

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Shah SP, Köbel M, Senz J, et al. Mutation of *FOXL2* in granulosa-cell tumors of the ovary. *N Engl J Med* 2009;360:2719-29. DOI: 10.1056/NEJMoa0902542.

Supplemental Appendix

Supplement to Shah SP, Köbel M, Senz J et al. Mutation of the *FOXL2* gene in granulosa cell tumors of the ovary

Contents

- I. Supplemental Methods and Results
- II. Supplemental Tables
- III Supplemental Figures
- IV References

I. Supplemental Methods and Results

Ethics

The tumor specimens analyzed in this study were collected via a BCCA prospective tumor banking protocols for ovarian tumors (Formation of Gynaecological Tumour Bank). Patients were approached for consent to donate tissue surplus to diagnostic requirements plus a blood sample, for use in a research ethics board (REB) approved research protocol. All patients were informed at the time of consent about the potential risks of loss of confidentiality arising from research and that none of the research study data would ever form part of the clinical record or be reported back to the care physicians. Specimen materials were released to investigators only with an REB approved study certificate, via a privacy guardian who anonymizes samples such that research investigators have no access to patient identifiers or information that could be used to directly identify a subject. The REB ethics protocol for genome-scale datasets stipulates that primary datasets from the granulosa cell tumor transcriptomes will not be released into the public domain, but can be made available via a tiered access mechanism to named investigators of institutions agreeing by a materials transfer agreement to honor the same ethical and privacy principles as the BCCA investigators.

Paired –End RNA Sequencing and Analysis

For each patient sample polyadenylated RNA was purified from 10ug of DNaseI (Invitrogen, Carlsbad, CA) treated total RNA using the MACS™ mRNA Isolation Kit (Miltenyi Biotec, Germany). Double-stranded cDNA was synthesized from the purified polyA+RNA using Superscript™ Double-Stranded cDNA Synthesis kit (Invitrogen,

Carlsbad, CA) and random hexamer primers (Invitrogen) at a concentration of 5 μ M. The resulting cDNA was sheared using a Sonic Dismembrator 550 (Fisher Scientific, Canada) and size separated by PAGE (8%). The 190-210bp DNA fraction was excised, eluted overnight at 4 °C in 300 μ l of elution buffer (5:1, LoTE buffer (3 mM Tris-HCl, pH 7.5, 0.2 mM EDTA)-7.5 M ammonium acetate) and purified using a QIAquick purification kit (Qiagen, Mississauga, ON). The sequencing library was prepared following the Illumina Genome Analyzer paired end library protocol (Illumina Inc., Hayward, CA) with 10 cycles of PCR amplification. PCR products were purified on Qiaquick MinElute columns (Qiagen, Mississauga, ON) and assessed and quantified using an Agilent DNA 1000 series II assay and Qubit fluorometer (Invitrogen, Carlsbad, CA) respectively. The resulting libraries were sequenced on an Illumina Genome Analyzer II following the manufacturer's instructions. Image analysis and basecalling was performed by the GA pipeline v1.0 (Illumina Inc., Hayward, CA) using phasing and matrix values calculated from a control phiX174 library run on each flowcell. Raw Quality scores were calibrated by alignment to the reference human genome (NCBI build 36.1, hg18) using ELAND (Illumina Inc., Hayward, CA).

Mapping

Short read sequences obtained from the Illumina Genome Analyzer were mapped to the reference human genome (NBCI build 36.1, hg18) plus a database of known exon junctions¹ using MAQ² in paired end mode. The exon junction sequences consisted of N-1 nucleotides of the donor exon 'spliced' to N-1 nucleotides of the acceptor exon for every known pair of exons in the human transcriptome, where N is the read length. This

was designed to rescue reads that would not map to the human reference genome due to their spanning a splice site. The maximum insert size (-a parameter) to MAQ was 500.

Single nucleotide variants (SNVs)

Single nucleotide variants were predicted using a novel Bayesian mixture model, *SNVmix*, whereby aligned reads were processed by counting the number of reference and non-reference bases at every coding nucleotide in the genome that had at least one read mapped to it. *SNVmix* is a probabilistic generative model that assumes the count data were generated by one of three genotype states: *aa* (homozygous for the reference base), *ab* (heterozygous), and *bb* (homozygous for the non-reference base); we considered the latter two genotypes to be an SNV. We used *SNVmix* to determine the probability of each of these states given the model parameters and the counts for each position in the transcriptome. For each location i , with observed data, we define three random variables:

$G_i \in \{aa, ab, bb\}$ represents the genotype, $N_i \in \{0, 1, \dots\}$ is the number of reads, and $a_i \in \{0, 1, \dots, N_i\}$ is the number of reference reads. The goal is to infer G_i from N_i and a_i .

We assume the following likelihood model for the data:

$p(a_i | G_i = k, N_i, \mu_{1,3}) \sim \text{Binomial}(a_i | N_i, \mu_k)$ where μ_k is the probability of a reference read for genotype k , where $k \in \{aa, ab, bb\}$. (We discuss how to estimate the μ parameters below.) We also have a prior on the genotype, $G_i | \pi \sim \text{Multinomial}(\pi, 1)$, where $\pi(k)$ is the prior probability of genotype k . By setting $\pi(aa) \gg \pi(ab), \pi(bb)$, we can encode our prior belief that most locations will be homozygous for the reference. Given knowledge of all the parameters, $\theta = (\mu_{1,3}, \pi)$ we can use Bayes rule to infer the posterior over genotypes, $\gamma_i(k) = p(G_i = k | a_i, N_i, \theta)$. We used the expectation maximization (EM)

algorithm to find a MAP estimate of the parameters given some training data. In our training data, we observed a_i and N_i for about 15,000 locations in a deeply sequenced breast cancer transcriptome. The 15,000 locations were determined by using data from an orthogonal assay, namely genotype calls from an Affymetrix SNP 6.0 genotyping array. We limited the positions by only considering coding positions on the array where the CRLMM algorithm³ predicted genotype with > 0.99 confidence. Given the 15,000 positions, we fit the model as follows using EM: in the E-step, we estimated $p(G_i = k | a_i, N_i, \theta)$ and in the M-step we estimated $\mu_{1,3}$ and π . We use a Dirichlet prior on π , with parameters $\delta = (1000, 100, 100)$, to reflect a strong prior belief that state aa 10 times more likely than states ab or bb . We use a Beta prior on each μ_k , with parameters α_k and β_k chosen so that the prior means are as follows: $\mu_{aa} \sim 1$, $\mu_{ab} \sim 0.5$, and $\mu_{bb} \sim 0$.

We therefore defined any position that CRLMM predicted to be heterozygous or homozygous for the non-reference allele to be a true positive (TP) and any position that CRLMM predicted to be homozygous for the reference allele to be a true negative (TN). We then computed $p(S_i) = \gamma_i(ab) + \gamma_i(bb)$ representing the posterior marginal probability of an SNV at position i as predicted by the model. This allowed us to compute a standard receiver operator characteristic (ROC) curve; SNVs were required to have $p(S) > 0.53$, (corresponding to a false discovery rate of 0.01) to be considered for further analysis.

Using these converged parameters and the threshold of 0.53, we applied the model to the aligned RNA-seq data from 15 ovarian cancer transcriptomes. Only bases with $> Q20$ base quality were considered in determining the counts in order to minimize errors. Predicted SNVs were cross-referenced against dbSNP version 129, the genomes

of James Watson ⁴ and Craig Venter ⁵, the genome of anonymous Yoruban ⁶ and Asian individuals ⁷, and six individuals from the 1000 genomes project (<http://www.1000genomes.org>) in order to eliminate any previously described germline variants.

Additional filtering was applied to the lists of SNVs. The junction database referred to above introduced alignment artifacts that appeared as SNVs in our data. These had a recognizable pattern of occurring in introns of a splice variant of a gene that overlapped an exon of a different splice variant due to sequence similarity of the intron sequence and the downstream exon. These tended to occur within a few nucleotides of the donor site of an exon, therefore we disregarded SNVs that fit this criteria, and were within 5bp of a donor site.

Any remaining SNVs were carried forward to subsequent analyses.

Insertions and deletions (indels)

Insertions and deletions were analysed using Maq's *indelpe* program. To reduce false positive predictions, we applied filters to the results so that we only carried forward indels for which there were supporting reads from both the forward and reverse directions, there were at least 2 reads supporting the indel and the ratio of mutant to wildtype reads was at least 0.1. In addition, indels with low complexity sequence on either side of the putative event were excluded from further analysis under the assumption that these were due to mis-alignment errors.

