

Transcriptome and methylome analysis of CNS germ cell tumor finds its cell-of-origin in embryogenesis and reveals shared similarities with testicular counterparts

Hirokazu Takami[○], Asmaa Elzawahry, Yasin Mamatjan[○], Shintaro Fukushima, Kohei Fukuoka, Tomonari Suzuki, Takaaki Yanagisawa, Yuko Matsushita, Taishi Nakamura, Kaishi Satomi, Shota Tanaka, Akitake Mukasa, Nobuhito Saito, Masayuki Kanamori, Toshihiro Kumabe, Teiji Tominaga, Keiichi Kobayashi, Motoo Nagane[○], Toshihiko Iuchi, Kaoru Tamura, Taketoshi Maehara, Kazuhiko Sugiyama, Koji Yoshimoto, Keiichi Sakai, Masahiro Nonaka, Akio Asai, Kiyotaka Yokogami, Hideo Takeshima, Yoshitaka Narita[○], Soichiro Shibui, Yoichi Nakazato, Natsuko Hama, Yasushi Totoki, Mamoru Kato, Tatsuhiko Shibata, Ryo Nishikawa, Masao Matsutani, and Koichi Ichimura[○] on behalf of the Intracranial Germ Cell Tumor Genome Analysis Consortium (the iGCT Consortium)

Division of Brain Tumor Translational Research, National Cancer Center Research Institute, Tokyo, Japan (H.T., S.F., K.F., Y.M., T.N., K.S., K.I.); Department of Neurosurgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan (H.T., S.T., N.S.); Division of Cancer Genomics, National Cancer Center Research Institute, Tokyo, Japan (A.E., N.H., Y.T., M.K., T.S.); MacFeeters Hamilton Centre for Neuro-Oncology Research, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada (Y.M.); Faculty of Science, Thompson Rivers University, Kamloops, British Columbia, Canada (Y.M.); Division of Pediatric Neuro-Oncology, Saitama Medical University International Medical Center, Saitama, Japan (K.F., T.Y.); Department of Pediatrics, Saitama Children's Medical Center, Saitama, Japan (K.F.); Department of Neuro-Oncology/Neurosurgery, Saitama Medical University International Medical Center, Saitama, Japan (T.S., R.N., M.M.); Department of Neurosurgery and Neuro-oncology, National Cancer Center Hospital, Tokyo, Japan (Y.M., Y.N., S.S.); Department of Neurosurgery, Graduate School of Medicine, Yokohama City University, Kanagawa, Japan (T.N.); Department of Diagnostic Pathology, National Cancer Center Hospital, Tokyo, Japan (K.S.); Department of Neurosurgery, Kumamoto University Hospital, Kumamoto, Japan (A.M.); Department of Neurosurgery, Tohoku University School of Medicine, Miyagi, Japan (M.K., T.K., T.T.); Department of Neurosurgery, Kitasato University, Kanagawa, Japan (T.K.); Department of Neurosurgery, Kyorin University Faculty of Medicine, Tokyo, Japan (K.K., M.N.); Department of Neurosurgery, Chiba Cancer Center, Chiba, Japan (T.I.); Department of Neurosurgery, Tokyo Medical and Dental University, Graduate School of Medical and Dental Sciences, Tokyo, Japan (K.T., T.M.); Department of Neurosurgery, Hiroshima University Faculty of Medicine, Hiroshima, Japan (K.S.); Department of Neurosurgery, Kyushu University Hospital, Fukuoka, Japan (K.Y.); Department of Neurosurgery, Kagoshima University Hospital, Kagoshima, Japan (K.Y.); Department of Neurosurgery, Shinshu Ueda Medical Center, Ueda, Japan (K.S.); Department of Neurosurgery, Kansai Medical University, Osaka, Japan (M.N., A.A.); Department of Neurosurgery, University of Miyazaki Faculty of Medicine, Miyazaki, Japan (K.Y., H.T.); Department of Pathology, Hidaka Hospital, Gunma, Japan (Y.N.); Department of Brain Disease Translational Research, Juntendo University Faculty of Medicine, Tokyo, Japan (K.I.)

Corresponding Author: Department of Brain Disease Translational Research, Juntendo University Faculty of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan (k.ichimura.uk@juntendo.ac.jp).

Abstract

Background. CNS germ cell tumors (GCTs) predominantly develop in pediatric and young adult patients with variable responses to surgery, radiation, and chemotherapy. This study aimed to examine the complex and largely unknown pathogenesis of CNS GCTs.

Methods. We used a combined transcriptomic and methylomic approach in 84 cases and conducted an integrative analysis of the normal cells undergoing embryogenesis and testicular GCTs.

Results. Genome-wide transcriptome analysis in CNS GCTs indicated that germinoma had a transcriptomic profile representative of primitive cells during early embryogenesis with high meiosis/mitosis potentials, while nongerminomatous GCTs (NGGCTs) had differentiated phenotypes oriented toward tissue formation and organogenesis. Co-analysis with the transcriptome of human embryonic cells revealed that germinomas had expression profiles similar to those of primordial germ cells, while the expression profiles of NGGCTs were similar to those of embryonic stem cells. Some germinoma cases were characterized by extensive immune-cell infiltration and high expression of cancer-testis antigens. NGGCTs had significantly higher immune-cell infiltration, characterized by immune-suppression phenotype. CNS and testicular GCTs (TGCTs) had similar mutational profiles; TGCTs showed enhanced copy number alterations. Methylation analysis clustered germinoma/seminoma and nongerminoma/nonseminoma separately. Germinoma and seminoma were co-categorized based on the degree of the tumor microenvironment balance.

Conclusions. These results suggested that the pathophysiology of GCTs was less dependent on their site of origin and more dependent on the state of differentiation as well as on the tumor microenvironment balance. This study revealed distinct biological properties of GCTs, which will hopefully lead to future treatment development.

Key Points

- Germinomas are characterized by the primordial germ cell-transcriptomic profile.
- Nongerminomatous GCTs have ESC-like expression and immune-suppression phenotype.
- CNS and testicular GCTs have commonality in biology, implying the mutual cell-of-origin.

Importance of the Study

We report an integrated clinical, transcriptomic, and methylomic analysis of CNS germ cell tumors (GCTs) registered in the iGCT Consortium. The study revealed that germinomas were characterized by primordial germ cell (PGC)-like features including high expression and low promoter methylation in genes related to meiosis/mitosis and sexual reproduction. Nongerminomatous GCTs (NGGCTs) resembled embryonic stem cells with enhanced expression directed toward tissue formation and organogenesis. Immune-cell infiltration was enhanced in NGGCTs, characterized

with an immune-suppression phenotype. Co-analysis with testicular GCTs demonstrated that germinoma and seminoma were clustered together on methylation, separated on expression, and were internally divided based on the degree of immune-cell burden. While the mutation frequency was similar, copy number alterations and nonseminomatous histological components were more prevalent in testicular GCTs. These analyses suggest the common ancestry of GCT (likely PGCs) regardless of the location and indicate their potential for the development of site-agnostic treatment.

