Down-Regulation of Stem Cell Genes, Including Those in a 200-kb Gene Cluster at 12p13.31, Is Associated with *In vivo* Differentiation of Human Male Germ Cell Tumors

James E. Korkola,[1] Jane Houldsworth,[1,2] Rajendrakumar S.V. Chadalavada,[1] Adam B. Olshen,[3] Debbie Dobrzynski,[2] Victor E. Reuter,[4] George J. Bosl,[2] and R.S.K. Chaganti[1,2]

[1]Cell Biology Program and Departments of [2]Medicine, [3]Epidemiology and Biostatistics, and [4]Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York

Abstract

Adult male germ cell tumors (GCTs) comprise distinct groups: seminomas and nonseminomas, which include pluripotent embryonal carcinomas as well as other histologic subtypes exhibiting various stages of differentiation. Almost all GCTs show 12p gain, but the target genes have not been clearly defined. To identify 12p target genes, we examined Affymetrix (Santa Clara, CA) U133A+B microarray (~83% coverage of 12p genes) expression profiles of 17 seminomas, 84 nonseminoma GCTs, and 5 normal testis samples. Seventy-three genes on 12p were significantly overexpressed, including *GLUT3* and *REA* (overexpressed in all GCTs) and *CCND2* and *FLJ22028* (overexpressed in all GCTs, except choriocarcinomas). We characterized a 200-kb gene cluster at 12p13.31 that exhibited coordinated overexpression in embryonal carcinomas and seminomas, which included the known stem cell genes *NANOG, STELLA*, and *GDF3* and two previously uncharacterized genes. A search for other coordinately regulated genomic clusters of stem cell genes did not reveal any genomic regions similar to that at 12p13.31. Comparison of embryonal carcinoma with seminomas revealed relative overexpression of several stem cell–associated genes in embryonal carcinoma, including several core "stemness" genes (*EBAF, TDGF1*, and *SOX2*) and several downstream targets of WNT, NODAL, and FGF signaling (*FGF4, NODAL*, and *ZFP42*). Our results indicate that 12p gain is a functionally relevant change leading to activation of proliferation and reestablishment/maintenance of stem cell function through activation of key stem cell genes. Furthermore, the differential expression of core stem cell genes may explain the differences in pluripotency between embryonal carcinomas and seminomas. (Cancer Res 2006; 66(2): 820-7)

Introduction

Adult male germ cell tumors (GCTs) are the most commonly diagnosed solid tumors in men ages 18 to 35 years (1). These tumors are thought to originate from transformed germ cells, and most commonly present as a testicular mass, although ~10% present at extragonadal sites (2). GCTs are unique in that they form distinct histologic subtypes that mimic different lineages of the developing human embryo. GCTs can be divided into two main groups based on the degree and type of differentiation (i.e., seminomas, which resemble undifferentiated primitive germ cells, and nonseminomas, which show varying degrees of embryonic and extraembryonic patterns of differentiation; refs. 2, 3). Nonseminomatous GCTs are further subdivided into embryonal carcinomas, which show early zygotic or embryonal-like differentiation, yolk sac tumors and choriocarcinomas, which exhibit extraembryonal forms of differentiation, and teratomas, which show somatic differentiation along multiple lineages (3). Both seminomas and embryonal carcinoma are known to express stem cell markers, such as *POU5F1* (4) and *NANOG* (5).

Cytogenetic analysis has identified 12p gain as the most frequent change in GCTs, occurring in ~100% of cases (6), regardless of histology. Gain of 12p is usually manifested as one or more copies of an isochromosome 12p [i(12p); ref. 7], with tandem duplication of 12p material occurring at a lower frequency (8). We and others have also observed the presence of high-level subregional amplifications of 12p (9–11). Given the frequency of 12p gain in all types of GCT and the presence of extra copies of 12p in some premalignant intratubular germ cell neoplasia (ITGCN; refs. 12, 13), it has been suggested that 12p gain may be involved in the early stages of germ cell tumorigenesis or may perhaps even be the initiating event (14). In contrast, others have reported that 12p gain does not occur frequently in pure ITGCN and instead is associated with progression (15, 16). Regardless, gain of 12p remains a genetic hallmark of adult male GCTs.

The 12p region has been studied extensively, and we have proposed that the proto-oncogene *CCND2* may be (one of) the targets of 12p gain (17). Other genes on 12p are also known to be overexpressed in GCT, including genes associated with stem cells (18, 19), but their importance in tumorigenesis is not well established. Interestingly, the formation of i(12p) in cultured human embryonic stem cells has been reported recently, which may give proliferative or self-renewal advantages to these cells (20). Amplification of 12p is commonly found in other human neoplasms, including esophageal (21) and ovarian (22) carcinomas, indicating that 12p target genes may have broader implications in oncogenesis.

We sought to identify target genes on 12p and infer their significance in germ cell tumorigenesis and to identify genes that are associated with the pluripotent phenotype of embryonal carcinoma tumors. We recently reported expression profiling on a group of 84 nonseminomatous GCT samples that were stratified into distinct histologic subtypes according to an expression-based classifier (23). We have analyzed these samples along with 17 seminoma samples to identify differentially expressed genes on 12p. We identified a 200-kb coordinately regulated region at 12p13.31 that was overexpressed in embryonal carcinomas and seminomas, which consisted of several stem cell–associated genes. No other

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: R.S.K. Chaganti, Cell Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021. Phone: 212-639-8121; Fax: 212-717-3541; E-mail: chagantr@mskcc.org.

doi:10.1158/0008-5472.CAN-05-2445

coordinately regulated genomic regions of stem cell gene clusters were evident. Comparison of embryonal carcinomas to seminomas identified overexpression of several core stem cell genes in embryonal carcinoma and differential activation of WNT, NODAL, BMP, and FGF pathways between the subtypes. These results suggest that 12p gain may play dual roles in GCT tumorigenesis: establishment or maintenance of a stem cell phenotype through gain of a cluster of 12p13.31 genes and conferring a proliferative advantage. Our results further suggest that differential expression of core stem cell genes may account for the differences in pluripotency of embryonal carcinoma and seminoma tumors.

Materials and Methods

Tumor specimens. Tumor tissues were collected under institutional review board–approved protocols at the Memorial Sloan-Kettering Cancer Center (New York, NY) between 1987 and 1999. The tumor samples consisted of 17 seminomas, 15 pure embryonal carcinomas, 15 pure teratomas, 10 pure yolk sac tumors, 2 pure choriocarcinomas, and 42 nonseminomatous GCTs with mixed histologies. Profiling of the nonseminomatous GCT specimens was reported recently (23). Using the histologic classifier that we developed recently (23), the mixed nonseminomatous GCT specimens were classified according to their predominant histology. For controls, five normal testis specimens from similarly aged patients with no evidence of testicular cancer were used. A pooled sample of all five normal specimens was also profiled and included in figures but was not used in identifying differentially expressed genes.

Cell culture. The pluripotent embryonal carcinoma cell lines 27X-1 and NT2/D1 were maintained as described previously (24). The 27X-1 cells were exposed to 10 μmol/L all-*trans*-retinoic acid (RA; Sigma, St. Louis, MO) to induce differentiation. Cells were reseeded every 2 to 3 days. Cells were harvested at 0, 2, 4, 8, 12, 18, and 24 hours and 3, 6, 8, and 10 days after treatment for isolation of RNA using RNeasy columns (Qiagen, Valencia, CA) according to the manufacturer's protocol. The time course was done in triplicate.