Determining GCT-specific events

Supplemental Figure 3 shows a flow chart describing the steps used to determine GCT-specific SNVs. We detected 29 SNVs present in 3 or more GCTs. 24 of these were

present in non-GCT cases. Of the remaining 5 SNVs, 3 were determined to be exon junction read artifacts. Of the remaining 2, ZNF141 chr4:357452 did not validate by PCR amplicon sequencing. The remaining SNV was FOXL2 chr3:140147853, which was carried forward for further analysis. 17 indels were predicted in 3 or more GCTs. 16 of these were found in other cases and the remaining indel was determined to be an alignment artifact due to low complexity sequence. An indel in RBBP6 chr16:24488600, a 1 bp insertion was detected in 3 GCTs early on in the data acquisition phase of this research. This indel did not validate by PCR (see Figure 3 in the main text). Note that as additional data was acquired, this indel no longer met the criteria described above. Please see supplemental data for all SNVs and indels found in 3 or more GCTs, and the occurrence of these events in non-GCT cases. The complete list of all SNVs and indels identified in the GCT cases will be available under the tiered access approach described above from the European Genotyping Archive.

Gene expression

We calculated the per-nucleotide depth for each position using only bases with $> Q20$. For each gene, we plotted the read depth against nucleotide position to determine an expression profile. The read depth was normalized against the total number of mapped reads for each case in order to account for variances in the number of reads generated across cases. These plots were used to qualitatively determine that these profiles matched the current standards of expression markers for the GCTs in comparison with non-GCTs.

Cross referencing GCTs against other ovarian cancers

We sequenced polyadenylated RNA from four GCTs, 10 ovarian carcinomas (five clear cell carcinomas, four endometrioid carcinomas, and one high grade serous carcinoma) and one serous borderline tumor derived cell line (SBOT-3.1). SBOT-3.1 has been previously shown to be genomically stable with a minimally abnormal karyotype⁸. Only SNVs and indel mutations that were unique to the GCTs and occurred in at least three of the four GCT cases were triaged for further validation.

Inferring copy number from Affymetrix SNP 6.0

The evaluation of genomic instability in the four index cases by Affymetrix SNP 6.0 high density genotyping arrays was performed using manufacturer's recommended procedures. The Affymetrix SNP 6.0 data were processed using the R-based *aroma.affymetrix* package⁹. Total copy number was estimated according to the CRMA procedure, calibrating for allelic cross talk, nucleotide-position probe sequence effects, robust probe summarization and correction for fragment length effects. We used a randomly selected set of 10 publicly available arrays from the 270 HapMap individuals (available from the Affymetrix website) to use as a reference to infer total copy number, then plotted the data to infer genomic instability.

Immunohistochemistry on GCTs

Tissue blocks were available from 32 GCT cases and used for TMA construction along with six other control ovarian carcinomas. A representative area of each tumor was selected by a pathologist and duplicate 0.6 mm tissue cores were extracted to construct a tissue microarray (TMA) (Pathology Devices Inc., Westminister, MD). Serial 4µm sections were cut for immunohistochemical (IHC) analysis. Calretinin (rabbit polyclonal

from Cell Marque Corp. (Rocklin, CA), dilution 1:200), and EMA (clone E29 from Dako (Carpinteria, CA), dilution 1: 200) were run using an automated protocol including heat antigen retrieval buffer CC1 (Ventana Benchmark, Tucson, AZ). Inhibin (clone R1 from AbD Serotec (Raleigh, NC), dilution 1:10) was stained on a Dako autostainer following antigen retrieval in 10 mM citrate buffer for 8 min in a microwave pressure cooker.

FOXL2 (goat polyclonal antibody from Imgenex Corp. (San Diego, CA), dilution 1:25) was run using an automated protocol with a standard CC2 buffer for heat antigen retrieval (Ventana Discovery XT, Tucson, AZ). Any expression was considered as positive.

Fluorescent In-Situ Hybridization on GCTs

TMA cores were assayed for FOXL2 copy number. 6 - μ m-thick sections were pre-treated as described previously¹⁰. FISH assays were performed using BACs (RP11-71B5 (ch3:140100841-140234525) and RP11-769B22 (chr3: 139997418-140186224)) from the British Columbia Genome Sciences Centre and a Vysis Spectrum Orange CEP 3 probe (3p11.1-q11.1) (Abbott Molecular Laboratories, Abbott Park, IL). BACs were directly labelled with Spectrum Green using a Nick Translation Kit (Abbott Molecular Laboratories, Abbott Park, IL). Analysis was done on a Zeiss Axioplan epifluorescent microscope. Images were captured using Metasystems Isis FISH imaging software (MetaSystems Group, Inc. Belmont MA).

RNA and DNA Extraction

Either eight micron sections, scrolls or 0.6mm cores from formalin fixed paraffin embedded (FFPE) tumors were used for DNA extraction. In cases with regional tumor infiltration on a slide, the tumor was circled on a matched stained section and FFPE

material from that area was scraped off the unstained sections with a fresh scalpel blade. Two paucicellular cases (GCT17 and GCT46) were subjected to laser capture microdissection (Veritas, Molecular Devices, Sunnyvale, CA). To obtain normal DNA from cases of mutation positive GCT, slides from 48 cases were reviewed to find tissue blocks with normal tissue. Standard methodologies for RNA and DNA extraction and the generation of cDNA from RNA were used.

Restriction Fragment Linked Polymorphism Analysis

For the paraffin samples, extraction was done using the RecoverAll Total Nucleic Acid Isolation kit for from Ambion (AppliedBiosystems, Foster City, CA) following the manufacturer's advised protocols. A 182bp fragment of the *FOXL2* gene was amplified from DNA using PCR with the following primers: 5'-CACAACTCAACGAGTGC-3' (forward) and 5'-TTGCCGGGCTGGAAGTGCG-3' (reverse). After denaturation at 94°C for 1 min, DNA was amplified over 35 cycles (94°C 30 sec, 70°C 30 sec, 72°C 30 sec) using a Tetrad 2 Peltier Thermal Cycler (MJ Research, Ramsey, MN). Final extension was at 72°C for 5 min. PCR products were digested with BstNI (New England Biolabs, Ipswich, MA) at 60°C for 1 hour and resolved on a 3% agarose gel. The 402C<G mutation creates a BstN1 restriction site which yields fragments of 90 and 92bp.

TaqMan Assay

A Custom TaqMan SNP Genotyping assay (Applied Biosystems, Foster City, CA) was performed according to manufacturer's instructions. The sequences of the primers were as follows: 5'-GCGCAAGGGCAACTACTG-3' (forward) and

5'CGGTAGTTGCCCTTCTCGAA-3' (reverse). The wildtype specific probe (5'-FAM dye -CATGTCTTCCCAGGCCG-NFQ (non-fluorescent quencher)) and mutation specific probe (5'-VIC dye-CATGTCTTCGCAGGCCG-NFQ) were included in the genotyping master mix and reactions were performed in a 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) using 20 ng of DNA. After denaturation at 95°C for 10minutes, DNA was amplified over 40 cycles (95°C 15seconds, 60°C 1 minute).

Direct Sequencing

FOXL2 was specifically PCR amplified from DNA using the following primers: 5'-CCAGTACATCATCGCGAAGTTCCCG-3' (forward) and 5'-CTCCGGCCCCGAAGAGCC-3' (reverse). Primers distinct from the PCR-RFLP assay were used to prevent misleading results due to primer site polymorphisms. After denaturation at 94°C for 1 min, DNA was amplified over 35 cycles (94°C 30 sec, 66°C 30 sec, 72°C 30 sec) using an MJ Research (Ramsey, MN) Tedrad. Final extension was at 72°C for 5 min. PCR products were purified using a MinElute PCR purification kit (Qiagen, Valencia, CA) and bi-directionally sequenced using an ABI BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) and an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA).

Homology model of the human FOXL2 forkhead domain bound to DNA

Homology models were derived from the structure of the interleukin enhancer binding factor 1 forkhead domain bound to DNA (53% identical to FoxL2 at the amino acid

sequence level, PDB ID: 2C6Y), using the SWISS Model homology modeling server (<http://swissmodel.expasy.org/SWISS-MODEL.html>).

Results

Using a combination of RFLP, Taqman and direct sequencing five cases out of 69 GCTs did not yield usable results. Three were repeatedly positive by the TaqMan assay but were mutation negative by Sanger sequencing. Two were paucicellular and required laser capture microdissection. These tested negatively by the TaqMan assay but did not produce enough DNA for sequencing. The final unusable result came from GCT80, a case with a source diagnosis of GCT that was mutation positive by both TaqMan assay and sequencing, however, we were unable to access material for pathology review.