CNS germ cell tumors (GCTs) develop mostly in pediatric and young adult populations with a strong male preponderance. The WHO classification recognizes five main histological subtypes of tumors in the CNS: germinoma, teratoma (mature and immature), choriocarcinoma, yolk sac tumor, and embryonal carcinoma. Tumors often appear as a combination of any of these subtypes.¹

Germinoma, accounting for approximately 50%–60% of GCTs, is generally chemo- and radiation-sensitive and has a comparatively good prognosis. However, about 10% of germinoma cases are known to have a dismal clinical course.^{2,3} Nongerminomatous GCTs (NGGCTs), which

include all other types of CNS GCTs, are generally more resistant to therapy and require more intensive chemotherapy and radiation therapy. These high-dose chemotherapy and radiation therapies; however, potentially lead to post-therapeutic long-term sequelae including secondary malignancy and cognitive decline during follow-up; this is more prominent because many of the patients are in the pediatric and adolescent ages during treatment.^{4–6}

GCTs reflect the normal development of gametes and embryos, and their cells of origin are considered PGCs that migrate from the yolk sac to the genital ridge.⁷ The multidirectional differentiation of GCTs reflects normal

embryogenesis; however, their tumorigenic, unorganized development remains enigmatic. The pathogenesis of GCTs is possibly caused by the failure of apoptosis and the arrest in the maturation of PGCs into normal gametes till the pediatric and young adult period, with oncogenic transformation resulting from genetic abnormalities, such as MAPK and PI3K pathway mutations and/or copy number alterations, including the gain of 12p.^{8–10} The aberrantly directed movement of PGCs away from gonads and/or failed apoptosis of these cells are considered the underlying mechanism of extragonadal GCT development.

However, there has been no consensus on the cell of origin of GCT. Some believe GCTs originate from misplaced embryonic cells at various stages of embryogenesis during neural tube formation.¹¹

Transcriptome analysis of CNS GCTs revealed that germinomas were characterized by the expression of genes involved in self-renewal and immune response, while NGGCTs were characterized by that involved in neuron differentiation, the Wnt/beta-catenin pathway, invasiveness, and epithelial-mesenchymal transition.¹² However, the study was performed on a relatively small number of cases ($n = 21$); therefore, a more thorough investigation into the pathogenesis of GCTs is warranted.

In this study, we investigated the pathogenesis of CNS GCTs from a multidimensional perspective through transcriptomic analysis, integrative analysis of the transcriptome and methylome, comparative analysis between CNS GCTs and their testicular counterparts, and developmental analysis by incorporating previously published data for normal human embryonic cells, such as PGCs.^{13–16} We observed the distinct molecular profiles of germinomas and NGGCTs, positioning of GCTs in the lineage of normal embryogenesis, and overall similarities and partial differences between CNS and testicular GCTs.

Materials and Methods

See [Supplementary Materials and Methods](#).

Results

Unsupervised Hierarchical Clustering Clearly Separates GCTs by their Histological Subtypes

Expression data of 61 GCT samples and two normal samples (brain and adult testes) were analyzed using the unsupervised hierarchical clustering method ([Figure 1A](#)). A heatmap was created using 2044 transcripts with CV > 3. This distinctly separated germinoma and mixed GCTs/NGGCTs (as NGGCTs inclusive), the latter of which included two out of three TGCTs and two normal samples. All germinoma cases, except for two, were clustered together. In the NGGCT cluster, five YST cases and two mixed GCTs with a YST component aggregated specifically. The highly expressed transcripts in the YST cases shown as * ($n = 46$) in the heatmap are delineated in [Supplementary Table S2](#). The gene ontology for these genes analyzed

using DAVID revealed “extracellular region part” as the top ontology term.

Genes Characteristic to Germinoma and NGGCT

To identify the differentially expressed genes and expression profiles in germinoma and NGGCTs, the expression of 35 germinomas and 23 NGGCTs were compared ([Figure 1B](#)). Germinoma showed 5912 significantly upregulated transcripts (false discovery rate (FDR)-adjusted P value < .05, and FC > 2) and NGGCTs showed 4015 significantly upregulated transcripts. The highly expressed protein-coding transcripts in germinomas included abundant characteristic PGC/meiosis markers such as *PIWIL1*, *DAZL*, *DDX4*, *NANOS3*, and *ERVW-1* and cancer-testis antigen (CTA) genes such as *MAGEB1*, *MAGEB16*, and *MAGEC2*. Gene ontology and functional category analysis of these differentially expressed genes using DAVID indicated “spermatogenesis”, “meiosis”, “tumor antigen” and “immune response” among the top signature ontology terms for germinoma. NGGCTs were characterized by terms such as, “extracellular matrix”, “homeobox”, and “skeletal muscle development” among others ([Supplementary Tables S3 and S4](#)).

To further investigate the expression profile of germinoma, gene set enrichment analysis (GSEA; MSigDB) was applied to 15 germinoma cases with tumor content >50% upon methylation analysis and compared with 11 NGGCTs without a germinoma component. As [Figure 1C](#) shows that the immune-cell infiltration is relatively homogeneous in NGGCT cases compared with germinoma cases, all NGGCT cases were included in the analysis irrespective of the tumor content. There was no statistically significant (FDR-adjusted P value < .25) gene set in NGGCTs; however, germinoma was characterized by “mitotic cell cycle”, “meiosis”, “double strand break repair”, “nucleoplasm”, “DNA damage checkpoint”, “nuclear DNA directed RNA polymerase complex”, “chromatin assembly” among others ([Supplementary Figure S1](#); [Supplementary Table S5](#)).

Expression of Genes Representative of Embryo Development

The expression of genes that are representative of and are highly expressed during the various stages of normal embryonic development were investigated across the histological subtypes of GCTs. This was based on the hypothesis that the transcriptomic signatures of each histological subtype mirror that of developing embryonic cells, confirmed by the GCTs’ histological similarities to various normal developmental tissues as well as the hypothesis that GCTs’ cell-of-origin are the PGCs.

Genes were selected and categorized into the following six groups: “PGC markers (early and late)”, “pluripotency”, “meiosis”, “primitive endoderm/mesoderm”, “trophoblast”, and “gonadal somatic/control”. Thirty-two genes were selected based on the key reports on the study of normal human embryonic development.^{13–18} Expression of the 32 representative genes in the 61 samples of

