The driver mutational landscape of ovarian squamous cell carcinomas arising in mature cystic teratoma

Susanna L. Cooke[1*], Darren Ennis[1*], Lisa Evers[1], Suzanne Dowson[1], Mei Yen Chan[1], James Paul[2], Lynn Hirschowitz[3], Rosalind Glasspool[4], Naveena Singh[5], Sarah Bell[6], Elizabeth Day[6], Agata Kochman[7], Nafisa Wilkinson[8], Philip Beer[9], Sancha Martin[1], David Millan[6], Andrew V. Biankin[1], Iain A. McNeish[1,4^] on behalf of the Scottish Genomes Partnership.

1. Institute of Cancer Sciences, University of Glasgow, Glasgow, UK
2. Cancer Research UK Clinical Trials Unit, Glasgow, UK
3. Department of Pathology, Birmingham Women's NHS Trust, Birmingham, UK
4. Beatson West of Scotland Cancer Centre, Glasgow, UK
5. Department of Pathology, Barts Health NHS Trust, London, UK
6. Department of Pathology, Queen Elizabeth University Hospital, Glasgow, UK
7. Department of Pathology, Monklands Hospital, Airdrie, UK
8. Department of Pathology, University College London Hospital, London, UK
9. Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

*: equal contribution

^: to whom correspondence should be addressed:
Professor Iain McNeish
Institute of Cancer Sciences, University of Glasgow
Wolfson Wohl Cancer Research Centre
GLASGOW
G61 1QH
iain.mcneish@glasgow.ac.uk

Running title

Squamous cell carcinoma arising in ovarian mature cystic teratoma

Keyword

Ovary; mature cystic teratoma; squamous cell carcinoma; *TP53*; bi-allelic

Conflict of interest statement

The authors have no conflicts of interest to declare.

Statement of translational relevance

Ovarian squamous cell carcinomas (SCC) arising in mature cystic teratomas (MCT) are rare malignancies of poor prognosis. There are no published data on mutations in these tumors to guide future possible clinical trials. Using archival samples from four large UK gynecologic cancer centers, we have performed next generation sequencing on 25 SCC samples with their associated MCT. MCT had few copy number or single nucleotide variants. SCC were all HPV negative, but had high mutation burden, with frequent abnormalities in *TP53, PIK3CA* and *CDKN2A*, at frequencies similar to lung SCC. Strikingly, 40% of the *TP53* mutations were bi-allelic, which may be associated with improved outcome. This is the first genomic analysis of these rare tumours. Our data suggest that patients with SCC arising in ovarian MCT could be included in any SCC-specific trial and may, like lung SCC, benefit from immune checkpoint inhibition.

Abstract

Purpose

We sought to identify the genomic abnormalities in squamous cell carcinomas (SCC) arising in ovarian mature cystic teratoma (MCT), a rare gynaecological malignancy of poor prognosis.

Experimental design

We performed copy number, mutational state and zygosity analysis of 151 genes in SCC arising in MCT (n=25) using next-generation sequencing. The presence of high/intermediate risk HPV genotypes was assessed by quantitative PCR. Genomic events were correlated with clinical features and outcome

Results

MCT had a low mutation burden with a mean of only 1 mutation per case. Zygosity analyses of MCT indicated four separate patterns, suggesting that MCT can arise from errors at various stages of oogenesis. A total of 244 abnormalities were identified in 79 genes in MCT-associated SCC, and the overall mutational burden was high (mean 10.2 mutations per megabase). No SCC was positive for HPV. The most frequently altered genes in SCC were *TP53* (20/25 cases, 80%), *PIK3CA* (13/25 cases, 52%) and *CDKN2A* (11/25 cases, 44%). Mutation in *TP53* was associated with improved overall survival. In 8/20 cases with *TP53* mutations, two or more variants were identified, which were bi-allelic.

Conclusions

Ovarian SCC arising in MCT has a high mutational burden with *TP53* mutation the most common abnormality. The presence *TP53* mutation is a good prognostic factor. SCC arising in MCT share similar mutation profiles to other SCC. Given their rarity, they should be included in basket studies that recruit patients with SCC of other organs.

Introduction

Mature cystic teratoma (MCT) of the ovary (also known as dermoid cyst and benign cystic teratoma) is a common benign gynaecological tumour, usually arising before the menopause (1). The origin of MCT remains uncertain. Karyotype analyses suggested that there may be five separate types arising from different replication errors during meiosis (2) and that MCT may represent primary oocytes that have escaped from meiotic arrest (3).

Because MCT contain all three germ cell layers, they often display multiple differentiated tissue types, including teeth, bone and hair. However, secondary transformation into invasive malignancy can also occur, at reported rates of 0.1 - 1% (4, 5). Up to 80% of these transformed teratomas contain squamous cell carcinoma (SCC), with the remaining 20% containing adenocarcinoma, thyroid carcinoma or carcinoid tumours (6).