II. Supplemental Tables

Supplemental Table 1. *FOXL2* mutation status in tumor and normal tissue as well as basic clinical parameter and histopathological review diagnosis for the requested cases of potential GCT provided by each institution. AGCT:adult-type granulosa cell tumor; JGCT:juvenile-type granulosa cell tumor; SLCT:Sertoli-Leydig cell tumor; N/A: not available. Note that GCT78 and GCT59 are recurrences associated with the primary tumors GCT29 and GCT76 respectively. The 149 additional ovarian tumors analyzed included 81 high-grade serous, 7 clear cell, 12 endometrioid, 1 mixed undifferentiated/endometrioid, 4 mixed endometrioid/clear cell, 10 mucinous carcinomas, 10 mucinous borderline tumors, 4 mucinous borderline tumors with intraepithelial carcinoma, 7 low grade serous carcinomas, 8 serous borderline tumors, 2 adenocarcinoma NOS, 1 ovarian small cell carcinoma hypercalcemic type, and 2 carcinosarcomas.

GCT Study ID	Source Hospital	Histology after arbitrary	Tumour Mutation Status	Normal Mutation Status	Additional Comments	Age at Surgery	Stage At Diagnosis	Primary or Recurrent
INDEX CASES								
GCT26	VGH	AGCT	Positive	Negative	Index Case	79	IIIc	Recurrence
GCT28	VGH	AGCT	Positive	Negative	Index Case	78	N/A	Recurrence
GCT77	VGH	AGCT	Positive	N/A	Index Case	41	Ia	Primary
GCT78	VGH	AGCT	Positive	N/A	Index Case; Recurrence associated with GCT29	40	IIa	Recurrence
FIRST SERIES – GRANULOSA CELL TUMORS								
GCT01	VGH	AGCT	Positive	Negative		78	IIb	Recurrence
GCT02	VGH	AGCT	Positive	Negative		41	Ia	Primary
GCT03	VGH	AGCT	Positive	Negative		26	Ia	N/A
GCT04	VGH	AGCT	Positive	Negative		39	Ic	Primary
GCT05	VGH	AGCT	Positive	N/A		56	Ic	Primary
GCT06	VGH	AGCT	Positive	Negative		65	Ia	Primary
GCT07	VGH	AGCT	Positive	Negative		61	Ia	Primary
GCT08	VGH	AGCT	Positive	Negative		42	Ia	Primary
GCT09	VGH	AGCT	Positive	Negative		66	Ia	Primary

GCT10	VGH	AGCT	Positive	Negative		75	Ia	Primary
GCT12	VGH	AGCT	Positive	Negative		46	Ia	Recurrence
GCT13	VGH	AGCT	Positive	Negative		48	Ia	Primary
GCT14	VGH	AGCT	Positive	Negative		53	Ia	Primary
GCT15	VGH	AGCT	Positive	Negative		74	Ic	Primary
GCT16	VGH	AGCT	Positive	Negative		48	Ia	Primary
GCT19	VGH	AGCT	Positive	Negative		61	Ia	Primary
GCT23	Richmond General	AGCT	Positive	Negative		52	Ia	Primary
GCT25	VGH	AGCT	Positive	Negative		48	Ia	Primary
GCT27	VGH	AGCT	Positive	Negative		41	IIb	Recurrence
GCT29	VGH	AGCT	Positive	Negative		39	IIa	Recurrence
GCT30	VGH	AGCT	Positive	Negative		77	Ia	Primary
GCT31	VGH	AGCT	Positive	Negative		49	Ia	Primary
GCT32	VGH	AGCT	Positive	Negative		37	Ic	Recurrence
GCT34	Montreal	AGCT	Positive	Negative		56	N/A	N/A
GCT35	Montreal	AGCT	Positive	Negative		48	N/A	N/A
GCT36	Montreal	AGCT	Positive	Negative		39	N/A	N/A
GCT37	Montreal	AGCT	Positive	N/A		N/A	N/A	N/A
GCT38	Montreal	AGCT	Positive	N/A		55	N/A	N/A
GCT39	Montreal	AGCT	Positive	Negative		53	N/A	N/A
GCT40	Montreal	AGCT	Positive	Negative		38	N/A	N/A
GCT41	Toronto General	AGCT	Positive	N/A		40	Ia	Primary
GCT43	Toronto General	AGCT	Positive	Negative		73	Ia	Primary
GCT44	Toronto General	AGCT	Positive	Negative		61	Ia	Primary
GCT47	Toronto General	AGCT	Positive	Negative		36	Ic	Recurrence
GCT48	Toronto General	AGCT	Positive	Negative		40	Ic	N/A
GCT49	Toronto General	AGCT	Positive	Negative		43	Ic	Recurrence
GCT50	Toronto General	AGCT	Positive	Negative		37	Ic	Recurrence
GCT51	Toronto General	AGCT	Positive	Negative		44	Ia	Primary
GCT52	Toronto General	AGCT	Positive	Negative		36	Ia	Recurrence
GCT53	Toronto General	AGCT	Positive	Negative		39	Ia	Recurrence
GCT54	Toronto General	AGCT	Positive	Negative		67	IIb	Recurrence
GCT55	Toronto General	AGCT	Positive	Negative		44	IIb	Primary
GCT56	Toronto General	AGCT	Positive	Negative		59	IIb	Primary
GCT57	Toronto General	AGCT	Positive	Negative		54	Ia	Primary
GCT58	Toronto General	AGCT	Positive	Negative		50	Ia	Primary
GCT59	Toronto General	AGCT	Positive	Negative	Recurrence associated with GCT76	37	Ia	Recurrence
GCT60	Toronto General	AGCT	Positive	Negative		66	Ia	Primary

GCT61	Toronto General	AGCT	Positive	N/A		26	Ia	Primary
GCT62	Toronto General	AGCT	Positive	Negative		51	Ia	Primary
GCT63	Ottawa	AGCT	Positive	N/A		35	I	N/A
GCT64	Ottawa	AGCT	Positive	N/A		52	I	N/A
GCT65	Ottawa	AGCT	Positive	N/A		62	II	N/A
GCT66	Ottawa	AGCT	Positive	N/A		54	I	N/A
GCT67	Ottawa	AGCT	Positive	N/A		53	IIb	N/A
GCT68	Australia	AGCT	Positive	Inconclusive		52	I	Primary
GCT71	Australia	AGCT	Positive	Negative		53	IIIc	Recurrence
GCT72	Australia	AGCT	Positive	Inconclusive		42	IIIc	Primary
GCT73	Australia	AGCT	Positive	N/A		57	N/A	Recurrence
GCT74	Australia	AGCT	Positive	Negative		32	IIIb	Primary
GCT76	Toronto General	AGCT	Positive	N/A		36	Ia	Recurrence
GCT79	Toronto General	AGCT	Positive	Negative		36	Ia	Primary
GCT11	VGH	AGCT	Negative	Negative		45	Ia	Primary
GCT22	VGH	AGCT	Negative	Negative		75	Ia	Primary
GCT33	VGH	AGCT	Negative	Negative		38	Ia	N/A
GCT24	VGH	JGCT	Negative	Negative		21	Ia	Primary
GCT45	Toronto General	JGCT	Negative	Inconclusive		35	Ia	Primary
FIRST SERIES: OTHER SEX CORD-STROMAL TUMORS								
GCT18	VGH	Thecoma	Positive	Negative		55	Ia	Primary
GCT21	VGH	Fibroma with sex cord elements	Negative	Negative		68	Ia	Primary
GCT42	Toronto General	Sertoli Cell tumor	Negative	Negative		22	Ia	Primary
FIRST SERIES: NON-GCT AND NON-SEX CORD-STROMAL TUMORS								
GCT69	Australia	High-grade neoplasm other than GCT	Negative	Negative		49	III	Primary
FIRST SERIES: INCONCLUSIVE CASES AND UNREVIEWED CASES								
GCT17	VGH	AGCT	Inconclusive	Negative	Laser capture microdissection	22	Ia	Primary
GCT20	VGH	AGCT	Inconclusive	N/A		90	IIb	Primary
GCT46	Toronto General	AGCT	Inconclusive	N/A	Laser capture microdissection	34	Ia	Primary
GCT70	Australia	AGCT	Inconclusive	N/A		48	N/A	Recurrence
GCT75	Australia	AGCT	Inconclusive	N/A		79	Ib	Primary
GCT80	Ottawa	N/A	Positive	N/A	No slide available for review	40	IIa	N/A
SECOND SERIES: GRANULOSA CELL TUMORS								
GCT81	England	AGCT	Positive					
GCT84	England	AGCT	Positive					