RNA isolation from tumor tissues. Fresh tumor tissue was grossly dissected to remove normal contamination with no further microdissection. Tumor tissue was homogenized using a Brinkman homogenizer, and RNA was isolated using RNeasy Midi columns (Qiagen) according to the manufacturer's protocol. RNA was quantified by UV absorbance and quality was verified by A_{260}/A_{280} ratios and integrity of 28S and 18S bands on denaturing agarose gels. Samples that showed evidence of DNA contamination were digested with RNase-free DNase (Qiagen) followed by purification through RNeasy Mini columns (Qiagen).

cDNA synthesis, labeling, and hybridization. The cDNA synthesis and cRNA labeling for hybridization to U133A and U133B microarrays (Affymetrix, Santa Clara, CA) was done as described previously (23, 24). The size of the labeled cRNA was evaluated on an Agilent Bioanalyzer (Agilent, Foster City, CA) to ensure probe quality for hybridization. Hybridization, washing, and imaging were done as described previously (23).

Reverse transcription-PCR expression analysis. Total RNA (1 μg) was reverse transcribed using 300 ng random hexamers (Invitrogen, Carlsbad, CA), deoxynucleotide triphosphates at a final concentration of 500 μmol/L, and 100 units SuperScript II reverse transcriptase (Invitrogen) at 42°C. The genes tested were *NANOG, STELLA*, AY151139 (the putative homeobox C14 gene, *HOXC14*), *CLECSF7, GDF3*, FLJ40451 (*NANOG2*), Hs.129302, *EBAF, TDGF1, SOX2, SOX17, GIP*, and *CD24*, with *ACTB* as a control. PCR was carried out with 30 pmol of each primer in a total volume of 30 μL using 20 to 50 ng of target cDNA using 2.5 units *Pfu*I turbo polymerase (Stratagene, La Jolla, CA) or 1 unit HotStarTaq (Qiagen). The product was run on 1.4% agarose gels and stained with ethidium bromide for visualization. Primers were designed to span introns to distinguish between mRNA and DNA for all genes, except Hs.129302, which has no introns, and *CLECSF7*. Primer sequences and PCR conditions are available as Supplementary Table S1.

Immunohistochemistry. Formalin-fixed, paraffin-embedded mixed GCT specimens were sectioned at 4 μm, baked on Superfrost slides for 30 minutes at 60°C, deparaffinized using xylenes, and then rehydrated through an ethanol series. For SOX2 staining, the sections were pretreated with 3% hydrogen peroxide and blocked using conventional methods. SOX2 antibody (R&D Systems, Minneapolis, MN; diluted 1:100) was incubated on the section overnight at 4°C. Secondary antibody and streptavidin were applied each for 1 hour at room temperature. 3,3′-Diaminobenzidine tetrahydrochloride (Sigma) was used as the chromagen for detection with hematoxylin counterstain. To detect GDF3, the same protocol was used, except that sections were pretreated with trypsin for 30 minutes at 37°C. Primary antibody for GDF3 staining was a goat polyclonal antibody (Santa Cruz Biotechnology, Santa Cruz, CA; diluted 1:100).

Data analysis. Background subtraction, normalization, and $\log_2$ transformation of the array data were done using the Robust Multi-Array Average (RMA) method (25) within the Bioconductor for R software package (26). Genes were annotated according to the June 23, 2004 Affymetrix annotation. In accordance with Minimum Information About a Microarray Experiment guidelines (27), both RMA-processed and raw data files (.CEL and .EXP) for all 101 GCTs and the normal testis specimens have been deposited in the Gene Expression Omnibus public repository (GEO accession no. GSE 3218).[5] To identify significantly differentially expressed transcripts, the maxT function in multtest in bioconductor for R was used. This is a permutation-based method to adjust P values to correct for multiple comparisons (28). We compared expression in each histologic subtype to the five normal testis samples and used 10,000 permutations within maxT (adjusted $P < 0.05$ is considered significant). Fisher's exact test was used to test whether overexpressed clones from 12p were overrepresented compared with clones from other chromosome arms. For validation of stem cell–associated genes in the 27X-1 cell lines, a one-sided paired t test was used to test for differences in expression between the baseline time point and 10-day time point. Visualization of the data was done using the Cluster/Treeview software package (29).

Results

Identification of target genes on 12p. We have recently published results from expression profiling of 84 nonseminomatous GCT tumor specimens (23). We analyzed these specimens along with 17 new seminoma profiles to identify differentially expressed genes between tumors of each of the histologic subtypes and normal testes specimens. There were 6,038 transcripts that showed a significant change in expression in at least one histologic subtype relative to normal, with 2,613 of those transcripts being overexpressed (Supplementary Table S2). Seminomas displayed the most changes, whereas choriocarcinoma had the least. Specifically for 12p, there were 107 transcripts (89 genes) that showed a significant change in at least one of the tumor types relative to normal testis, with 86 transcripts being overexpressed. We currently estimate that there are ~400 genes on 12p, consistent with our previous estimates (9). We identified 555 transcripts representing 331 genes on the U133A+B microarrays for an estimated coverage of 83%. The complete list of transcripts from 12p showing significantly different expression is included as Supplementary Table S3. Although genes from other chromosomes clearly play a role in GCT etiology and progression, the importance of the gain of 12p is further exemplified by the overrepresentation of overexpressed transcripts from 12p compared with overexpressed transcripts from all other chromosomes by Fisher's exact test ($P < 0.001$), consistent with previous findings (30).

There were 21 transcripts (16 genes) on 12p that were significantly down-regulated in at least one of the histologic subtypes relative to normal testes. These included several genes

[5] http://www.ncbi.nlm.gov/geo/.