Due to their rarity, there are few large published series and no prospective clinical trials in MCT-derived SCC, and the recent Gynecologic Cancer Intergroup (GCIG) consensus review concluded that there were insufficient data to provide clear guidance on optimal treatment (7). However, the prognosis for these SCC is poor, especially for stage II - IV disease [summarised in (6)]. As a consequence, there is a need to characterise the genomic features of these tumours, to identify opportunities for patients to enrol in clinical trials.

Here we present analysis of the mutational state, copy number and zygosity of 151 cancer genes, as well as HPV status, in 25 cases of MCT with SCC from four large UK gynaecological cancer centres.

Patients and methods

Study conduct, survival analyses and patient samples

All samples were acquired and utilised under the auspices of the NHS Greater Glasgow and Clyde Biorepository following approval by West of Scotland Research Ethics Committee 4 (Reference 10/S0704/60). Overall survival was calculated from the date of diagnosis to the date of death or the last clinical assessment where known. Overall survival was calculated by log-rank test (Mantel-Cox) using Prism v6.0 (GraphPad, San Diego, CA), whilst multivariate analysis was calculated using a Cox Proportional Hazard model in SPSS; $p<0.05$ was considered to be significant.

Formalin-fixed paraffin-embedded specimens were identified from the pathology archives of participating centres. Following review by an expert gynaecological pathologist, areas of squamous cell carcinoma (SCC), mature cystic teratoma (MCT) and normal tissue were marked for macro-or microdissection. Laser capture microdissection (LCM) was performed on consecutive cresyl violet-stained sections using a Leica CTR6500 microscope (Leica Microsystems, Milton Keynes, UK). Sections were collected onto LCM-compatible polyethylene terephthalate (PET) frame slides (Leica Microsystems, Milton Keynes, UK).

DNA extraction and quantification.

DNA was extracted from 10 x 10 µm sections using QIAmp DNA FFPE Tissue Kit (Qiagen, UK) according to manufacturer's protocol. For macrodissected samples, paraffin was removed by the xylene/ethanol method. For microdissected samples, paraffin was removed by the heptane method. Briefly, 500 µl heptane was added to the sample, followed by 10 mins incubation at room temperature. 25 µl methanol was added, followed by centrifugation and aspiration. 1ml of absolute ethanol was then added to the pellet, followed by centrifugation and aspiration. The pellet was then air dried. From here the standard QIAmp extraction method was followed as per manufacturer's protocol. DNA was quantified on a Qubit 2.0 Fluorometer (Life Technologies, Paisley, UK).

HPV Analysis

High/intermediate risk HPV serotypes (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68) were assessed in 50ng SCC DNA using real-time PCR [HPV High Risk Taqman PCR Kit (Norgen BioTek, Thorold, ON, Canada)] using a CFX96 Touch Real-Time PCR System (BioRad, Watford, UK).

Sequencing and bioinformatic analyses

50-200ng DNA was sheared with a Covaris LE220 focused-ultrasonicator (Covaris, Woburn, MA) to produce 100-200bp fragments. Libraries were generated using SureSelect XT standard protocol (Agilent Technologies, Santa Clara, CA) for low-input and FFPE samples. Pre-capture libraries were quantified as above, then assessed for library size and impurities with a High Sensitivity DNA BioAnalyzer Chip (Agilent Technologies). Libraries were captured with 120nt biotinylated RNA baits designed for the ClearSeq Comprehensive Cancer Panel (Agilent Technologies), covering 151 cancer-associated genes (Table S1). Hybridization for capture libraries less than 3Mb was performed according to manufacturer's protocol. Post-capture PCR incorporated primers with unique 8-bp indexes (Agilent Technologies) to facilitate multiplexing. Final captured-libraries were quantified with a Qubit Fluorometer High Sensitivity DNA Assay (Invitrogen) and assessed for size distribution and quality on a High Sensitivity DNA BioAnalyzer Chip (Agilent Technologies). Pooled capture-libraries were sequenced on a NextSeq 500 instrument (Illumina, San Diego, CA) or MiSeq instrument (Illumina) with 2 x 76-bp, paired-end reads according to manufacturer's instructions.

A full description of mapping, QC, point mutation and indel calling, copy number calling (Table S2) and zygosity analysis is provided in Supplementary Methods.

Results

Patients and samples

Thirty-one cases were originally identified. Samples from three patients were missing. There were low DNA yields (<50ng) in two cases, and sequencing failed in one SCC case, leaving 77 samples from 25 cases that were successfully sequenced (Fig. 1; see Table S3 for QC). Twenty cases had sample trios (normal, MCT, SCC; Table 1), five of which had two or more regions of SCC and one of which had two MCT samples. Four cases had normal-SCC pairs and one case had MCT-SCC with no matching normal. In the sequenced cases, median age at diagnosis was 51.0 years (range 25 - 86) and 56% (14/25) had stage I disease (Table 1).