GCT85	England	AGCT	Positive		
GCT86	England	AGCT	Positive		
GCT87	England	AGCT	Positive		
GCT88	England	AGCT	Positive		
GCT89	England	AGCT	Positive		
GCT90	England	AGCT	Positive		
GCT91	England	AGCT	Positive		
GCT92	England	AGCT	Positive		
GCT93	England	AGCT	Positive		
GCT94	England	AGCT	Positive		
GCT95	England	AGCT	Positive		
GCT96	England	AGCT	Positive	Negative	
GCT97	England	AGCT	Positive	Negative	
GCT98	England	AGCT	Positive	Negative	
GCT99	England	AGCT	Positive		
GCT100	Australia	AGCT	Positive	Negative	
GCT101	Australia	AGCT	Positive	Negative	
GCT102	Australia	AGCT	Positive	Negative	
GCT103	Australia	AGCT	Positive	Negative	
GCT104	Australia	AGCT	Positive	Negative	
GCT105	Australia	AGCT	Positive	Negative	
GCT106	Australia	AGCT	Positive	Negative	
GCT107	Australia	AGCT	Positive		
GCT108	VGH	AGCT	Positive		
GCT109	VGH	AGCT	Positive		
GCT82	England	JGCT	Negative		
GCT83	England	JGCT	Negative		
SCST10	VGH	JGCT	Negative		
SCST11	VGH	JGCT	Positive		
SCST12	VGH	JGCT	Negative		
SCST18	VGH	JGCT	Negative		
SCST24	VGH	JGCT	Negative		
SCST27	VGH	JGCT	Negative		
		malignant			
SECOND SERIES: OTHER SEX CORD-STROMAL TUMORS					
SCST28	Pittsburgh	SLCT	Negative		
SCST29	Pittsburgh	SLCT	Negative		
SCST30	Pittsburgh	SLCT	Negative		
SCST31	Pittsburgh	SLCT	Negative		
SCST32	Pittsburgh	SLCT	Negative		
SCST33	Pittsburgh	SLCT	Negative		
SCST34	Pittsburgh	SLCT	Negative		
SCST35	Pittsburgh	SLCT	Negative		
SCST36	Pittsburgh	SLCT	Negative		
SCST37	Pittsburgh	SLCT	Negative		
SCST39	Pittsburgh	SLCT	Negative		
SCST40	Pittsburgh	SLCT	Negative		
SCST41	Pittsburgh	SLCT	Negative		
SCST42	Pittsburgh	SLCT	Negative		
SCST43	Pittsburgh	Steroid cell tumor	Negative		

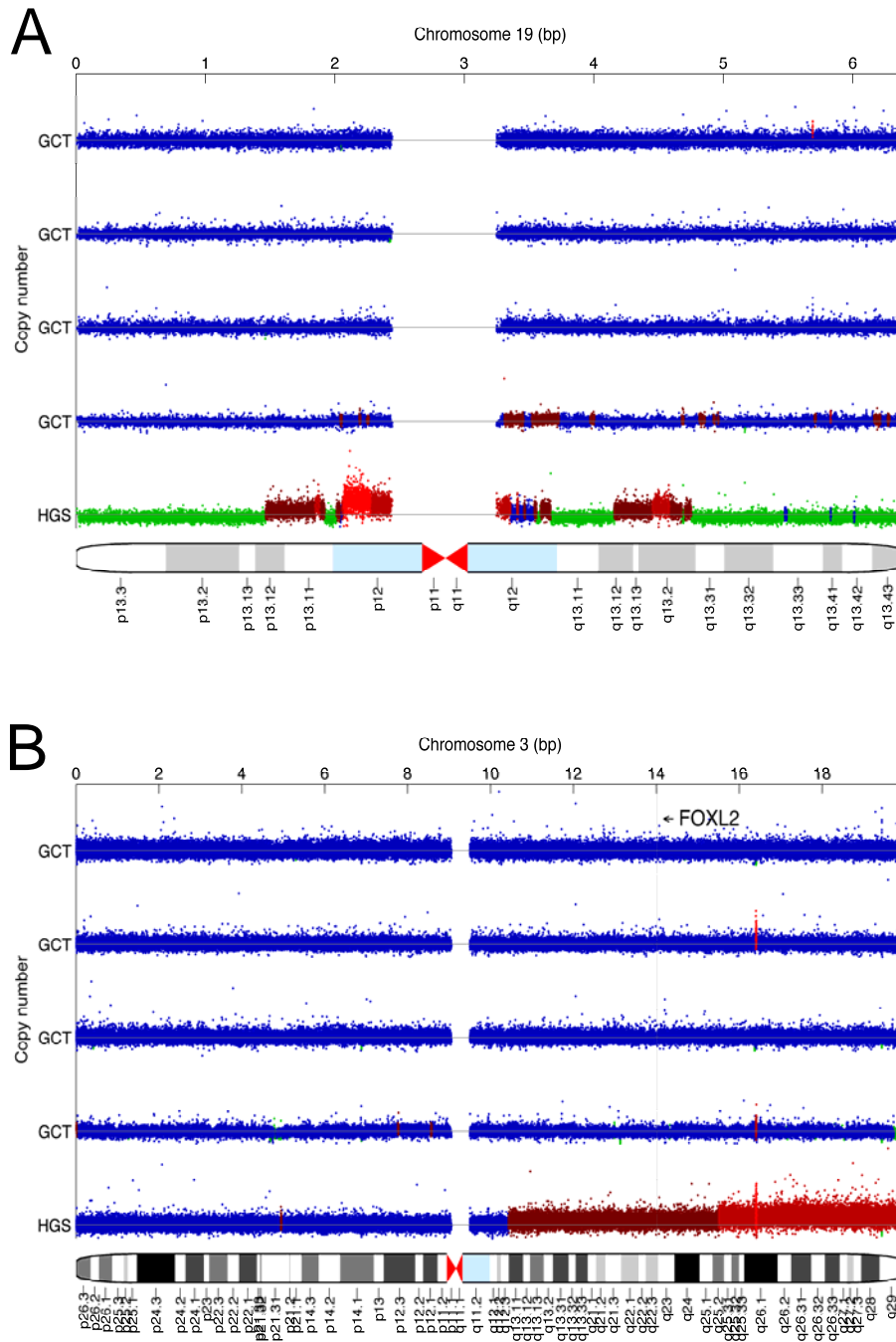
SCST44	Pittsburgh	Steroid cell tumor	Negative	
SCST45	Pittsburgh	Steroid cell tumor	Negative	
SCST46	Pittsburgh	Steroid cell tumor	Negative	
SCST48	Pittsburgh	Steroid cell tumor	Negative	
SCST49	Pittsburgh	Steroid cell tumor	Negative	
SCST50	Pittsburgh	Steroid cell tumor	Negative	
SCST52	Pittsburgh	Steroid cell tumor	Negative	
SCST53	Pittsburgh	Steroid cell tumor	Negative	
SCST54	Pittsburgh	Steroid cell tumor	Negative	
SCST56	VGH	Thecoma	Negative	
SCST59	VGH	Thecoma	Negative	
SCST60	VGH	Thecoma	Positive	
SCST61	VGH	Thecoma	Negative	
SCST63	VGH	Thecoma	Negative	
SCST64	VGH	Thecoma	Positive	
SCST66	VGH	Thecoma	Negative	
SCST69	VGH	Thecoma	Negative	
SCST62	VGH	Thecoma	Negative	
SCST67	VGH	Thecoma	Negative	
SCST96	Australia	Thecoma	Negative	
SCST97	Australia	Thecoma	Negative	
SCST98	Australia	Thecoma	Negative	
SCST70	VGH	Fibroma	Negative	
SCST70	VGH	Fibroma	Negative	
SCST71	VGH	Fibroma	Negative	
SCST72	VGH	Fibroma	Negative	
SCST73	VGH	Fibroma	Negative	
SCST74	VGH	Fibroma	Negative	
SCST75	VGH	Fibroma	Negative	
SCST76	VGH	Fibroma	Negative	
SCST77	VGH	Fibroma	Negative	
SCST78	VGH	Fibroma	Negative	
SCST79	VGH	Fibroma	Negative	
SCST80	VGH	Fibroma	Negative	
SCST81	VGH	Fibroma	Negative	
SCST82	VGH	Fibroma	Negative	
SCST83	VGH	Fibroma	Negative	
SCST84	VGH	Fibroma	Negative	
SCST85	VGH	Fibroma	Negative	
SCST86	VGH	Fibroma	Negative	
SCST87	VGH	Fibroma	Negative	
SCST88	VGH	Fibroma	Negative	
SCST89	VGH	Fibroma	Negative	
SCST92	Australia	Fibroma	Negative	