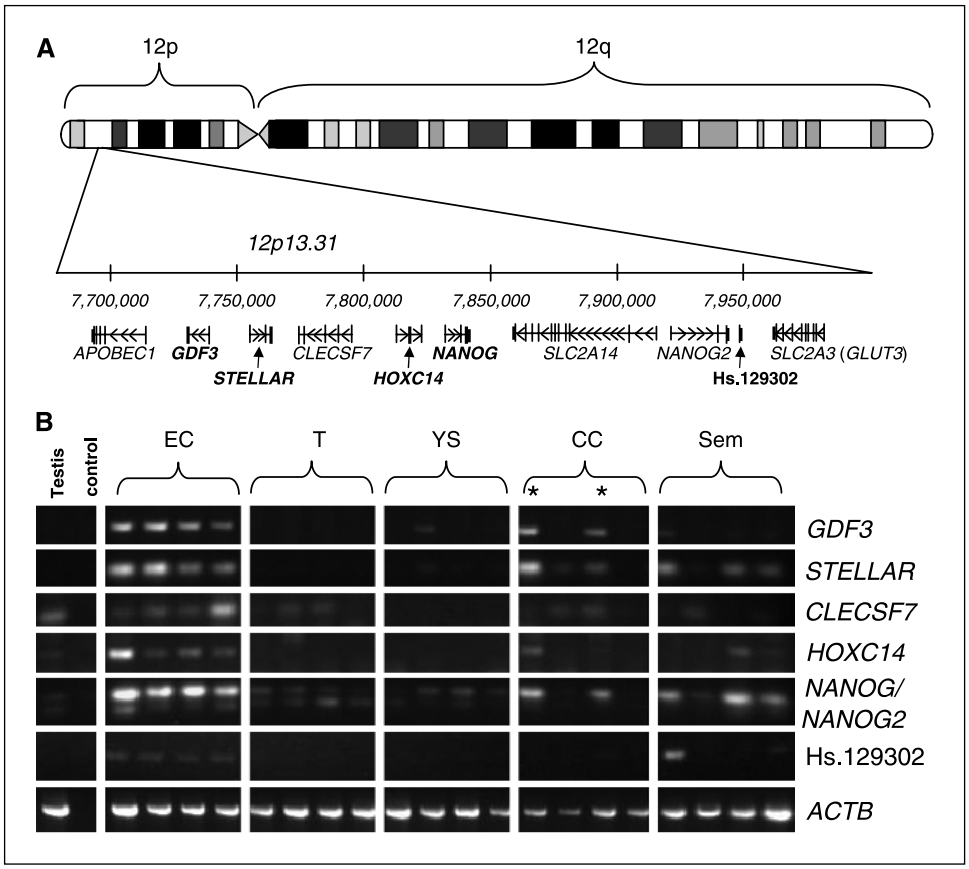

Figure 1. *A,* genomic organization of chromosome 12, with details of a 200-kb region of 12p13.31 (adapted from http://genome.ucsc.edu). Genes in bold indicate those that showed a stem cell–like pattern of expression by expression microarray and/or RT-PCR. *B,* RT-PCR analysis of 12p13.31 genes in pure GCT samples showing differential gene expression in undifferentiated (embryonal carcinomas and seminomas) versus differentiated (teratoma, yolk sac tumor, and choriocarcinoma) tumors. *Asterisks,* two mixed tumor samples that were classified as choriocarcinoma samples, which contained embryonal carcinoma components. *NANOG* and *NANOG2* were detected with the same set of primers, with *NANOG* presenting a 394-bp band (*top*) and *NANOG2* presenting a 346-bp band (*bottom*). *EC,* embryonal carcinoma; *T,* teratoma; *YS,* yolk sac tumor; *CC,* choriocarcinoma; *Sem,* seminoma.

related to meiosis, sperm function, and/or spermatogenesis, such as *PLCZ1, CAPZA3, GSG1, SOX5*, and *ACRBP*, which had significantly lower levels of expression in all GCTs relative to normal. Other down-regulated genes included *H2AJ* and *C12ORF2*. There were also some genes that showed lower levels of expression in only a subset of the tumor types, including the uncharacterized IMAGE clone 5295478 (down-regulated in seminomas) and *CSDA*, which was significant only in yolk sac tumor specimens.

The remaining 86 significant transcripts (73 genes) on 12p were overexpressed in at least one of the tumor subtypes compared with normal testis specimens. There were two genes that were up-regulated in all tumor subtypes regardless of histology. These were *SLC2A3* (*GLUT3*), which has been described previously in GCTs (19), and *REA*. Four genes showed significantly higher expression in all but one histologic subtype, including the previously identified *CCND2* (refs. 17, 19; which was high in all but choriocarcinoma), *FLJ22028* (high in all but choriocarcinoma), and the metabolic genes *GAPDH* and *TPI1* (which were high in all subtypes, except teratoma). Eight genes were overexpressed in three subtypes, including *CD9*, *FLJ22662* (high in all but teratoma and yolk sac tumors) and *TEAD4, LRP, TERA,* and *PTPN6* (high in all but teratoma and choriocarcinoma). The remaining genes were overexpressed in two or fewer histologic subtypes.

Particularly interesting was the overexpression of 14 transcripts on 12p seen only in embryonal carcinomas and seminomas, including several genes implicated in stem cell and GCT biology (5, 30). These included *AK3* and *GABARAPL1*. Most importantly, there were two previously identified stem cell–associated genes that mapped within 100 kb of each other on 12p13.31 [i.e., *STELLA* (*DPPA3*) and *NANOG*]. A previously unreported gene, Hs.129302, located close to *STELLA* and *NANOG* (Fig. 1*A*), also displayed comparable overexpression in embryonal carcinomas and seminomas, confirming the existence of a coordinately regulated region of stem cell–associated genes that had been suggested previously (18). Further examination revealed that the pluripotency-associated gene *GDF3* (18), which also maps to this region, was overexpressed only in embryonal carcinoma specimens (Fig. 1*A*). In contrast, *SLC2A14*, a gene that lies between *NANOG* and Hs.129302 within this region, was highly expressed in all tumor samples (data not shown). The coordinately regulated 200-kb gene cluster seemed to be flanked terminally by *APOBEC1* and proximally by *SLC2A3* and *FOXJ2* (not shown in Fig. 1*A*), none of which were overexpressed only in embryonal carcinomas and seminomas.

Detailed mapping of the region from the May 2004 freeze[6] showed that several other genes that were not on the microarrays also mapped to this 200-kb region, including AY151139 (*HOXC14*), a C-type lectin (*CLECSF7*), and a *NANOG* pseudogene FLJ40451 (*NANOG2*) as shown in Fig. 1*A*. We did reverse transcription-PCR (RT-PCR) analysis on a set of tumors consisting of four pure specimens of each of embryonal carcinoma, teratoma, yolk sac tumor, and seminoma and four choriocarcinoma specimens (two pure, two mixed) as an internal validation of the expression of genes from the array that mapped between *GDF3* and Hs.129302 (excluding *SLC2A14*) and to test expression for the three genes that were not included on the arrays. *NANOG2* and *CLECSF7* were not

[6] http://genome.ucsc.edu.