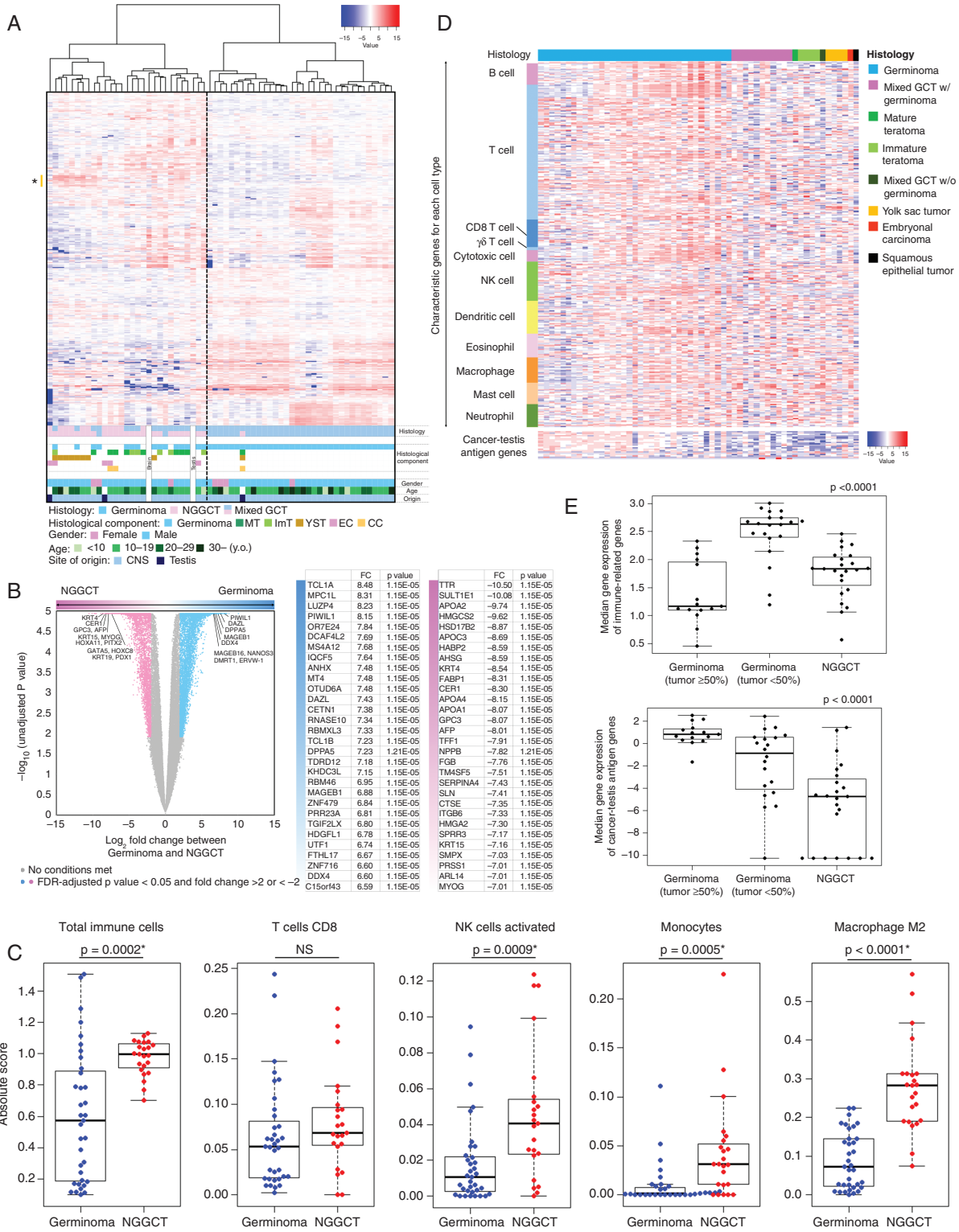


Fig. 1 (A) Heatmap with unsupervised hierarchical clustering of transcriptome profiles of 61 germ cell tumors (GCTs) and 2 normal samples (brain and testis) shows expression profiles GCTs. Each row represents a gene, and each column represents a sample. The level of expression (log-transformed RPKM) is represented in a color scale as depicted at the top right. Histology, histological components, gender, and age are indicated at the bottom. Germinoma and nongerminomatous GCTs (NGGCTs) are clearly separated. Genes highly expressed in yolk sac tumors are marked as * on the left side, which were used for gene ontology analysis. (B) Detection of histology-specific differences in the gene expression of germinoma and NGGCTs. Volcano plot illustrates the log two-fold change (FC) in gene expression (germinoma vs. NGGCT) on the x-axis and log-converted

GCTs and normal tissues are presented as bar graphs (Supplementary Figure S2). PGC marker genes (early and late) were highly expressed in germinomas, and genes for the late stage of PGC were highly expressed in the testis. Pluripotency genes were expressed in germinomas and other histological subtypes of GCTs; in addition, some pluripotency genes (*POU5F1* and *NANOG*) were highly expressed in embryonal carcinoma. Meiosis-related genes were highly expressed in germinomas and testes. Genes, which represent differentiation in specific directions, such as primitive endoderm, mesoderm, or trophoblast, exhibited distinct expression patterns, and they were expressed variably in germinomas but predominantly in NGGCTs.

Immune-Cell Infiltration in GCT

The 480 previously selected immune-related genes¹⁹ were variably expressed, especially among germinomas, possibly reflecting the variable ratio of germinoma and immune-cells in the samples (Figure 1D). To minimize the influence of the germinoma cell-immune-cell ratio, germinoma cases were divided into two groups depending on the cutoff of the tumor cell percentage estimated in H-E pathological specimens. Germinomas with a small tumor content exhibited abundant immune-cell infiltration when compared with germinomas with high tumor content and NGGCTs. The CTAs were highly expressed in germinomas with high tumor content when compared to that in the nongerminomas, which can be regarded as another germinoma characteristic ($P < .0001$) (Figure 1E).

CIBERSORT was used to probe the amount of immune-cell infiltration. On an average, germinomas showed variable amounts of immune-cell infiltration, while NGGCTs showed solid and higher immune-cell infiltration. The amount of CD8+T cells did not differ between germinomas and NGGCTs. However, the populations of natural killer cells (NK cells), monocytes, and macrophages M2 were significantly enriched in NGGCTs (Figure 1C, Supplementary Figure S3).

Subdivision of Germinoma on Transcriptome and Integrated Analysis With Methylation

Germinomas were best divided into two groups, based on unsupervised hierarchical clustering and silhouette width score analysis (Figure 2A, Supplementary Figure S4A). To

rule out the possibility of an incidental separation in clustering, a second analysis was performed without the NGGCT cases, using different sets of transcripts selected based on $CV > 3$ ($n = 582$). This showed the same clustering and the germinoma cases were again subdivided into two groups (Supplementary Figure S4B). This indicated that Group 1 had noticeably lower tumor content compared to Group 2.

We selected 208 transcripts out of 2044, having $CV > 3$ (Figure 2), which were visibly highly expressed in Group 2, and subjected them to GO analysis. This group of transcripts included pluripotency genes such as *UTF1*, *GDF3*, *GBX2*, and *TCL1A*, and meiosis genes such as *DAZL*, *MAEL*, *PIWIL2*, and *SYCE1*. Enriched ontology terms were “meiotic prophase”, “regulation of transcription, DNA-dependent” or “sexual reproduction”, all of which represented PGC characteristics (Supplementary Table S6). In Figure 2, 374 highly expressed Group 1 transcripts reflect abundant immune-related genes, such as those of B-cells, T-cells, NK cells, and dendritic cells. GO analysis of Group 1 transcripts ranked “immune response” as the top term with extremely low P values (Supplementary Table S7).

For investigating the clinical features, we compared age, sex, location of tumors, tumor content in the pathological specimen, and progression-free survival after chemotherapy and radiation therapy covering the ventricles. Figure 2B shows that the peak age was 10–19 years for Groups 2, while it was 20 years or older for Group 1. There were no significant differences in sex and tumor location (Figure 2C and D); however, there was a marked difference in the tumor content (Figure 2E). There was no significantly different prognosis between the two germinoma groups (Figure 2F). One case (GCT02) was omitted from the analyses for tumor location and prognosis as lacking these information.

Next, 32 cases having both expression and methylation data were analyzed for clustering, based on methylation data. Consequently, germinoma cases were subdivided into two or four subgroups (Supplementary Figure S4A and B).

Finally, data from both platforms were integrated, and two clusters were created by similarity network fusion (SNF) analysis, which were concordant with the methylation subgroups and mostly overlapped with expression subgroups (Supplementary Figure S6A). These integrated subgroups were characterized by distinct tumor content and marginally different prognosis (Supplementary Figure S6B and C).

Fig. 1 Continued

unadjusted p value between germinoma and NGGCT on the y -axis. Differentially expressed transcripts (FDR-adjusted P value $< .05$ and $FC > 2$ or < -2) between germinoma and NGGCT are highlighted in blue and pink, respectively. The top 30 protein-coding genes with the highest ratio between germinoma versus NGGCTs are shown in the table on the right. (C) Heatmap with 480 immune-related genes is displayed in the order of immune categories and histology. (D) Germinomas with low tumor content ($< 50\%$) were strongly infiltrated by immune-cells ($P < .0001$). Genes related to cancer-testis antigens were highly expressed in germinoma, especially among cases with high tumor content ($P < .0001$). (E) The degrees of immune-cell infiltration calculated using CIBERSORT for selected cell types. Immune-cells were more populated in NGGCTs, although some germinoma cases showed quite high immune-cell infiltration ($P = .0002$). Activated natural killer cells, monocytes, and M2 macrophages were more dominant in NGGCTs.