SCST94	Australia	Fibroma	Negative	
SCST95	Australia	Fibroma	Negative	

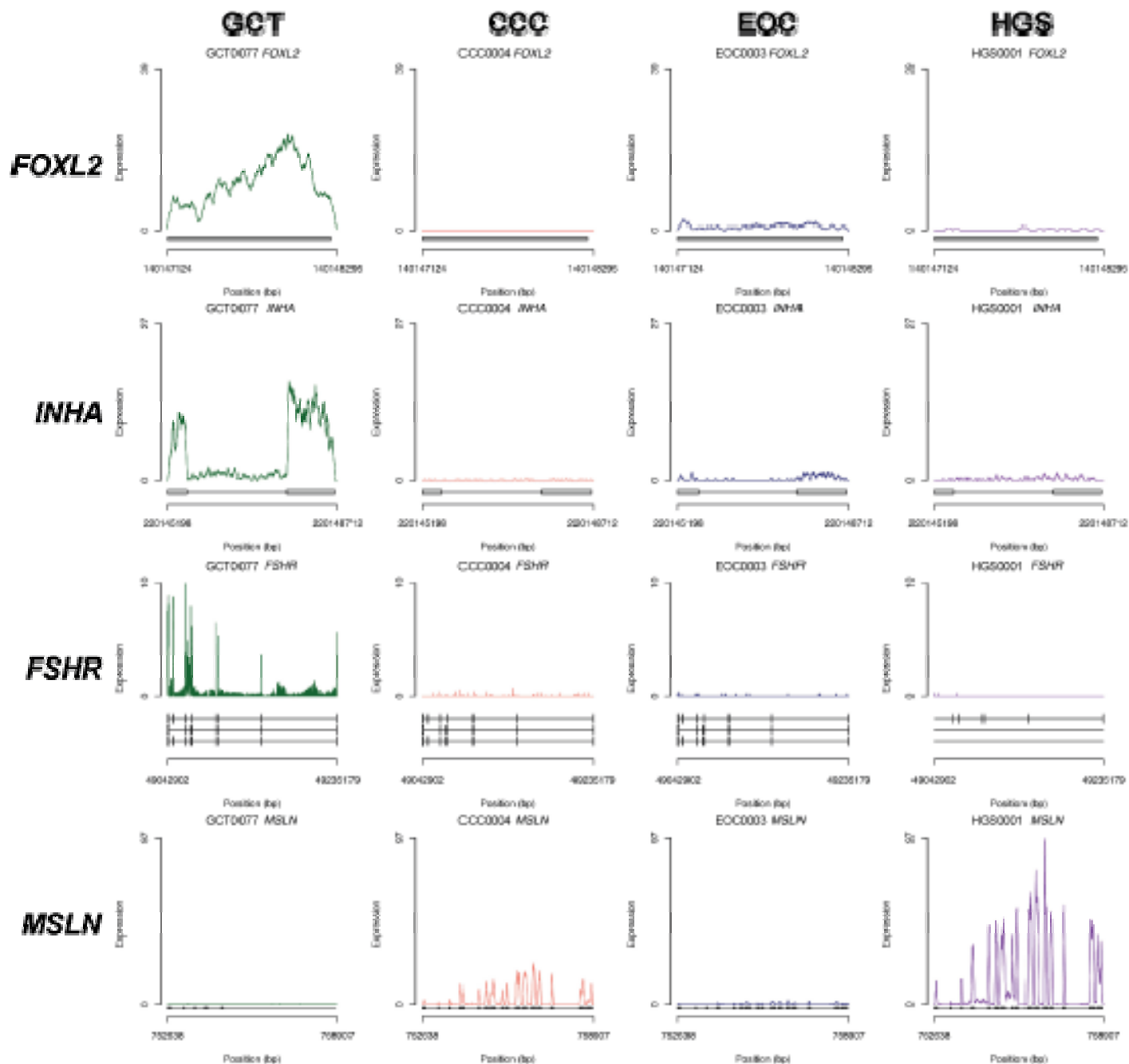
Supplemental Table 2. Summary of fluorescent *in situ* hybridization analysis of GCTs and non-GCT samples. BACs were used to generate a signal at the *FOXL2* locus which was normalized to the signal from the chromosome 3 centromeric probe (CEP3 probe: 3p11.1-q11.1). (*) indicates hemizygous/homozygous cases.

GCT ID	FOXL2 signal	CEP3 signal	Amplification ratio
GCT0001	2.5	3.1	0.8
GCT0002	4.1	4.6	0.9
GCT0003	3.4	3.4	1.0
GCT0004	3.2	3.7	0.9
GCT0005	3.5	3.7	0.9
GCT0006	3.5	4	0.9
GCT0007	3.7	4.6	0.8
GCT0008	3.6	4.6	0.8
GCT0009 (*)	2.1	3.7	0.6
GCT0010	3.8	4	1.0
GCT0011	2.8	3.2	0.9
GCT0012	4.4	3.9	1.1
GCT0013	3.8	4.6	0.8
GCT0014	4.5	4.8	0.9
GCT0015	4	5.5	0.7
GCT0016	3.2	3.6	0.9
GCT0018	3.2	3.8	0.8
GCT0019	3.2	3.6	0.9
GCT0020 (*)	3.8	4.9	0.8
GCT0021	4	4.7	0.9
GCT0022	4.3	4	1.1
GCT0023	3.2	4.9	0.7
GCT0024	3.1	2.6	1.2
GCT0025	3.2	3.2	1.0
GCT0026 (*)	3.3	3.3	1.0
GCT0027	4.3	4.9	0.9
GCT0029	3	3.7	0.8
GCT0030	2.8	4.2	0.7
GCT0031	3	4.9	0.6
GCT0032	3.4	4	0.9
GCT0033	3.9	4.6	0.8
GCT0078	1.6	2.2	0.7
Non-GCT	2.9	3.9	0.7
Non-GCT	2.6	2.6	1.0
Non-GCT	2.9	2.9	1.0
Non-GCT	3	2.5	1.2
Non-GCT	2.5	3.2	0.8