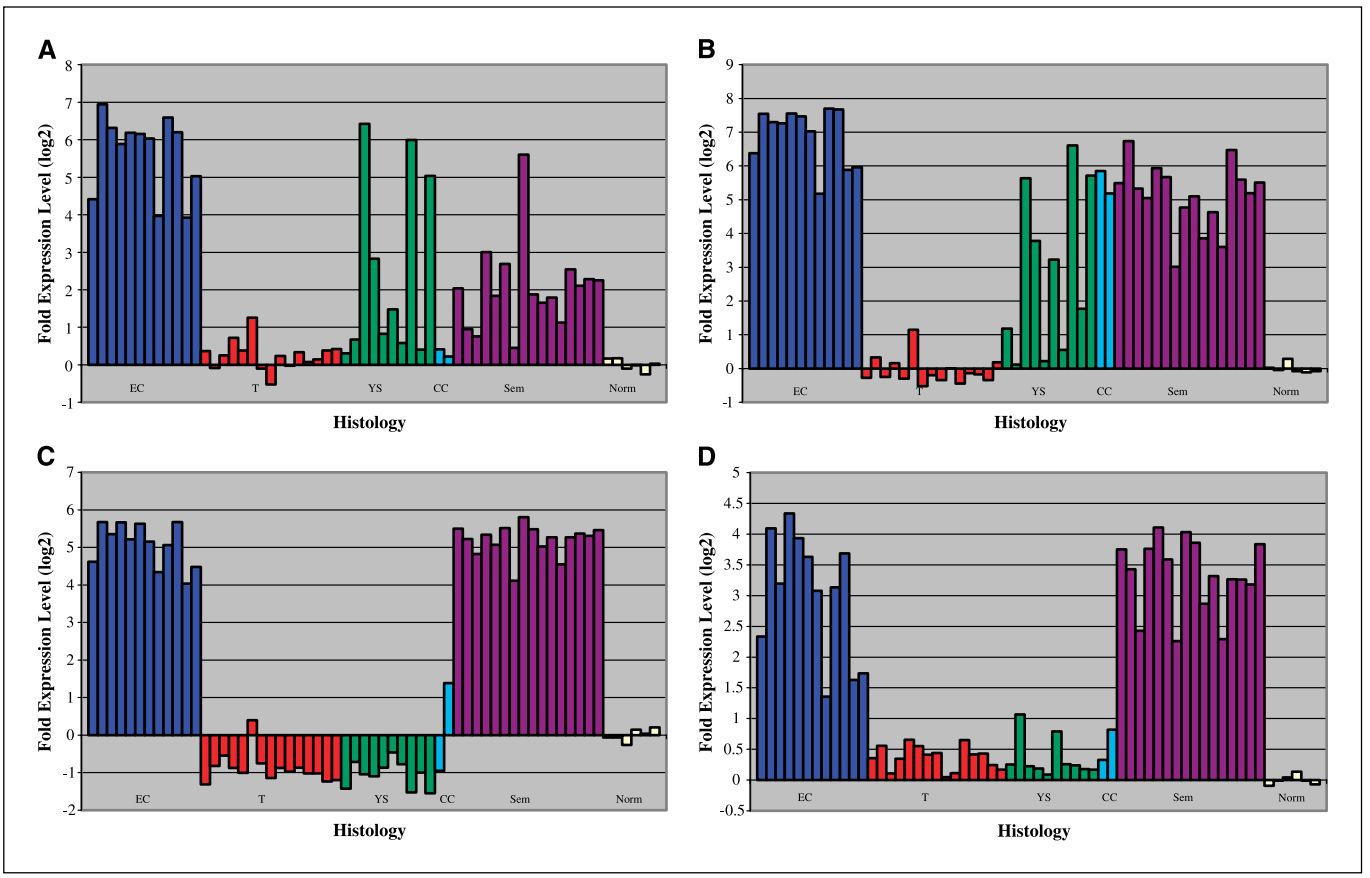

Figure 2. Expression of stem cell–related genes in pure nonseminomatous GCT and seminoma specimens. Gene expression levels (log_2) are shown relative to average normal testes in pure embryonal carcinoma, teratoma, yolk sac tumor, choriocarcinoma, seminoma, and normal testis (*Norm*) specimens for 12p13.31 stem cell–related genes (*A*) *GDF3*, (*B*) *STELLA*, (*C*) *NANOG*, and (*D*) Hs.129302.

highly expressed in the samples (Fig. 1*B*). *STELLA, NANOG*, and Hs.129302 shared a similar expression pattern: high in embryonal carcinomas and seminomas and low in teratoma, yolk sac tumors, and pure choriocarcinoma. *GDF3* and *HOXC14* exhibited high levels of expression primarily in the embryonal carcinoma samples, with some seminoma specimens showing detectable expression. The expression levels from the microarray data for *GDF3, STELLA, NANOG*, and Hs.129302 in pure embryonal carcinoma, teratoma, yolk sac tumor, choriocarcinoma, and seminoma samples relative to normal testes are shown in Fig. 2.

Expression patterns of additional genes associated with stem cell phenotypes. Given the expression of known stem cell markers, such as *POU5F1*, within embryonal carcinomas and seminomas and the described cluster of genes on 12p with a similar expression pattern, we sought to evaluate all other chromosomes to determine if any additional stem cell–associated clusters existed. For the initial analysis, transcripts that were strongly correlated with the core stem cell gene *POU5F1* were selected. We found 41 such transcripts (36 genes), all with a correlation value of $\geq$0.80 (Supplementary Table S4). Of these 41 transcripts, all but 3 were also found to be significantly differentially expressed between embryonal carcinomas or seminomas and normal testes specimens, including all of the 12p genes. Several of the genes were reported previously to be targets of *POU5F1*, including *UPP1, UTF1, REX1* (*ZFP42*), and *NANOG* (31, 32). Considering chromosomal location, 12p was the most highly represented and, not surprisingly, included three of the genes described above (*STELLA, NANOG*, and Hs.129302) plus *TEAD4* and *PHC1*, which were located ~5.0 and ~1.0 Mb away from the 12p13.31 cluster, respectively. In addition, four genes on 10q shared this expression pattern, consisting of *SLC25A16, FRAT2, UTF1*, and *VENTX2*. These last two genes were located close to one another at 10q26.3, separated by 6.6 kb. We also observed three genes that shared the stem cell–like expression pattern on each of 1p and 7q, but none of these genes were located close to one another on the chromosomes. Genes representative of this type of expression pattern are shown in Supplementary Fig. S1.

To further validate if these genes behaved in a manner consistent with stem cell function, we examined the changes in their expression in the pluripotent 27X-1 embryonal carcinoma cell line on induction of differentiation with RA along a parietal yolk sac tumor lineage. Only 8 of the 36 genes that correlated with *POU5F1* did not show significant decreases in expression on treatment with the morphogen, consistent with the hypothesis that they may be involved in stem cell maintenance or markers of an undifferentiated state (Fig. 3). One of these 8 genes, *PYPAF3* (*NALP7*), showed a significant increase in expression following differentiation. The remaining 7 genes were not significantly altered, including the known stem cell markers *ZFP42* and *STELLA*. Our results confirm that a large number of stem cell–associated genes are highly expressed in embryonal carcinomas and seminomas as reported previously (33).

Differential gene expression in embryonal carcinoma versus seminomas. To identify genes that account for the differences in pluripotency between embryonal carcinomas and seminomas, we examined these specimens for differentially expressed genes. There were 777 transcripts (~660 genes) that were differentially expressed between the two tumor types by maxT analysis. We focused on the 31 transcripts that showed ≥4-fold expression in seminomas compared with embryonal carcinoma samples, and the 40 and 120 transcripts that were expressed at least 4- or 2-fold higher, respectively, in embryonal carcinomas compared with seminomas. The 25 most differentially expressed transcripts for each tumor class are shown in Supplementary Table S5A and B. Several of these genes have not been identified in previous screens of embryonal carcinomas compared with seminomas. Compared with embryonal carcinomas, seminomas showed high levels of expression of *GIP, TCL1A*, and *XIST*. Interestingly, among the genes more highly expressed in embryonal carcinomas compared with seminomas were *SOX2*, which has been reported previously (30), as well as *TDGF1* and *FGF4*. All of these genes have been identified as being essential genes that define "stemness" (34). We examined the expression levels of these three genes along with the other four defining markers of "stemness" (*ZFP42, POU5F1, EBAF*, and *THY1*) in embryonal carcinomas and seminomas and found that all were highly expressed in embryonal carcinoma but only *ZFP42, POU5F1*, and *THY1* showed high levels of expression in seminomas (Supplementary Fig. S2). We validated several of the differentially expressed genes (*EBAF, SOX2, TDGF1, CD24, GIP, SOX17*, and *KIT*) by RT-PCR analysis in a subset of the tumors that had been used for expression profiling. As expected, all genes showed an expression pattern that was consistent with the data from the microarray analysis (Supplementary Fig. S3). We also did immunohistochemical staining for GDF3 and SOX2 protein in ITGCN, seminoma, and embryonal carcinoma specimens that were not used for expression profiling. As expected, SOX2 was highly expressed in the nucleus of embryonal carcinoma samples (Fig. 4*C*) but not in seminomas (Fig. 4*B*) or ITGCN (Fig. 4*A*). Interestingly, SOX2 was expressed in *in situ* embryonal carcinoma lesions (data not shown). GDF3 showed strong cytoplasmic staining in embryonal carcinoma, with weaker staining in seminomas and ITGCN (Supplementary Fig. S4A-C). There was also some evidence of staining of spermatogonia but not in Sertoli cells in these specimens.