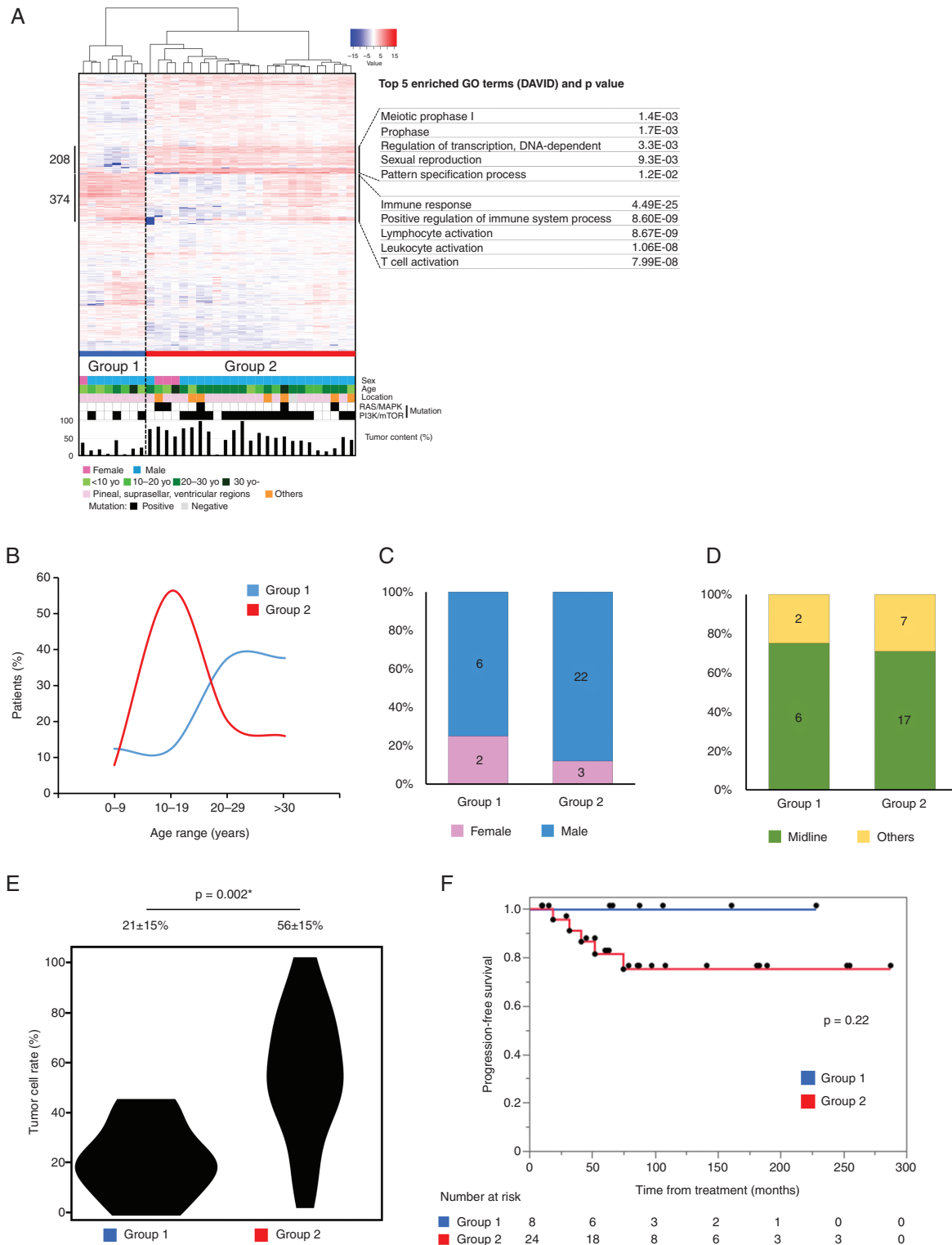


Fig. 2 (A) Heatmap representation of hierarchical clustering of germinoma cases using the 2084 genes used in Figure 1A. Each row represents a gene, and each column represents a sample. The level of expression (log-transformed RPKM) is represented with a color scale, as described previously. Three groups were named Groups 1 and 2. Highly expressed genes in Group 2 (208) and those in Group 1 (373) were analyzed using DAVID, and their characteristic gene ontology, together with representative genes are described in the right column. (B) Age distribution in each subgroup of germinoma. Group 2 had a peak age at 10–19, while Group 1 showed a peak age at 20 years or older. (C) Differences in gender distribution are not significant between the groups. (D) Tumor location is described as midline structure or other sites. There were no significant differences in the tumor locations based on the groups. (E) Tumor content in each group is indicated with the violin plot. Group 2 had a significantly higher

Correlation Between Expression and Methylation for Identifying Methylation-Associated Genes in Germinoma and NGGCT

To investigate the relationship between methylation and expression, and to define groups of genes whose expression might be positively regulated by methylation, we compared 15 germinoma cases with tumor content of more than 50% on methylation-based analysis and 23 NGGCT cases. A total of 7,545 protein-coding genes with methylation probes within CpG islands and transcription start site 200 were selected, in which methylation beta values were averaged if more than one probe existed in each gene (Supplementary Figure S7A). The signature genes, including 15 PGC-related genes, 19 meiosis-related genes, and 7 pluripotency-related genes were mapped. In particular, PGC- and meiosis-related genes were less methylated and more highly expressed in germinomas than in NGGCTs. There were 206 genes with a fold change >2 , and 84 genes with absolute methylation difference >0.3 . Twenty-one genes were both hypomethylated and highly expressed in germinomas compared to NGGCTs. Gene ontology analysis by DAVID showed terms associated with piRNA, gamete generation/fertilization, and meiosis among others (Supplementary Figure S7B; Supplementary Table S8).

Principal Component Analysis of Expression in GCTs and Normal Embryonic Cells

GCTs may reflect the developmental stages of embryogenesis; therefore, we compared the expression of GCTs in the histological subtypes and normal embryonic cells and integrated to establish a three-dimensional PCA (Figure 3A). During the early embryonic development from the zygote through the epiblast (curved arrow in blue), PGC and ESC appeared to branch out. NGGCTs were located close to ESCs, whereas germinomas were close to PGCs. A two-dimensional PCA analysis indicated similar results: normal embryogenic cells were aligned in their sequential development and GCTs were positioned next to ESC/PGC (Supplementary Figure S8A); PGCs and ESCs diverged after epiblast, clustering with germinoma, and NGGCT, respectively (Supplementary Figure S8B). These results suggest that NGGCTs may have a phenotype similar to that of ESCs, while germinoma to PGCs.

We performed an unsupervised hierarchical clustering analysis using the group of genes representing “pluripotency”, “early stage of PGC”, and “late stage of PGC”, obtained from previous reports. This reiterated that germinoma was clustered with PGC, while NGGCTs with ESCs. Oocytes, zygotes, 2-, 4-, and 8-cell stages, and morulae were clustered in a separate group (Figure 3B).

Shared Copy Number Status Among Different Histological Components in Mixed GCTs

Four representative mixed GCT cases were subjected to tissue microdissection and methylation analysis. The copy number status was similar in all the analyzed cases (Supplementary Figure S9).

Comparison of Pathological and Genomic Backgrounds of CNS and Testicular GCTs

The 84 CNS GCTs were combined with 134 TGCTs from The Cancer Genome Atlas (TCGA) to make 218 GCTs in total (Figure 4A). Histopathologically, the ratios of germinoma in CNS GCTs and seminoma in TGCTs were similar (43/84, 51.2% vs. 64/134, 47.8%); however, NGGCTs without a germinoma component were less frequent (22/84, 26.2%) than the nonseminomatous germ cell tumors (NSGCTs) without a seminoma component (53/133, 39.9%, $P = .042$). The histological components, YST and EC were more frequent in TGCTs. The frequency of mutations in *KIT*, *KRAS*, and *NRAS* was similar between CNS GCTs and TGCTs in each histological category (Figure 4B). The copy number alterations was more prominent in TGCTs, although the pattern was similar (Figure 4C). Statistically significant differences in copy number alterations and their frequency between CNS GCTs and TGCTs included a gain of chromosomes 7, 8, and 12p, and loss of chromosomes 4, 5, 10, 11, 13, 18.