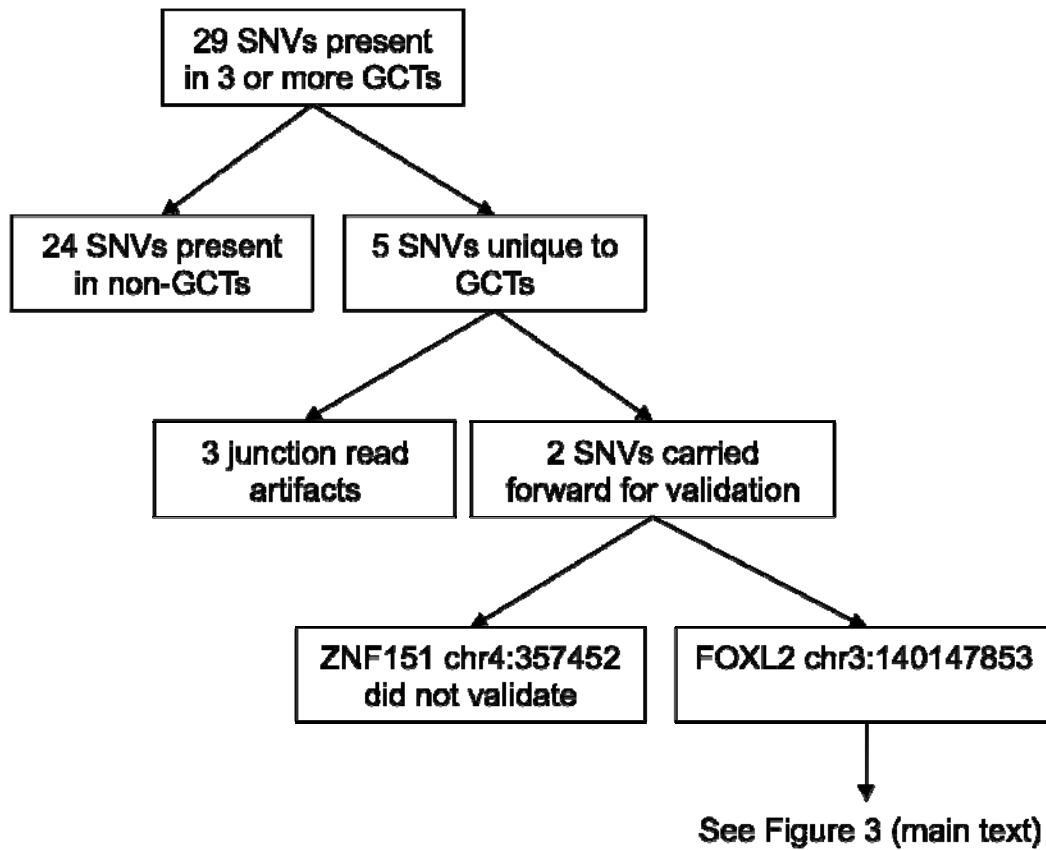
III. Supplemental Figures



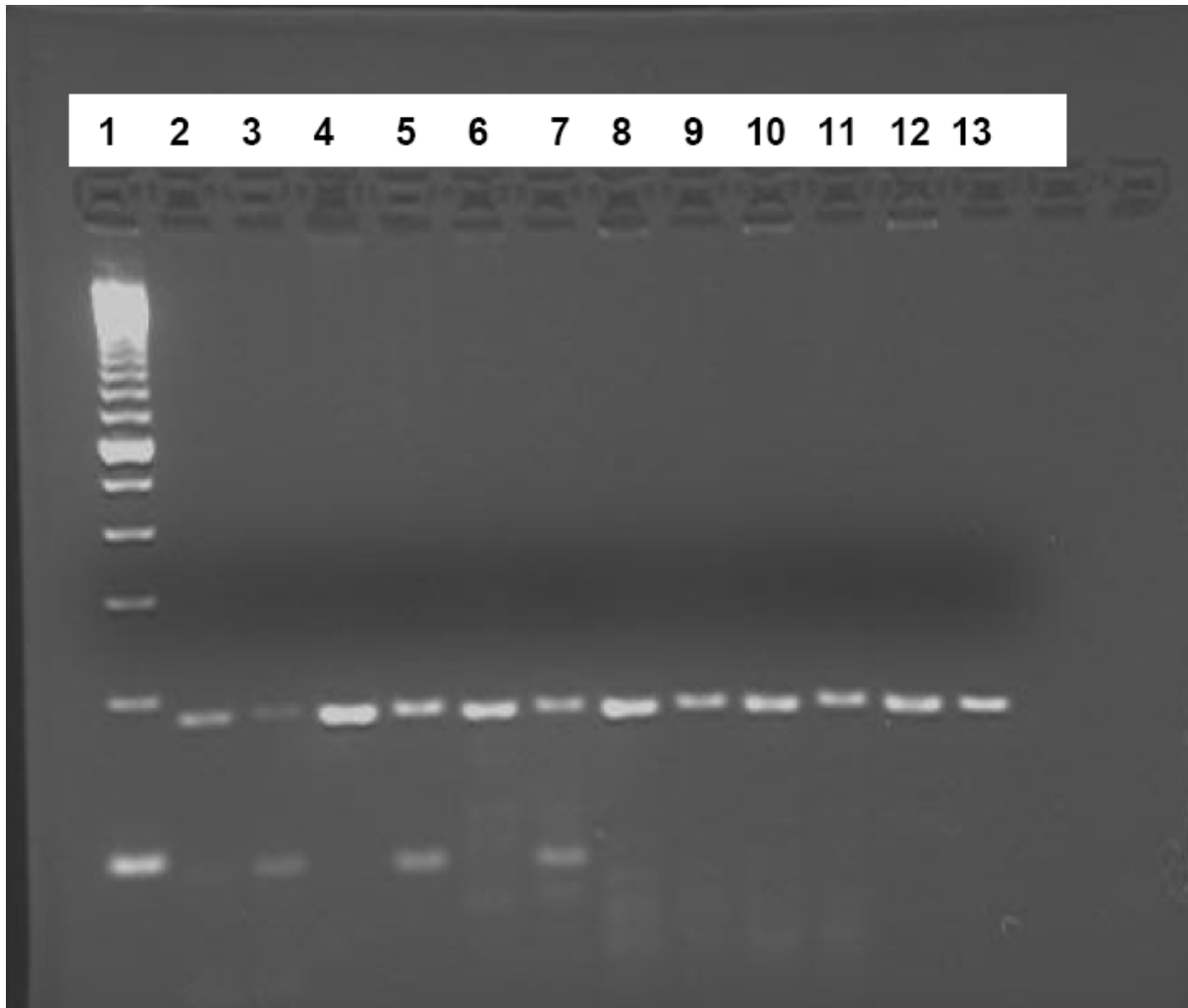
Supplemental Figure 1. Copy number assessment using high density genotyping arrays showed the lack of copy number alterations in the GCTs (four index cases) compared to a high grade serous cancer. Chromosome 19 (A) is used to illustrate this difference. The *FOXL2* gene is located on chromosome 3 (B).



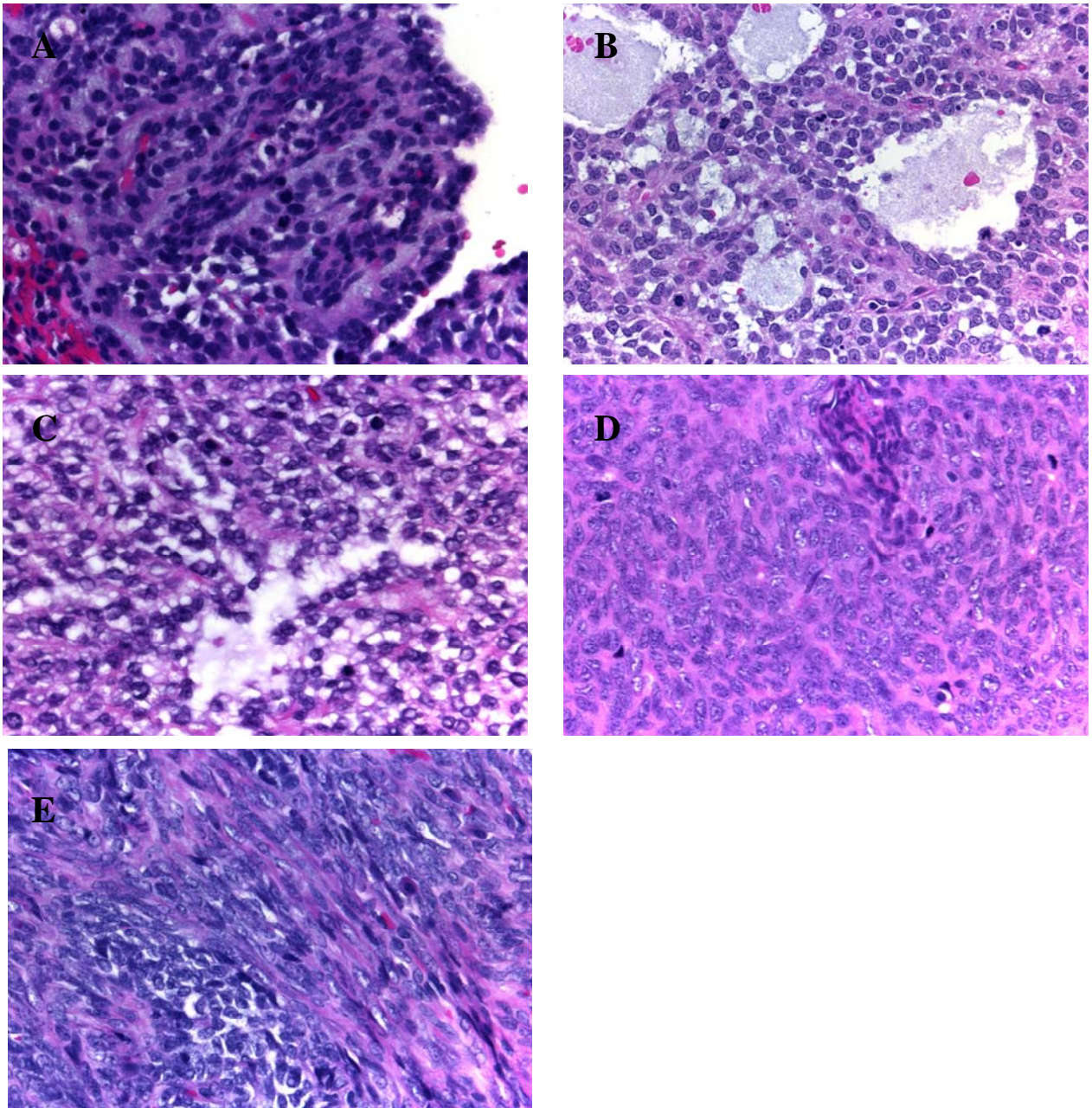
Supplemental Figure 2. Expression level (y-axis: shown as the number of reads mapping to each position of the gene) of four selected genes, FOXL2, InhibinA (*INHA*), Follicle Stimulating Hormone Receptor (*FSHR*) and Mesothelin (*MSLN*) in one index GCT case compared to three ovarian carcinomas of different subtypes (CCC – clear cell carcinoma, EOC – endometrioid ovarian carcinoma, HGS – high-grade serous carcinoma). *INHA* and *FSHR* are known to be expressed in GCTs and *MSLN* expression is a marker for serous differentiation¹¹.



