of FZD or act as a coreceptor (43). Whereas the Wnt/β-catenin pathway is essential to maintain the pluripotency phenotype in hES cells (41, 44) and seminomatous tumors (8, 12), overexpression of more Wnt pathway genes in YSTs may lead to differentiation, as well as deregulated growth. Our study is the first to show that overexpression of the Wnt/β-catenin pathway genes occurs in YSTs across the pediatric age spectrum. YSTs also showed enrichment for cell-to-cell interaction and other signaling pathways (Fig. 5), with overexpression of genes in the TGF-β/BMP pathway almost exclusively seen in YSTs. Whereas tight junction genes, notably GJA1 and CLDN6, have previously been reported to be important in defining the stem cell phenotype (44), we observed a significant enrichment for overexpression of tight junction, focal adhesion, and adherens junction genes in YSTs, consistent with the differentiated state of the tumor cells.

From a clinical perspective, the discriminatory genes that we have identified may be of value in the diagnosis or differential diagnosis of GCTs, particularly in small biopsy samples. We did not look at associations with clinical outcome in pediatric malignant GCTs, as these tumors are largely curable with platinum-based combination chemotherapy. Our data provide biological evidence supporting clinical differences between pediatric and adult malignant GCTs. Resistance to cisplatinum chemotherapy is extremely rare in pediatric malignant GCTs, unlike in adults (45). Moreover, whereas less toxic carboplatin chemotherapy is associated with an inferior outcome in adult malignant TGCTs (46), this has not been observed for pediatric malignant GCTs. Our evidence of differences in global gene expression supports the view that chemotherapeutic treatments can be different between children and adults without adversely affecting outcome (1, 2).

Several of our findings also have direct clinical implications. Overexpression of CD99 (MIC2) in YSTs suggests its use as a specific marker of stromal cells (47) when investigating gonadal tumors may be inappropriate. TSPAN1, a possible marker of invasive potential in adult TGCTs (48), was only overexpressed in three pediatric GCTs, all from male patients, raising the possibility of a gender bias to TSPAN expression levels (Supplementary Fig. S8). Cardiac troponin C (TNNT1), which we showed (and confirmed by immunohistochemistry; Fig. 4) to be overexpressed in YSTs, may be a useful adjunct to AFP in the diagnosis and monitoring of GCTs in childhood, particularly during investigation of sacrococcygeal teratomas in infants, where serum AFP levels are typically raised (49). Moreover, as cardiac troponin is released from cells, there may be potential to use it as a serum marker for YST diagnosis in children. Finally, YSTs showed overexpression of both ERBB4 (HER4; rank 16) and ERBB2 (HER2; rank 1952), raising the possibility of heterodimerization of the proteins, which has been described in childhood medulloblastomas (50) and may represent a cell surface therapeutic target.

In conclusion, we have shown that pediatric malignant GCTs differ at the gene expression level on the basis of histology and also differ from their adult TGCT histologic counterparts. Consequently, an understanding of the biology of pediatric malignant GCTs cannot be inferred from adult GCT studies alone. Our findings also support the use of different therapeutic strategies in pediatric compared with adult malignant GCTs, thereby reducing the lifelong morbidity associated with treatment in childhood without affecting overall survival (2).
Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments

References