The top 5,000 probes of standard deviation (SD) were used in the 450K methylation array data and all GCTs were clustered into two representative groups based on the tSNE analysis using the highest silhouette score (Figure 4D). The two clusters in tSNE were closely related to histology; one of the clusters (“Group A” in tSNE) was dominated by germinomas and seminomas and the other (“Group B”) was composed of mixed GCTs and NGGCTs/NSGCTs. The heatmap showed three well-separated clusters.

On analyzing the germinoma/seminoma cases separately, they were classified into two groups, based on the silhouette width score analysis (Figure 5A, Supplementary Figure S10). These groups were characterized by their degree of tumor content. Gene ontology analysis was performed for transcripts with fold change >2 and q -value < 0.05 . The highest tumor content group (“Group I”) was characterized by the gene ontology terms, such as “spermatogenesis”, “synapsis”, and “multicellular organism development”, which mostly represent male reproductive system and organ development. Group II, with low tumor content, was characterized by the gene ontology terms, such as “immune response”, “inflammatory response”, and “integral component of plasma membrane”, together summarized as immune response (Figure 5A, Supplementary Tables S9 and S10).

Fig. 2 Continued

(F) Progression-free survival (PFS) in each subgroup was shown using Kaplan–Meier curves. The differences in PFS among these groups were not significant ($P = .22$, log-rank test); however, Group 1 was better in prognosis. One case (GCT02) was omitted from the analyses for tumor location and prognosis as it lacked this information.

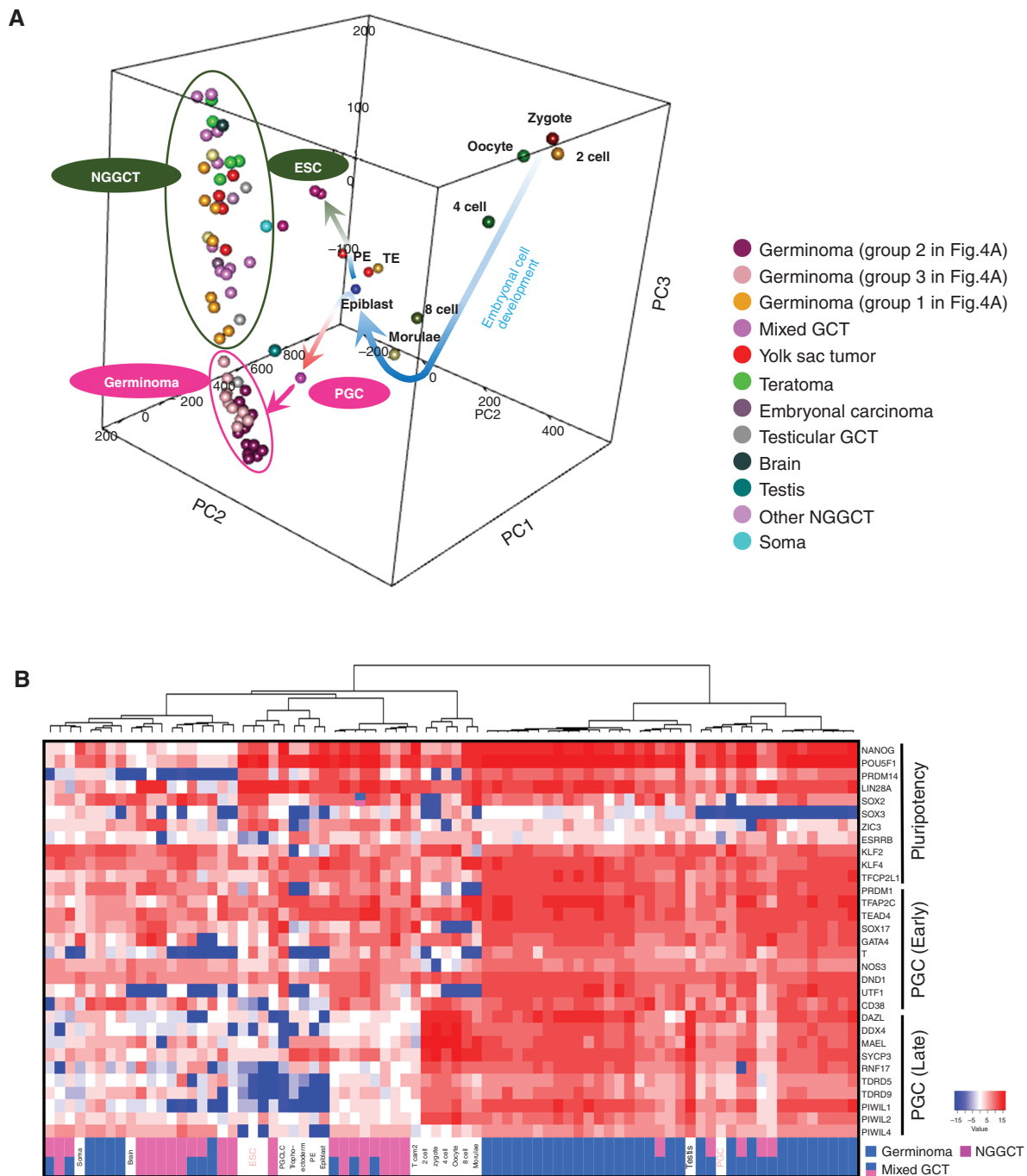


Fig. 3 Transcriptomic comparison between normal embryonic cells at various developmental stages and GCTs of all histopathological variants. (A) Three-dimensional principal component analysis showed that primordial germ cells (PGCs) and embryonic stem cells (ESCs) diverged from the morulae/epiblast. Germinomas (circled in pink) were located further away in the direction of PGCs and nongerminomatous GCTs (circled in dark green) were situated close to ESCs. (B) Supervised hierarchical clustering of the same set of embryonic cells and the current GCT cases using PGC (early and late)-related, and pluripotency-related genes as described in a previous report.¹⁵ Germinomas were mostly clustered with PGC, nongerminomatous GCT with ESC, and the normal embryonic cells from zygote to morulae were clustered separately. Epi, epiblast; PE, primitive endoderm; PGCLC, primordial germ cell-like cell; TE, trophoctoderm; Tcam2, testicular germ cell tumor cell line.

The transcriptomes of CNS GCTs and TGCTs were compared using 2901 transcripts with a CV > 3 (Figure 5B). tSNE distinctly separated germinomas from seminomas;

they were separately clustered from NGGCTs/NSGCTs; however, NGGCTs and NSGCTs were clustered together. GSEA showed "protein localization to Golgi apparatus"