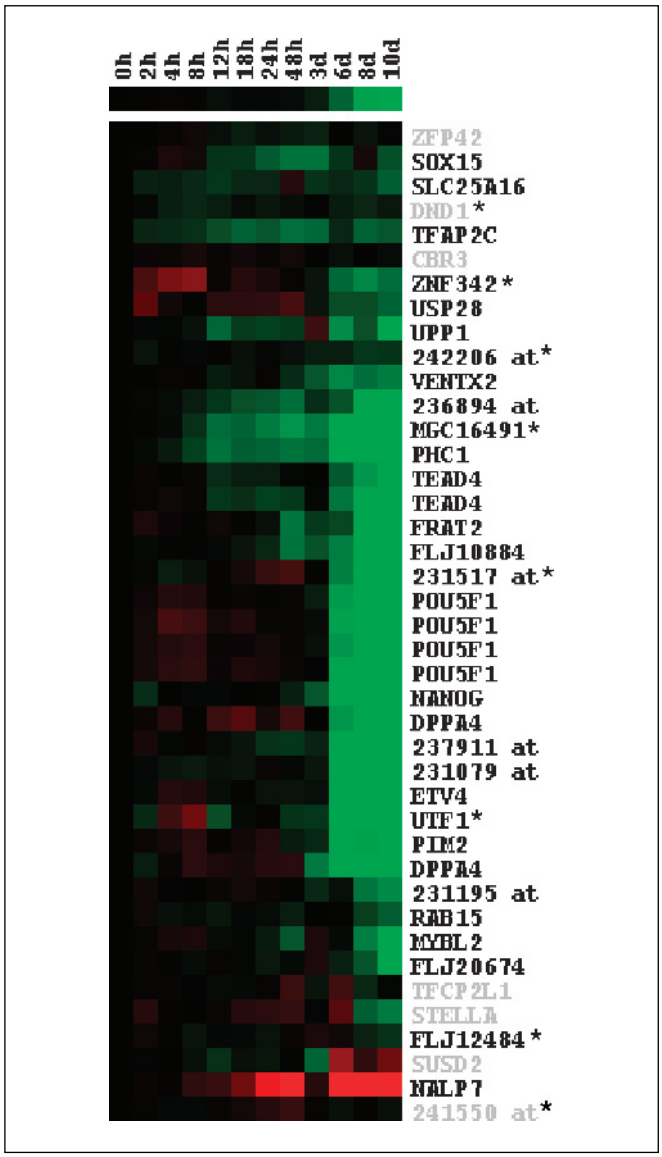

Figure 3. Heat map showing expression levels of genes that were highly correlated with *POU5F1* at various times following RA treatment in 27X-1 cells. *Green,* low relative expression; *red,* high relative expression. *Gray,* genes that were not significantly differentially expressed in embryonal carcinoma relative to testes by maxT analysis; *asterisks,* genes that were not significantly differentially expressed at 10 days compared with time 0 in 27X-1 cells by *t* test.

Activation of WNT signaling in embryonal carcinoma versus seminomas. Many of the differentially expressed genes between embryonal carcinomas and seminomas have been implicated in developmental signaling pathways (Fig. 5). We observed differential overexpression of genes involved in NODAL, FGF, BMP, and WNT signaling between embryonal carcinoma and seminoma samples. In particular, the WNT signaling pathway was highly represented as reported previously (30, 35). In embryonal carcinoma, *WNT5A* and *FZD7* were overexpressed relative to seminomas as reported previously (30) as well as the TCF-like transcription factor *TCF7L* and several previously identified downstream targets of WNT signaling, including versican (36), connexin 43 (36), *JUN* (37), and *FGF4* (38). In contrast, seminomas showed overexpression of a different WNT ligand (*WNT2B*) and several known downstream WNT targets, including the homeobox gene *CDX1* (39) and *NRCAM* (40). We also did staining for β-catenin, which is normally translocated to the nucleus on activation of WNT signaling. This revealed intense staining in the membrane and some staining in the cytoplasm of embryonal carcinoma specimens but no nuclear staining (Fig. 6*B*), whereas in seminomas staining was limited to the membranes (Fig. 6*A*). We also examined β-catenin localization in the pluripotent embryonal carcinoma cell line NT2/D1 by immunocytochemistry, which revealed that staining was limited to the membranes and cytoplasm (data not shown).

Discussion

GCTs are characterized by gain of 12p in ~100% of cases (41). Gain of 12p seems to occur early during GCT tumorigenesis, as some ITGCN lesions show evidence of 12p gain (13) and overexpression of 12p genes, such as *NANOG* (5). However, whether gain is associated with initiation or progression is still unclear. Several candidate genes on 12p have been suggested, but

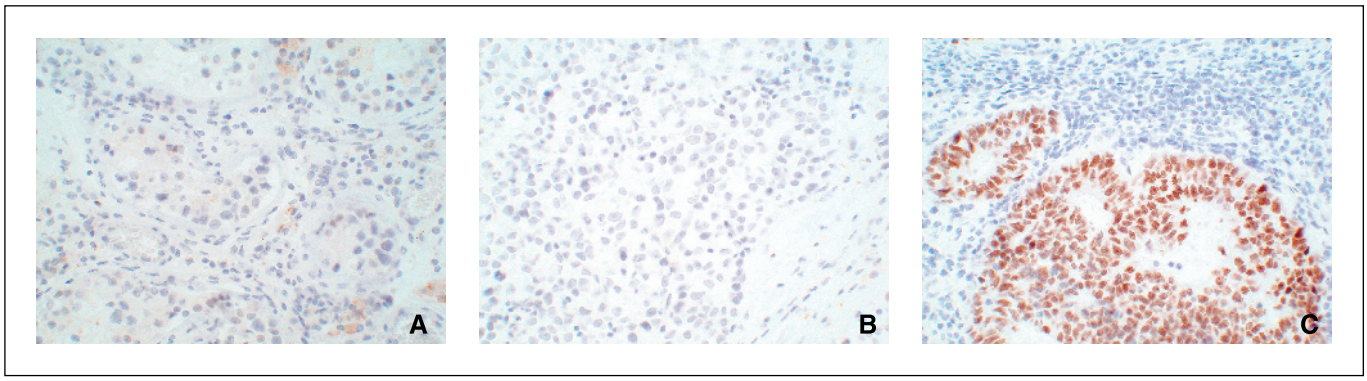

Figure 4. Immunohistochemical analysis of SOX2 expression in embryonal carcinoma, seminoma, and ITGCN. Nuclear SOX2 staining in (*A*) ITGCN, (*B*) seminoma, and (*C*) embryonal carcinoma.