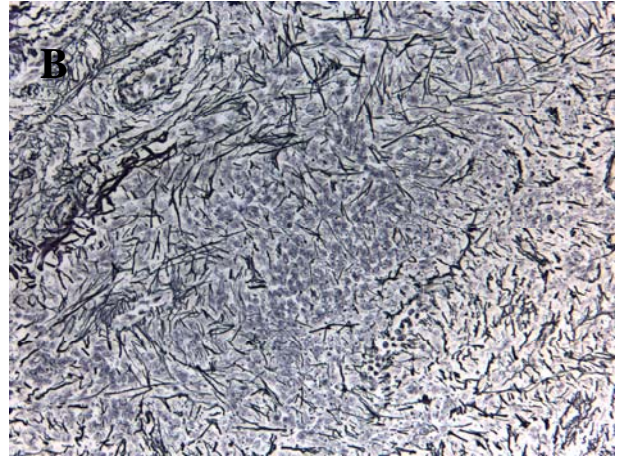
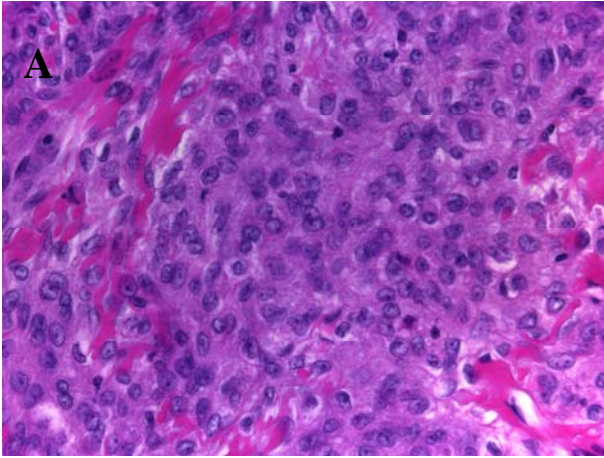
Supplemental Figure 3. Workflow diagram showing the triaging of SNVs to determine GCT-specific events.



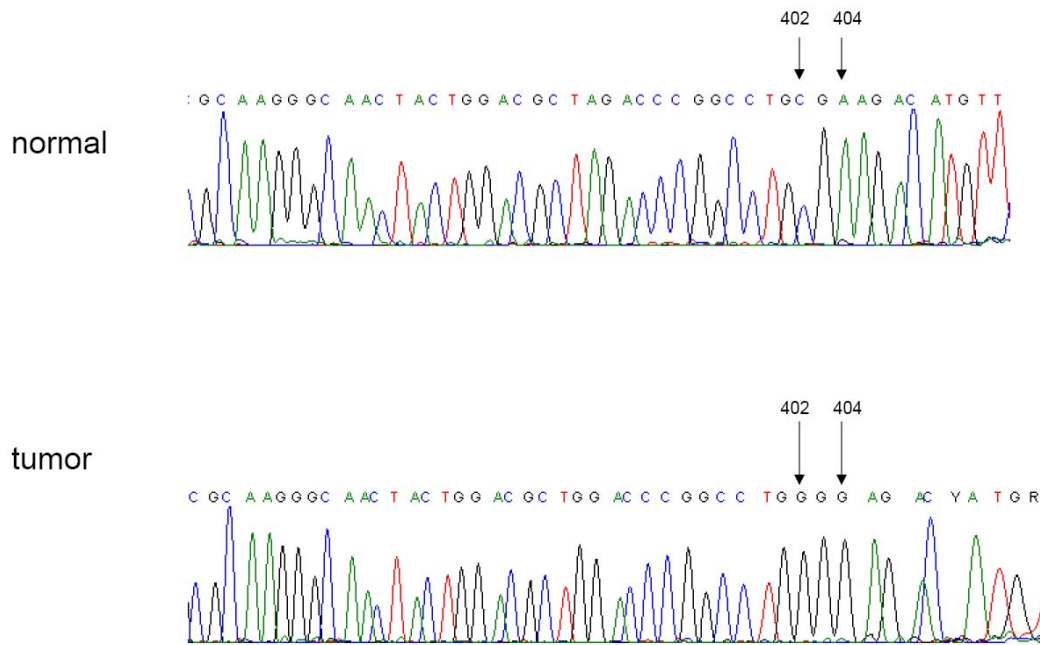
Supplemental Figure 4. Restriction fragment length polymorphism analysis showing three mutation positive GCTs and three mutation negative non-GCT ovarian carcinomas. Lanes from left to right: 1) 100bp ladder, 2) GCT26 uncut PCR product, 3) GCT26 PCR product digested with BstNI, 4) GCT28 uncut PCR product, 5) GCT28 PCR product digested with BstNI, 6) GCT78 uncut PCR product, 7) GCT78 PCR product digested with BstNI, 8) Mucinous cancer uncut PCR product, 9) Mucinous cancer PCR product digested with BstNI, 10) Endometrioid cancer uncut PCR product, 11) Endometrioid cancer PCR product digested with BstNI, 12) Mucinous cancer uncut PCR product, 13) Mucinous cancer PCR product digested with BstNI



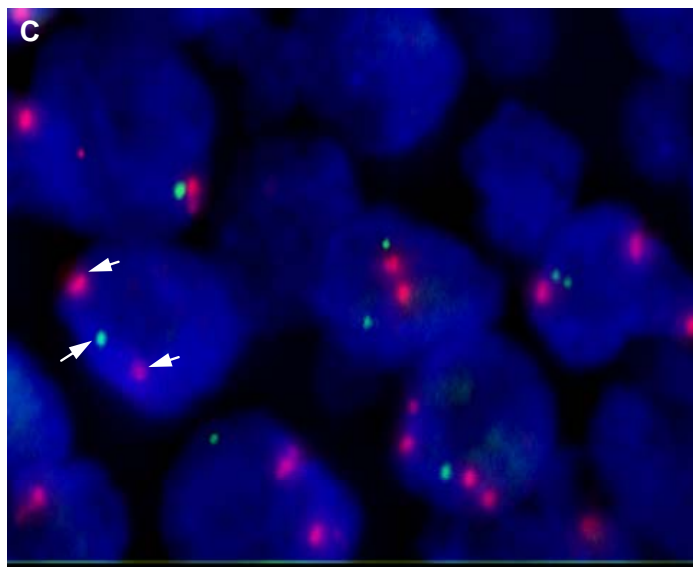
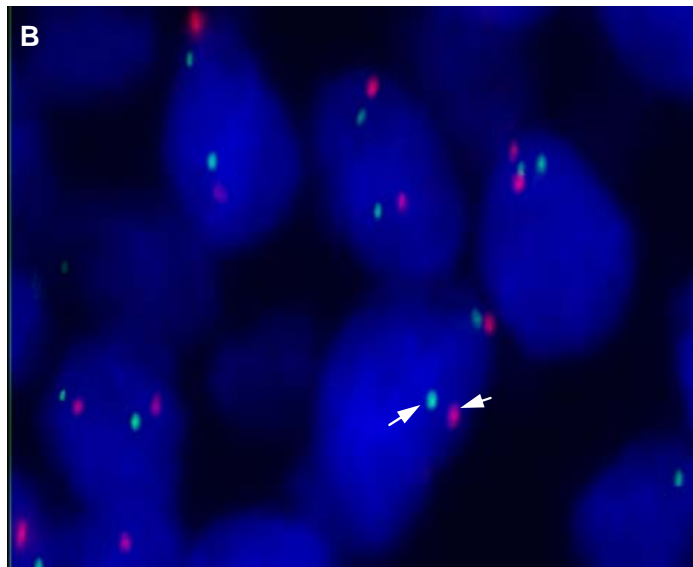
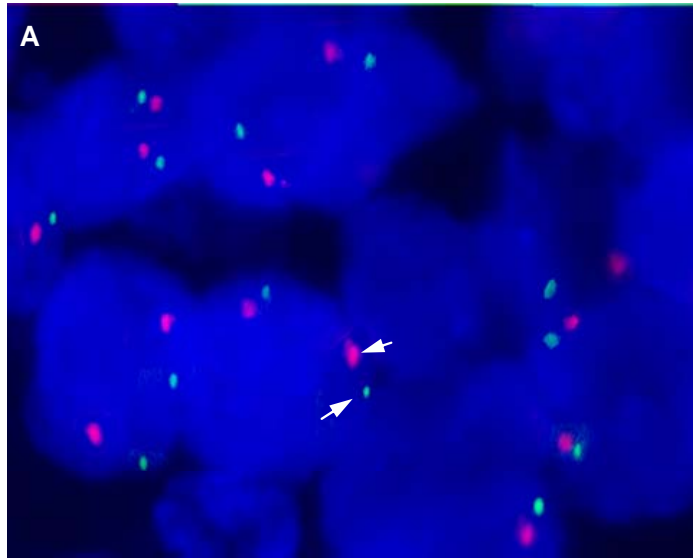
Supplemental Figure 5. Histopathology of the five mutation negative GCTs. (A) GCT24, (B) GCT45, The only two juvenile-type GCTs in the study show on morphological examination a diffuse pattern with numerous follicles which contain basophilic secretions. Nuclei are eu- to hyperchromatic lacking the typical nuclear grooves of an adult-type GCT. Mitotic rate is increased. (C) GCT11, (D) GCT22, Both tumors were considered to be consistent with GCTs, however, they lack typical features. INHA was not expressed in both tumors by immunohistochemistry, in contrast to all other GCTs available for tissue microarray construction. (E) GCT33, Nests of granulosa cells merging with a prominent fibrothecomatous component, a feature that was also observed in two mutation positive



Supplemental Figure 6. Histopathology of the mutation positive non-GCT. (A) Histopathological features of the case GCT18 that showed indeterminate features between a thecoma and a granulosa cell tumor on H&E (400x). (B) Reticulin stain (400x) reveals fibrils surrounding nests or larger areas of granulosa cells which is diagnostically used for GCTs.