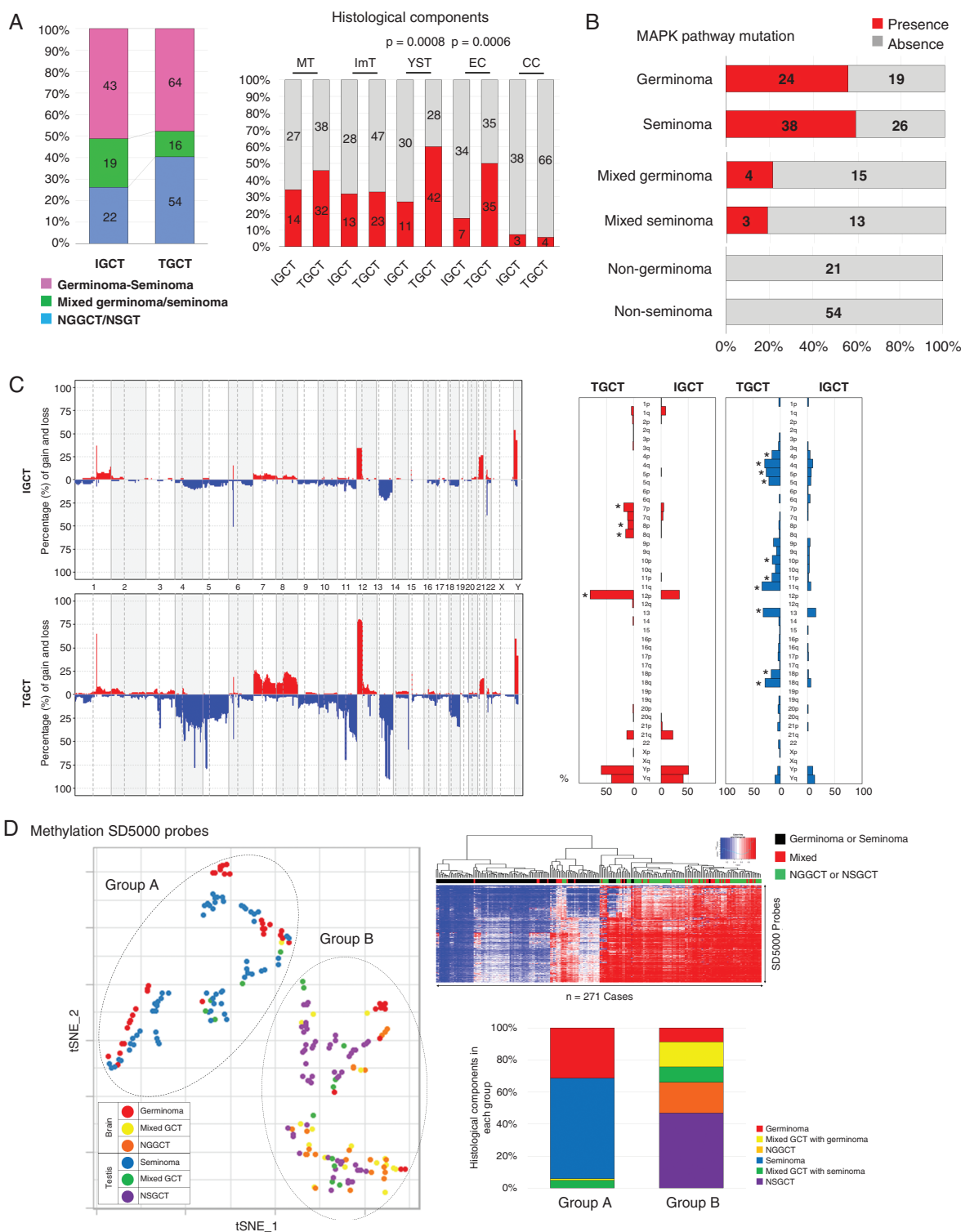


Fig. 4 Comparison between intracranial GCTs (IGCTs) and testicular GCTs (TGCTs). (A) The percentage of germinoma and seminoma in each category was similar; however, the rates of nongerminomatous GCTs (NGGCTs) and nonseminomatous GCTs (NSGCTs) were different (26% vs. 40%, $P = .04$). Yolk sac tumor (YST) and embryonal carcinoma (EC) are more frequently found in TGCTs. (B) The presence of MAPK pathway mutations is similar in both IGCT and TGCT. (C) Copy number alterations were more prominent in TGCT. Statistically different frequencies of copy number alterations between IGCT and TGCT included a gain of chromosomes 7, 8, 12p, and loss of chromosomes 4, 5, 10, 11, 13, 18. (D) Using the top 5000 probes of standard deviation (SD) on 450K methylation analysis, all the GCTs were clustered into two groups. The heatmap indicated 2 well-demarcated clusters, one with abundant demethylated probes, and the other with medium to highly methylated probes. The two clusters were related to histology; “Group A” was dominated by germinoma and seminoma and “Group B”, by mixed GCTs and NGGCT/NSGCT.

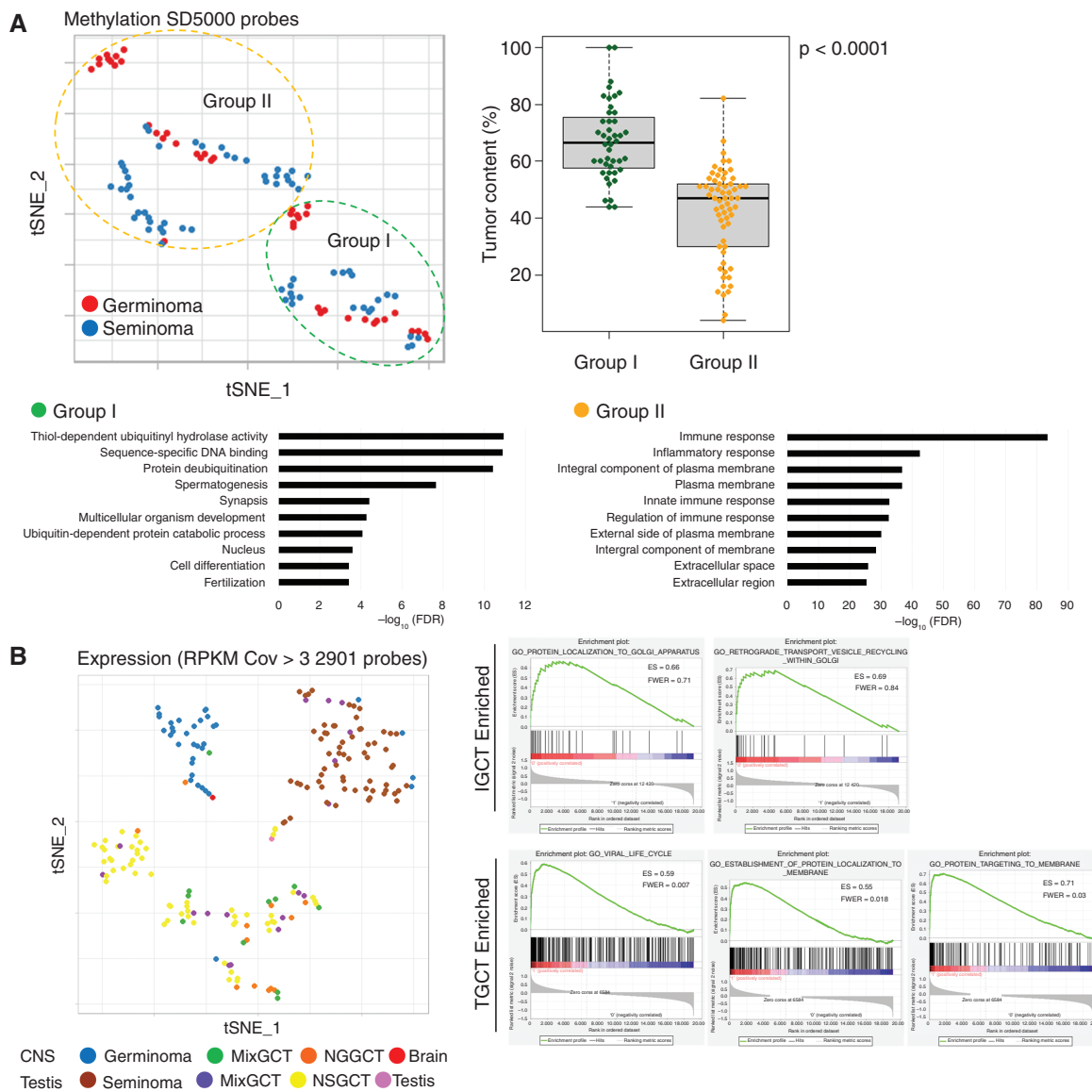


Fig. 5 (A) Germinoma/seminoma was classified into two groups characterized by the degree of tumor content. The highest tumor content group (“Group 1”) was characterized by the gene ontology terms representing male reproductive system and organ development. Group 2 was characterized by the gene ontology terms representing immune response. (B) Transcriptome of IGCT and TGCT are compared using the 2901 probes with the coefficient of variance >3. tSNE separates germinoma, seminoma, and others distinctly.

and “retrograde transport vesicle recycling within Golgi” as characterizing CNS GCTs, and “viral life cycle,” “establishment of protein localization to membrane,” “protein targeting to membrane” as characterizing TGCTs.