these have not been based on comprehensive expression screens. We analyzed expression profiles from a panel of 101 GCT specimens and 5 normal testis samples using Affymetrix U133A+B microarrays, which provide ~83% coverage of 12p transcripts. We identified differentially expressed genes within histologic subtypes of GCT, including 107 transcripts on 12p. Several genes were altered in all or all but one histologic subtypes. These included *CCND2*, which we have identified previously as potentially important in GCT etiology by reinitiating cell cycle and providing a proliferative advantage for the tumor (2). The *REA* (repressor of estrogen activation) gene, which regulates transcriptional activity, was also highly expressed in all tumor subtypes. The *GLUT3* glucose transporter was highly expressed in all tumor types relative to normal testis as reported previously (19). This gene also maps to the 12p13.31 region that harbors the stem cell cluster that we have defined but does not share the stem cell–related expression pattern of the other genes in this region. Perhaps most interesting is the overexpression in three tumor types (excluding yolk sac tumor and teratoma) of *CD9*, a tetraspanin receptor. Tetraspanin proteins have been implicated in cell adhesion, migration, and proliferation (42). *CD9* has been implicated in the maintenance of undifferentiated embryonic stem cells in the mouse (43). Furthermore, the murine form of *CD9* is the receptor for pregnancy-specific glycoprotein 17 (44). We have shown recently that high levels of expression of pregnancy-specific glycoproteins are predictive of choriocarcinoma histology (23). At this time, it is unclear as to what role many of these genes play in GCT tumorigenesis.

We observed that several genes that map to 12p13.31 and have been implicated in stem cell maintenance (18) were coordinately overexpressed in the seminomas and/or pluripotent embryonal carcinoma samples. These included *GDF3, STELLA*, Hs.129302 (a potential pseudogene designated as *FEENEE* by AceView),[7] and the homeobox genes *HOXC14* and *NANOG*. A search for other coordinately regulated stem cell–associated regions failed to find any other such genomic regions. However, we did observe that the stem cell gene *UTF1* and *VENTX2* on 10q26.3, separated by only 6.6 kb, shared this expression pattern. Interestingly, *VENTX2* is also a homeobox gene, which have been found to cluster together in the genome and be coordinately expressed (45, 46) as we observed on 12p13.31. Other than *UTF1*, the closest proximal or distal genes to *VENTX2* that were present on the array did not show a stem cell pattern of expression. Furthermore, this region has not been reported to be frequently gained in GCTs, suggesting that there is not a coordinately regulated cluster of stem cell genes at 10q26.3. Our results thus confirm the presence of a 200-kb region on 12p that is coordinately regulated in pluripotent GCTs, such as embryonal carcinoma, and identifies two additional genes (Hs.129302 and *HOXC14*) that map to this region.

Our results also suggest that the stem cell–associated genes within the 12p13.31 region may be coregulated to maintain expression only in "stem cell–like" cells. Elevated expression of genes mapped to this region have been reported previously in seminoma and embryonal carcinoma (30) and in cultured human embryonic stem cells (20). We have further observed that these and other stem cell–related genes are down-regulated in response to differentiation in cultured embryonal carcinoma cells and that expression levels are lower in the differentiated yolk sac tumor, choriocarcinoma, and teratoma samples. Together, these results suggest multiple roles for the selection for 12p gain during GCT tumorigenesis (i.e., providing a proliferative advantage and reestablishment and/or maintenance of a stem cell phenotype). This would be most consistent with 12p gain occurring early, perhaps at tumor initiation. In support of this notion, i(12p) has been detected in ITGCN (13), the precursor to invasive GCT. Furthermore, overexpression of genes on 12p, such as *NANOG* and *TEAD4*, have been observed in ITGCN lesions (5).

Based largely on immunohistochemical and morphologic criteria, seminomas have been suggested to resemble undifferentiated germ cells. In contrast, embryonal carcinoma resembles pluripotent embryonic stem cells and has the capacity to differentiate. In support of this notion, another study identified genes that are differentially expressed between the two tumor types and, based on differential gene expression, postulated that seminomas more closely resemble primordial germ cells, whereas embryonal carcinoma represents a more primitive ectoderm-like cell (30). A recent study also examined differentially expressed genes in embryonal carcinomas and seminomas and identified 169 genes (35), of which 56 transcripts (40 genes) overlapped with genes identified in our study, including ones overexpressed in seminomas [*KIT*, which was known previously (47), *GAL, CPEB1*, and *SOX17*] and genes overexpressed in embryonal carcinoma (*BCAT1, GDF3*, and *DNMT3B*). Our results indicated that embryonal carcinoma specimens show expression of genes that

[7] http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/.

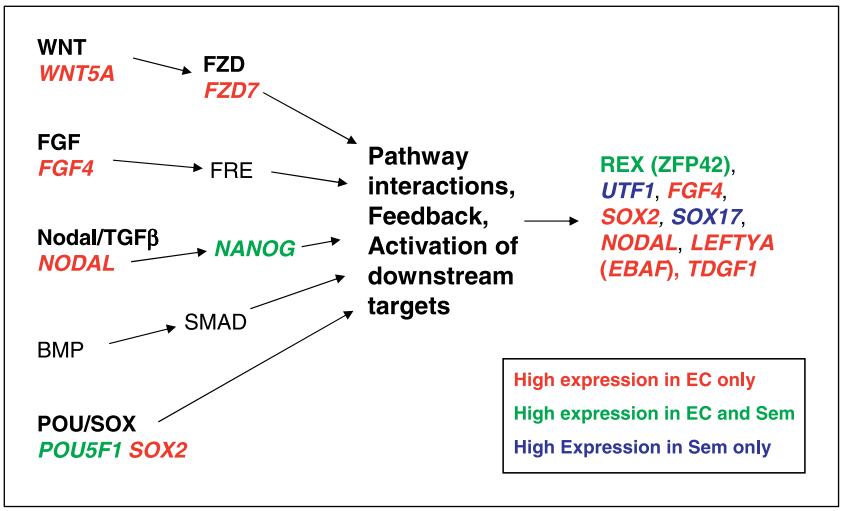

Figure 5. Signaling pathways active in stem cells that show differential expression between embryonal carcinoma and seminoma.

have been implicated in stem cell function to a much greater extent than seminoma samples. Indeed, of the seven genes that have been proposed as unequivocal markers of human embryonic stem cells (34), only *POU5F1, THY1*, and *ZFP42* were highly expressed in seminoma samples, whereas all seven were expressed in embryonal carcinoma samples. The four additional genes expressed in embryonal carcinoma only (*TDGF1, EBAF, SOX2*, and *FGF4*) have all been implicated as targets of WNT, TGFB, FGF, BMP, and NODAL signaling pathways during embryogenesis (48), suggesting that these signaling pathways may play an important role in germ cell tumorigenesis as well. *EBAF, TDGFI*, and *STELLA* have all recently been identified as genes that are highly expressed in nonseminoma GCTs (49), although our results indicate that expression of these genes is limited to embryonal carcinoma samples. WNT signaling has been suggested previously to explain the differences between embryonal carcinoma and seminoma (30), although in the present study we saw evidence of WNT signaling in both, albeit with different ligands and downstream targets. β-Catenin, which is normally found in cell membranes, can be translocated to the nucleus on activation of canonical WNT signaling and thus is a marker for active WNT signaling. Staining for nuclear β-catenin was negative in both embryonal carcinoma tumors and cell lines. However, in the pluripotent murine teratocarcinoma cell line F9, RA-induced differentiation results in stabilization of β-catenin (50). Furthermore, in this study, it was found that inhibition of WNT signaling by introduction of a dominant-negative Tcf4 molecule attenuates the response of F9 cells to RA (50). Our results suggest that highly active WNT signaling may not be important for maintenance of an undifferentiated, pluripotent phenotype in embryonal carcinoma cells, but activation of WNT signaling may be required for a response to a differentiation signal in these cells. Our results further suggest that other signaling pathways may also be important in maintaining pluripotency in embryonal carcinoma cells.