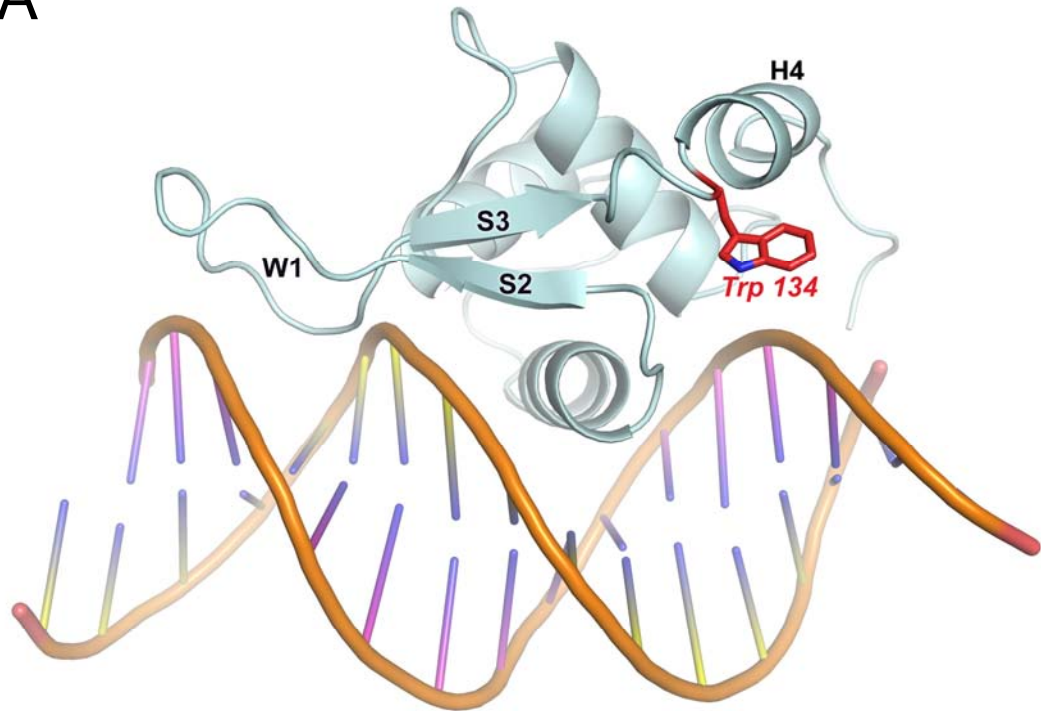


Supplemental Figure 7. GCT44 with a double mutation 402C>G and 404 A>G in the *FOXL2* gene.

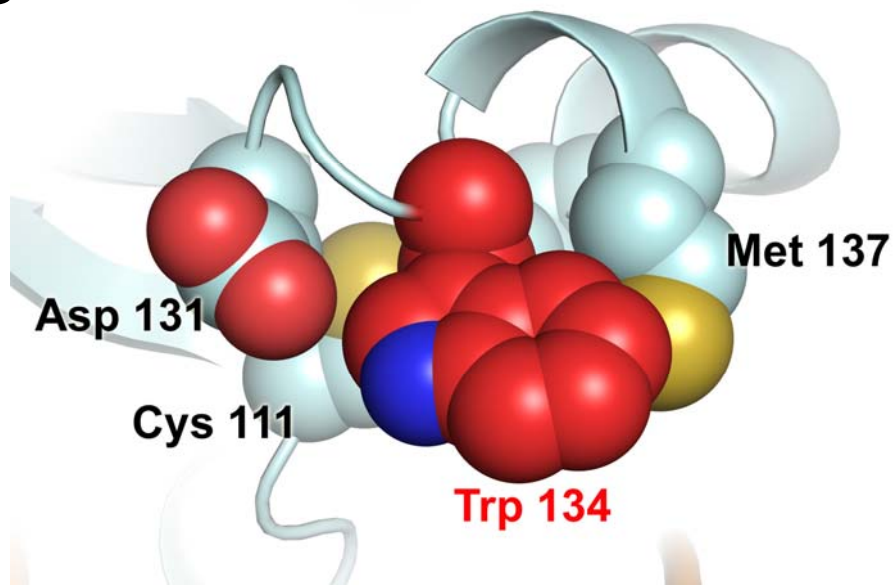


Supplemental Figure 8. FISH images of GCT and non-GCT samples. FISH assays were performed using two BACS to visualize the *FOXL2* locus (green) and a CEP3 probe to visualize the centromeric region of chromosome 3 (red). (A) a non-GCT case; (B) a GCT case heterozygous for the *FOXL2* mutation; (C) a GCT case (GCT9) hemizygous/homozygous for the *FOXL2* mutation. Arrows indicate green (*FOXL2* locus) and red (chromosome 3 centromere) signals.

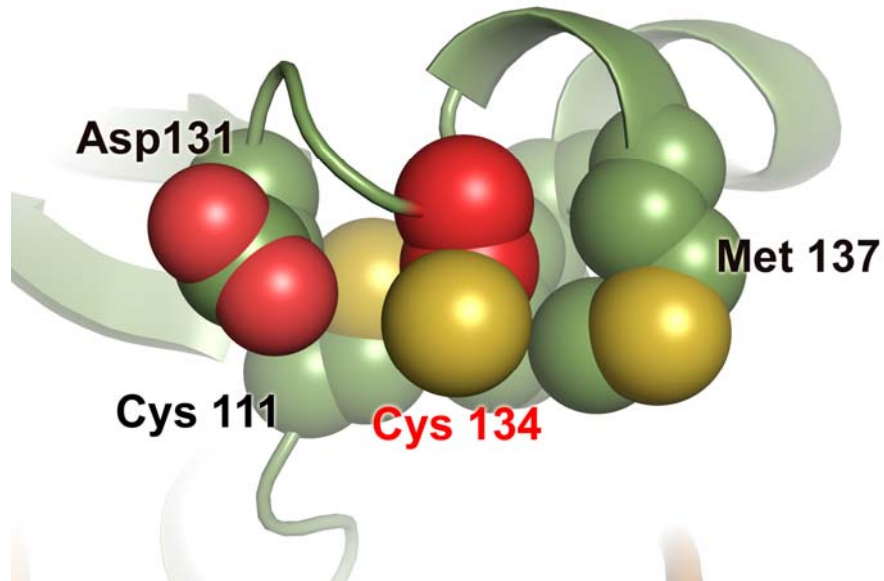
A



B



C



Supplemental Figure 9. Homology model of the human FOXL2 forkhead domain bound to DNA. (A) Overview of the model of the C134W mutant, highlighting the position of Trp 134. Key secondary structure elements, wing 1 (W1), b-strands 2 and 3 (S2 and S3) and helix 4 (H4) are labelled. (B) Details of the packing of Trp 134 on the surface of the FOXL2 forkhead domain. (C) Details of the packing of the wild type residue, Cys 134, on the surface of the FOXL2. Note that no rearrangements of the surrounding residues are required to accommodate the C134W mutation

IV References

1. Morin R, Bainbridge M, Fejes A, et al. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques* 2008;45:81-94.
2. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 2008;18:1851-8.
3. Carvalho B, Bengtsson H, Speed TP, Irizarry RA. Exploration, normalization, and genotype calls of high-density oligonucleotide SNP array data. *Biostatistics* 2007;8:485-99.
4. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008;452:872-6.
5. Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. *PLoS Biol* 2007;5:e254.
6. Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 2008;456:53-9.
7. Wang J, Wang W, Li R, et al. The diploid genome sequence of an Asian individual. *Nature* 2008;456:60-5.
8. Woo MM, Salamanca CM, Miller M, et al. Serous borderline ovarian tumors in long-term culture: phenotypic and genotypic distinction from invasive ovarian carcinomas. *Int J Gynecol Cancer* 2008.
9. Bengtsson H, Irizarry R, Carvalho B, Speed TP. Estimation and assessment of raw copy numbers at the single locus level. *Bioinformatics* 2008;24:759-67.
10. Makretsov N, He M, Hayes M, et al. A fluorescence in situ hybridization study of ETV6-NTRK3 fusion gene in secretory breast carcinoma. *Genes Chromosomes Cancer* 2004;40:152-7.
11. Kobel M, Kalloger SE, Boyd N, et al. Ovarian Carcinoma Subtypes Are Different Diseases: Implications for Biomarker Studies. *PLoS Med* 2008;5:e232.