Discussion

This study is the first to analyze the combined genome-wide transcriptome and methylome of CNS GCTs. We addressed the long-standing question of the cell-of-origin of CNS GCTs from a transcriptome perspective. This study elucidated the transcriptomic resemblance of GCTs to the

stages of normal embryonic cell development. This pathophysiological investigation was previously discussed from an epigenetic point of view,²⁰ and this study corroborates those findings from a transcriptomic perspective. In addition, this study compared the molecular characteristics of CNS GCTs with their testicular counterparts.

Transcriptomic Differences in Germinomas and NGGCTs

Germinoma showed a transcriptomic profile distinct from NGGCTs. All germinoma cases except two were grouped

together. The two outlier germinoma cases (GCT27 and 103) were either characterized by extremely low tumor content (<5%) or elevated tumor markers (HCG 1051 IU/l and AFP 33.1 ng/ml). In the NGGCT group, YST formed a separate cluster, suggesting that YST may have a distinct etiology among the NGGCTs. Mature teratomas were also separately clustered; however, a component of teratoma is often found in other mixed NGGCTs. GSEA demonstrated that germinomas were characterized by signatures related to mitosis, meiosis, and normal germ cells, while NGGCTs were characterized by extracellular activities and organogenesis. This suggests that germinomas are closely related to cells at a primitive stage of germ cell development, which is the PGCs while, NGGCTs are closely related to cells in a more differentiated state, with a direction towards developing tissues and organs.

These findings were also emphasized by the expression of specific genes characteristic of embryonal development and those representing early-stage PGCs and meiosis were highly expressed in germinomas. This finding corroborates the long-standing hypothesis of an early-stage migrating PGC as the cell-of-origin of germinoma.²¹ In contrast, pluripotency genes were found to be highly expressed in embryonal carcinoma, which is consistent with the hypothesis that embryonal carcinoma may be a transformed counterpart of ESCs.²² Genes representing later stages of embryogenesis were highly expressed in NGGCTs, including YST. These findings are in line with the hypothesis that NGGCTs reflect more differentiated embryogenic cells in the direction of organogenesis and extra-embryonic tissues.²³ As the current study included only a small subset of the expression data of ESCs and PGCs, it was limited in its capacity to identify the genes dividing these primitive cells that may be commonly observed between germinoma and NGGCTs. Further study is warranted to investigate this aspect.

The germinoma cases were clustered into two distinct groups. The differences in the tumor content possibly contributed to this clustering, as germinomas often have a considerable amount of immune-cells coexisting with tumor cells, which is histopathologically termed as “two-cell pattern”.¹ Two or four subgroups were created when germinoma cases were analyzed based on methylation (Supplementary Figure S5A). Integrated analysis of data from the two platforms by SNF resulted in two subgroups, well-correlated with methylation and specifically, with expression subgroups. These clustering patterns may simply reflect the tumor content; however, there was a difference in the frequency of MAPK-pathway mutations (15/19 vs. 5/13, $P = .06$), patients’ age, and a tendency toward worse prognosis in the group with a higher tumor content. Further molecular background needs to be investigated, in addition to the observed mutation frequencies. We recently reported that tumor content is a prognostic factor in germinoma.^{24,25} These findings suggest the possibility of stratification of treatment intensity, depending on the tumor to immune-cell balance and warrant further investigation.

As demonstrated previously, germinoma is characterized by a genome-wide low methylation status.²⁰ In germinoma cases, we investigated the genes with a higher expression supposedly driven by low methylation. We identified 21 genes that had lower methylation (difference in beta values >0.3) and higher expression ($FC > 2$), in germinomas than in the NGGCTs. This group of genes included signature

genes for PGCs and meiosis and were characterized by the gene ontology terms of piRNA, gamete generation, and meiosis. This reiterated that germinomas had a distinct feature of early embryogenic status.

Another interesting finding is that while some germinomas harbor abundant immune-cells in the micro-environment, NGGCTs showed comparatively more abundant immune-cells in the CIBERSORT analysis. Activated natural killer cells, monocytes, and M2 macrophages were more enriched in NGGCTs. This monocyte-macrophage abundance was also reported in an independent CNS GCT cohort from a clinical perspective, and monocyte abundance in the preoperative cerebrospinal fluid was significantly associated with NGGCTs than with germinomas.⁶ This immunosuppressive phenotype of NGGCTs is important from an immunotherapy point of view, as this can be one of the underlying causes for treatment resistance in NGGCTs and can offer an opportunity for immune therapy. We are currently working on validating this finding by pathological specimen in a larger cohort.

Finally, by integrating the expression data of GCTs and normal developing embryonic cells, PCA showed that germinomas belonged to the lineage of PGCs, whereas NGGCTs were similar to ESCs. This, along with the transcriptomic similarities between PGCs and germinomas, dovetails with the hypothesis that PGC is the progenitor of germinoma. Our earlier study demonstrated that the different histopathological components in the same specimens, such as in mixed germinoma with mature teratomas, shared the same mutations.²⁰ The current study further demonstrated that representative copy number alterations may be shared among the histological components in the same tissue. As the number of samples investigated is quite small ($n = 4$), further research needs to be performed to validate the findings. These findings suggested that the mutation and copy number alteration occurred in a common tumor progenitor cell before it differentiated into various tumor subtypes as the DNA methylation status changed.²⁰ Based on the findings that the genomic insults appeared to have occurred in a common progenitor cell, the integrative analysis of expression from normal developmental cells and GCTs, and with reference to the scheme presented earlier,¹⁵ we hypothesized that an initiating genomic event *or events* occur first in migrating PGCs, germinomas develop in PGCs that are germ cell-destined, and NGGCTs develop in cells that are embryonically and extra-embryonically pluripotent, differentiating away from germ cell destination (Supplementary Figure S11). It must be reiterated that this model was established on a limited number of samples for normal developmental cells and GCT samples, and was based on a limited extent of analysis. Thus, the genes and pathways represented in this study remain hypothetical and require further analytical and experimental validation. We are working on a single-cell analysis of mixed GCTs with various subtypes. Pseudotime trajectory analysis may help delineate a potential path for NGGCT development.

CNS and Testicular GCTs

There was a distinct difference in the frequency of copy number alterations and the expression profile; however, there were a number of commonalities between CNS

and testicular GCTs. The mutation patterns in the MAPK pathway and the profiles of copy number alterations were very similar. Importantly, methylation analysis grouped germinomas and seminomas, as well as NGGCTs and NSGCTs. As reported previously, germinoma is characterized by whole-genome hypomethylation²⁰ and this was found to be shared by seminoma. The similarity in the methylation profiles of CNS GCTs and TGCTs probably reflects the common cell-of-origin among GCTs, regardless of the location. In contrast, there was a prominent difference in the expression profiles of germinoma and seminoma, while those of NGGCT and NSGCT were clustered together. One possibility is the different pathophysiological backgrounds underpinned by genomic alterations, such as distinct chromosomal instabilities, despite the seemingly common cell-of-origin and similar pathological phenotypes. Another possibility is the different microenvironments of the respective locations, which may provide distinct transcription factors. In a previous TGCT transcriptomic analysis comparing seminoma and YSTs, seminoma showed higher expression of transcription factors for maintaining stemness, while YSTs had higher expression of cytokines, endoderm, and endothelial markers.²⁶ Our results are in line with the earlier reports, indicating that germinoma is the CNS counterpart of seminoma. The similarities in the genomic/epigenomic profiles of CNS and testicular GCTs may provide insights into the treatment of both CNS and testicular GCTs.