This study represents the largest expression profiling of GCTs to date, allowing the identification of previously unrecognized differentially expressed genes, including several novel candidate

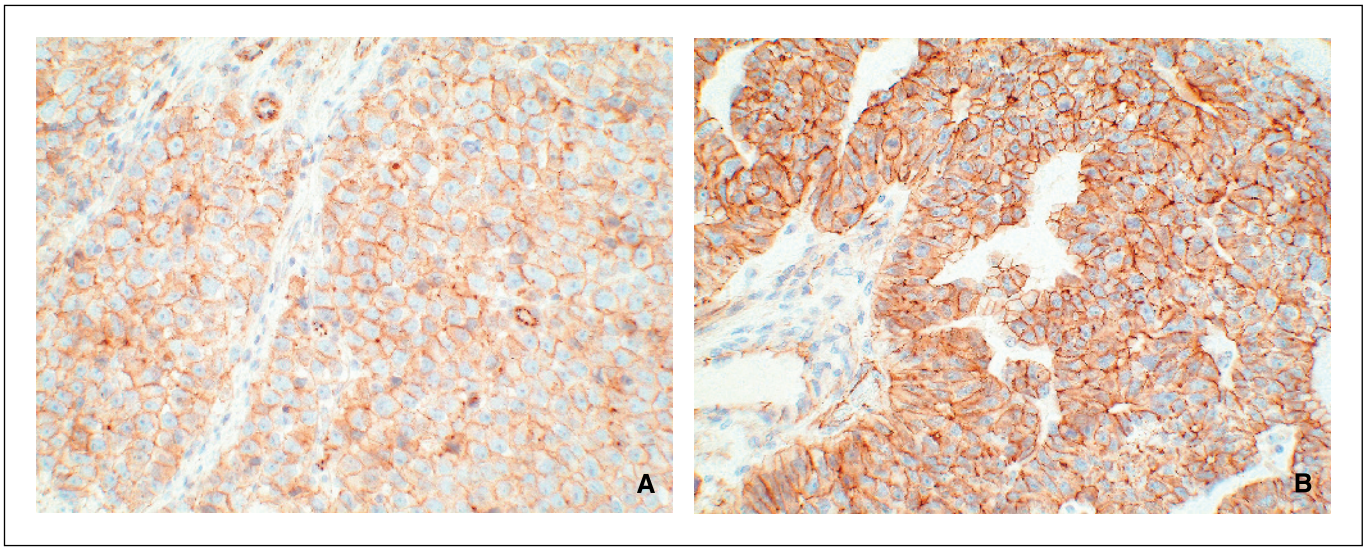

Figure 6. β-Catenin protein expression in seminoma and embryonal carcinoma. *A,* staining of seminoma for β-catenin is limited to membranes. *B,* staining of embryonal carcinoma for β-catenin reveals intense membrane staining, with some cytoplasmic staining also present.

target genes on 12p and stem cell–associated genes. Genetic changes represented by consistent gain or loss of chromosomes and chromosomal arms have been documented in many tumor types, although their biological significance remains unclear. Our results indicate that gain of the entire p arm of chromosome 12 may play multiple roles, driving the selection of this large genetic alteration during GCT tumorigenesis. Our study also raises several important questions that need to be addressed, such as the timing of overexpression of *POU5F1* and the homeobox/stem cell–related genes on 12p, the mechanism of silencing of stem cell genes on differentiation into yolk sac tumor, teratoma, and choriocarcinoma lineages, and the roles of WNT, FGF, and BMP signaling pathways and additional stem cell genes, such as *SOX2, TDGF1*, and *FGF4*, in the reestablishment of pluripotency and subsequent development of embryonal carcinoma tumors.

Acknowledgments

Received 7/12/2005; revised 10/3/2005; accepted 11/2/2005.

Grant support: Byrne Fund, Lance Armstrong Foundation, and NIH.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References