In conclusion, our analyses revealed that germinomas are biologically akin to PGCs, which include various characteristics influenced by tumor cell/immune-cell balance, and that NGGCTs belong to the lineage of ESCs. There were differences in copy number and expression profiles; however, the genomic and epigenomic signatures of CNS and testicular GCTs were similar. These findings will hopefully lead to further understanding of the pathophysiology of this enigmatic tumor and serve as the basis for exploring therapeutic options.

Supplementary material

Supplemental material is available at *Neuro-Oncology* online.

Keywords

embryogenesis | germ cell tumor | methylation | RNA sequence | transcriptome

Funding

This work was supported by JSPS KAKENHI Grant Number 19H03440 and 20K17918, The Tokyo Society of Medical Sciences, and by AMED under Grant Number 21ck0106692s0701.

Acknowledgments

The authors thank all clinician practitioners who took care of patients and made a great contribution to this study by providing specimen and clinical information. We thank all participating institutions of the iGCT Consortium for their invaluable support. We thank Tomoko Urushidate for excellent technical assistance with next-generation sequencing and data analysis.

Conflict of interest statement. K.I. received a research grant from EPS Corporation. Department of Brain Disease Translational Research is an endowment department supported with an unrestricted grant from Idorsia Pharmaceuticals Japan Ltd.

Authorship statement. Conceptualization and project administration: H.T., K.I.; supervision: Y.T., M.K., T.S., R.N., K.I.; formal analysis: H.T., A.E., Y.M., M.C., Y.T., M.K., T.S.; investigation: H.T., A.E., Y.M., S.F., Y.T., M.K., T.S., K.I.; writing—original draft: H.T., A.E., M.K., K.I.; manuscript review: H.T., A.E., Y.M., S.F., K.F., T.S., T.Y., Y.M., T.N., K.S., S.T., A.M., N.S., M.K., T.K., T.T., K.K., M.N., T.I., K.T., T.M., K.S., K.Y., K.S., M.N., A.A., K.Y., H.T., Y.N., S.S., Y.N., N.H., Y.T., M.K., T.S., R.N., M.M., K.I.

References

- Louis D, Ohgaki H, Wiestler O, Cavenee W. *WHO Classification of Tumours of the Central Nervous System* (Revised 4th Edition). International Agency for Research on Cancer (IARC): Lyon; 2016.
- Matsutani M, Sano K, Takakura K, et al. Primary intracranial germ cell tumors: a clinical analysis of 153 histologically verified cases. *J Neurosurg*. 1997;86(3):446–455.
- Takami H, Fukuoka K, Fukushima S, et al. Integrated clinical, histopathological, and molecular data analysis of 190 central nervous system germ cell tumors from the iGCT Consortium. *Neuro-oncology* 2019;21(12):1565–1577.
- Sands SA, Kellie SJ, Davidow AL, et al. Long-term quality of life and neuropsychologic functioning for patients with CNS germ-cell tumors: from the First International CNS Germ-Cell Tumor Study. *Neuro Oncol*. 2001;3(3):174–183.
- Mabbott DJ, Monsalves E, Spiegler BJ, et al. Longitudinal evaluation of neurocognitive function after treatment for central nervous system germ cell tumors in childhood. *Cancer* 2011;117(23):5402–5411.
- Takami H, Perry A, Graffeo CS, Giannini C, Daniels DJ. Novel Diagnostic Methods and Posttreatment Clinical Phenotypes Among Intracranial Germ Cell Tumors. *Neurosurgery* 2020;87(3):563–572.
- Oosterhuis JW, Stoop H, Honecker F, Looijenga LH. Why human extragonadal germ cell tumours occur in the midline of the body: old concepts, new perspectives. *Int J Androl*. 2007;30(4):256–63; discussion 263.
- Ichimura K, Fukushima S, Totoki Y, et al. Recurrent neomorphic mutations of MTOR in central nervous system and testicular germ cell tumors may be targeted for therapy. *Acta Neuropathol*. 2016;131(6):889–901.
- Wang L, Yamaguchi S, Burstein MD, et al. Novel somatic and germline mutations in intracranial germ cell tumours. *Nature* 2014;511(7508):241–245.

10. Shen H, Shih J, Hollern DP, et al. Integrated molecular characterization of testicular germ cell tumors. *Cell Reports* 2018;23(11):3392–3406.
11. Sano K, Matsutani M, Seto T. So-called intracranial germ cell tumours: personal experiences and a theory of their pathogenesis. *Neurol Res* 1989; 11(2):118–126.
12. Wang HW, Wu YH, Hsieh JY, et al. Pediatric primary central nervous system germ cell tumors of different prognosis groups show characteristic miRNome traits and chromosome copy number variations. *BMC Genomics* 2010;11:132.
13. Gkoutela S, Zhang KX, Shafiq TA, et al. DNA demethylation dynamics in the human prenatal germline. *Cell* 2015;161(6):1425–1436.
14. Guo F, Yan L, Guo H, et al. The transcriptome and DNA methylome landscapes of human primordial germ cells. *Cell* 2015;161(6):1437–1452.
15. Irie N, Weinberger L, Tang WW, et al. SOX17 is a critical specifier of human primordial germ cell fate. *Cell* 2015;160(1):253–268.
16. Yan L, Yang M, Guo H, et al. Single-cell RNA-Seq profiling of human pre-implantation embryos and embryonic stem cells. *Nat Struct Mol Biol* 2013;20(9):1131–1139.
17. Tang WW, Dietmann S, Irie N, et al. A unique gene regulatory network resets the human Germline Epigenome for development. *Cell* 2015;161(6):1453–1467.
18. Sasaki K, Yokobayashi S, Nakamura T, et al. Robust in vitro induction of human germ cell fate from pluripotent stem cells. *Cell Stem Cell* 2015;17(2):178–194.
19. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 2013;39(4):782–795.
20. Fukushima S, Yamashita S, Kobayashi H, et al. Genome-wide methylation profiles in primary intracranial germ cell tumors indicate a primordial germ cell origin for germinomas. *Acta Neuropathol* 2017;133(3):445–462.
21. Stevens Jr LC, Little CC. Spontaneous testicular teratomas in an inbred strain of mice. *Proc Natl Acad Sci U S A* 1954;40(11):1080.
22. Donovan PJ, de Miguel MP. Turning germ cells into stem cells. *Curr Opin Genet Dev* 2003;13(5):463–471.
23. Teilum G. Classification of endodermal sinus tumour (mesoblastoma vitellinum) and so-called “embryonal carcinoma” of the ovary. *Acta Pathol. Microbiol. Scand* 1965;64(4):407–429.
24. Takami H, Fukushima S, Aoki K, et al. Intratumoural immune cell landscape in germinoma reveals multipotent lineages and exhibits prognostic significance. *Neuropathol Appl Neurobiol* 2020;46(2):111–124.
25. Takami H, Satomi K, Fukuoka K, et al. Low tumor cell content predicts favorable prognosis in germinoma patients. *Neuro-Oncol Adv* 2021;3(1). PMID: 34549182.
26. Poynter JN, Bestrashniy JR, Silverstein KA, et al. Cross platform analysis of methylation, miRNA and stem cell gene expression data in germ cell tumors highlights characteristic differences by tumor histology. *BMC Cancer* 2015;15:769.