1. Bosl GJ, Motzer RJ. Testicular germ-cell cancer. N Engl J Med 1997;337:242–53.
2. Chaganti RS, Houldsworth J. Genetics and biology of adult human male germ cell tumors. Cancer Res 2000; 60:1475–82.
3. Ulbright TM. Germ cell neoplasms of the testis. Am J Surg Pathol 1993;17:1075–91.
4. Looijenga LH, Stoop H, de Leeuw HP, et al. POU5F1 (OCT3/4) identifies cells with pluripotent potential in human germ cell tumors. Cancer Res 2003;63:2244–50.
5. Almstrup K, Hoei-Hansen CE, Wirkner U, et al. Embryonic stem cell-like features of testicular carcinoma *in situ* revealed by genome-wide gene expression profiling. Cancer Res 2004;64:4736–43.
6. Chaganti RS, Houldsworth J, Bosl GJ. Molecular biology of adult male germ cell tumors. In: Vogelzang NJ, Scardino PT, Shipley WU, Coffey DS, editor. Comprehensive textbook of genitourinary oncology. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2000. p. 891–6.
7. Atkin NB, Baker MC. Specific chromosome change, i(12p), in testicular tumours? Lancet 1982;2:1349.
8. Rodriguez E, Houldsworth J, Reuter VE, et al. Molecular cytogenetic analysis of i(12p)-negative human male germ cell tumors. Genes Chromosomes Cancer 1993;8:230–6.
9. Bourdon V, Naef F, Rao PH, et al. Genomic and expression analysis of the 12p11-p12 amplicon using EST arrays identifies two novel amplified and overexpressed genes. Cancer Res 2002;62:6218–23.
10. Kraggerud SM, Skotheim RI, Szymanska J, et al. Genome profiles of familial/bilateral and sporadic testicular germ cell tumors. Genes Chromosomes Cancer 2002;34:168–74.
11. Zafarana G, Grygalewicz B, Gillis AJ, et al. 12p-amplicon structure analysis in testicular germ cell tumors of adolescents and adults by array CGH. Oncogene 2003;22:7695–701.
12. Meng FJ, Zhou Y, Giwercman A, Skakkebaek NE, Geurts van Kessel AD, Suijkerbuijk RF. Fluorescence *in situ* hybridization analysis of chromosome 12 anomalies in semen cells from patients with carcinoma *in situ* of the testis. J Pathol 1998;186:235–9.
13. Vos A, Oosterhuis JW, de Jong B, Buist J, Schraffordt Koops H. Cytogenetics of carcinoma *in situ* of the testis. Cancer Genet Cytogenet 1990;46:75–81.
14. Chaganti RS, Houldsworth J. The cytogenetic theory of the pathogenesis of human adult male germ cell tumors [review]. APMIS 1998;106:80–3; discussion 83–4.
15. Rosenberg C, Van Gurp RJ, Geelen E, Oosterhuis JW, Looijenga LH. Overrepresentation of the short arm of chromosome 12 is related to invasive growth of human testicular seminomas and nonseminomas. Oncogene 2000;19:5858–62.
16. Summersgill B, Osin P, Lu YJ, Huddart R, Shipley J. Chromosomal imbalances associated with carcinoma *in situ* and associated testicular germ cell tumours of adolescents and adults. Br J Cancer 2001;85:213–20.
17. Houldsworth J, Reuter VE, Bosl GJ, Chaganti RS. Aberrant expression of cyclin D2 is an early event in human male germ cell tumorigenesis. Cell Growth Differ 1997;8:293–9.
18. Clark AT, Rodriguez RT, Bodnar MS, et al. Human STELLAR, NANOG, GDF3 genes are expressed in pluripotent cells and map to chromosome 12p13, a hotspot for teratocarcinoma. Stem Cells 2004;22:169–79.
19. Rodriguez S, Jafer O, Goker H, et al. Expression profile of genes from 12p in testicular germ cell tumors of adolescents and adults associated with i(12p) and amplification at 12p11.2-p12.1. Oncogene 2003;22: 1880–91.
20. Draper JS, Smith K, Gokhale P, et al. Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat Biotechnol 2004;22:53–4.
21. Yen CC, Chen YJ, Chen JT, et al. Comparative genomic hybridization of esophageal squamous cell carcinoma: correlations between chromosomal aberrations and disease progression/prognosis. Cancer 2001; 92:2769–77.
22. Schraml P, Schwerdtfeger G, Burkhalter F, et al. Combined array comparative genomic hybridization and tissue microarray analysis suggest PAK1 at 11q13.5-q14 as a critical oncogene target in ovarian carcinoma. Am J Pathol 2003;163:985–92.
23. Korkola JE, Houldsworth J, Dobrzynski D, et al. Gene expression-based classification of nonseminomatous male germ cell tumors. Oncogene 2005;24:5101–7.
24. Houldsworth J, Heath SC, Bosl GJ, Studer L, Chaganti RS. Expression profiling of lineage differentiation in pluripotential human embryonal carcinoma cells. Cell Growth Differ 2002;13:257–64.
25. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4: 249–64.
26. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004; 5:R80.
27. Brazma A, Hingamp P, Quackenbush J, et al. Minimum Information About a Microarray Experiment (MIAME)—toward standards for microarray data. Nat Genet 2001;29:365–71.
28. Westfall PH, Young SS. Resampling-based multiple testing: examples and methods for *P*-value adjustment. New York: John Wiley & Sons; 1993.
29. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95: 14863–8.
30. Sperger JM, Chen X, Draper JS, et al. Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. Proc Natl Acad Sci U S A 2003;100:13350–5.
31. Du Z, Cong H, Yao Z. Identification of putative downstream genes of Oct-4 by suppression-subtractive hybridization. Biochem Biophys Res Commun 2001; 282:701–6.
32. Wang P, Branch DR, Bali M, Schultz GA, Goss PE, Jin T. The POU homeodomain protein OCT3 as a potential transcriptional activator for fibroblast growth factor-4 (FGF-4) in human breast cancer cells. Biochem J 2003;375:199–205.
33. Okada K, Hirota E, Mizutani Y, et al. Oncogenic role of NALP7 in testicular seminomas. Cancer Sci 2004;95: 949–54.
34. Brivanlou AH, Gage FH, Jaenisch R, Jessell T, Melton D, Rossant J. Stem cells. Setting standards for human embryonic stem cells. Science 2003;300:913–6.
35. Skotheim RI, Lind GE, Monni O, et al. Differentiation of human embryonal carcinomas *in vitro* and *in vivo* reveals expression profiles relevant to normal development. Cancer Res 2005;65:5588–98.
36. Willert J, Epping M, Pollack JR, Brown PO, Nusse R. A transcriptional response to Wnt protein in human embryonic carcinoma cells. BMC Dev Biol 2002;2:8.
37. Mann B, Gelos M, Siedow A, et al. Target genes of β-catenin-T cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. Proc Natl Acad Sci U S A 1999;96:1603–8.
38. Kratochwil K, Galceran J, Tontsch S, Roth W, Grosschedl R. FGF4, a direct target of LEF1 and Wnt signaling, can rescue the arrest of tooth organogenesis in Lef1(−/−) mice. Genes Dev 2002;16:3173–85.
39. Lickert H, Domon C, Huls G, et al. Wnt/(β)-catenin signaling regulates the expression of the homeobox gene Cdx1 in embryonic intestine. Development 2000;127: 3805–13.
40. Conacci-Sorrell ME, Ben-Yedidia T, Shtutman M, Feinstein E, Einat P, Ben-Ze'ev A. Nr-CAM is a target gene of the β-catenin/LEF-1 pathway in melanoma and colon cancer and its expression enhances motility and confers tumorigenesis. Genes Dev 2002;16:2058–72.
41. Chaganti RS, Rodriguez E, Bosl GJ. Cytogenetics of male germ-cell tumors. Urol Clin North Am 1993;20: 55–66.
42. Maecker HT, Todd SC, Levy S. The tetraspanin superfamily: molecular facilitators. FASEB J 1997;11: 428–42.
43. Oka M, Tagoku K, Russell TL, et al. CD9 is associated with leukemia inhibitory factor-mediated maintenance of embryonic stem cells. Mol Biol Cell 2002;13:1274–81.
44. Waterhouse R, Ha C, Dveksler GS. Murine CD9 is the receptor for pregnancy-specific glycoprotein 17. J Exp Med 2002;195:277–82.
45. Ferrier DE, Holland PW. Ancient origin of the Hox gene cluster. Nat Rev Genet 2001;2:33–8.
46. Grier DG, Thompson A, Kwasniewska A, McGonigle GJ, Halliday HL, Lappin TR. The pathophysiology of HOX genes and their role in cancer. J Pathol 2005;205: 154–71.
47. Murty VV, Houldsworth J, Baldwin S, et al. Allelic deletions in the long arm of chromosome 12 identify sites of candidate tumor suppressor genes in male germ cell tumors. Proc Natl Acad Sci U S A 1992;89: 11006–10.
48. Rao M. Conserved and divergent paths that regulate self-renewal in mouse and human embryonic stem cells. Dev Biol 2004;275:269–86.
49. Almstrup K, Hoei-Hansen CE, Nielsen JE, et al. Genome-wide gene expression profiling of testicular carcinoma *in situ* progression into overt tumours. Br J Cancer 2005;92:1934–41.
50. Liu T, Lee YN, Malbon CC, Wang HY. Activation of the β-catenin/Lef-Tcf pathway is obligate for formation of primitive endoderm by mouse F9 totipotent teratocarcinoma cells in response to retinoic acid. J Biol Chem 2002;277:30887–91.

Down-Regulation of Stem Cell Genes, Including Those in a 200-kb Gene Cluster at 12p13.31, Is Associated with *In vivo* Differentiation of Human Male Germ Cell Tumors

James E. Korkola, Jane Houldsworth, Rajendrakumar S.V. Chadalavada, et al.

Cancer Res 2006;66:820-827.