Mutation and copy number analyses

Figure 2 shows a summary of key mutation and copy number alterations in both SCC and MCT - full details are given in Tables S4 and S5. MCT were genomically quiet, with few mutations per sample (median 0, mean 1, range 0 - 7), and only one sample showed high-level copy number alterations – amplification of *MYC* and *EGFR* in SC025 (Fig. S1A), the latter of which was shared with the arising SCC. Two further MCT (SC006 and SC013) showed evidence of low-level copy number gains (Fig. S1B), while the remaining 18 had normal copy number profiles (Fig. S1C).

SCC samples contained a median of 8 mutations per case (range 0 - 39), with an overall median mutational burden of 10.2 mutations per MB of sequenced DNA (range 0 - 49). This is similar to that of lung SCC (8.4 per MB (8, 9)), and higher than head and neck SCC (3.2-5.0/MB (10, 11)), but lower than skin SCC, which has one of the highest mutation rates of any malignancy (at least 50 mutations/MB (12)). All cases were negative for 13 high/intermediate risk HPV genotypes (Table S5).

The most commonly altered genes were *TP53* (20/25 cases, 80%) (Fig. 2B), including 4/20 (20%) cases with mutations in codon 285 (p.E285L/Q), *PIK3CA* (13/25 cases, 52%) and *CDKN2A*, both by homozygous deletion (Table S6) and by point mutation (11/25 cases, 44%), a pattern similar to lung (8) and skin SCC (12). Many SCCs showed evidence of highly rearranged genomes, with multiple low-level gains and losses (Fig. S2A). Amplification of MYC was called in two cases (SC005, SC021) (Table S6) and gain of *MYC*, falling just below the threshold for amplification calling, was seen in a further three cases

(SC011, SC020, SC025), making the potential prevalence as high as 20% (5/25) (Fig. S2B). The only other recurrent amplification was JAK2 in two cases (SC020 and SC025) (Table S6).

In five cases, there were two or more separate SCC samples. One of these (SC030) contained no detectable mutations, but in the remaining four cases, the *TP53* mutation was an early and shared event, even in samples that showed some heterogeneity between samples (Table S6). Two multi-sample cases (SC022, SC031) showed identical mutation profiles across samples, whilst the remaining two (SC014, SC025) showed some overlap: SC014 shared two mutations (*TP53, GNAQ)* with *PIK3CA* and *CDKN2A* changes in one sample only. SC025 had four mutations in one sample and three in the second, but with only the *TP53* variant being shared, and therefore the only one occurring prior to divergence of the two clones. Where there were multiple mutations in one gene (e.g. *DDX3X* in SC022; *APC* in SC031), there was complete concordance between samples.

TP53 mutations showed a striking pattern with 8/20 (40%) mutated cases containing two or more variants (Fig. 2A). Several pieces of evidence suggested strongly that these abnormalities are bi-allelic rather than resulting from the presence of two or more discrete clonal populations or the existence of kataegis-like clusters. Firstly, in SC006 the mutations were sufficiently close to be phased, but were not detected in the same sequencing read pair and thus must lie on separate alleles (Fig S3A). While the allele frequencies (18.4% and 13.1%) and copy number state (neutral, $\log_2$ ratio shift of 0.05) do not preclude the possibility that these mutations are in separate clones, the heterogeneity analysis of the multi-region sampled SCCs shows that even where sub-clones are present, they share the same *TP53* mutation (Table S7). Secondly, high *TP53* mutant allele frequencies in two cases can only be reconciled if the mutations are in the same cells. SC013 has *TP53* mutation frequencies of 47.2% and 55.8%. If these mutations are *in trans*, they must account for 100% of *TP53* alleles and thus cannot be in separate cells unless those cells are either homozygously mutated (and therefore bi-allelic for the same mutation) or have loss of the second copy to remove all wild-type alleles. However, the copy number state for *TP53* in SC013 is a $\log_2$ ratio shift of -0.35, which, at most, would indicate single-copy loss of *TP53* in a fraction of cells. The possibility that these mutations are *in cis* cannot be excluded here. The same logic holds for SC021, which has *TP53* mutation frequencies of 36.4% and 46.3% and a *TP53* copy number $\log_2$ ratio shift of 0.11, indicating no copy number alteration. Thirdly, two separate mutations were identified in SC005 at the same allele frequency at the same position (R280K at 21.4%, R280T at 22.9%), along with a third mutation (E285K) 14bp away that is in phase with the R280K mutation but always on a different allele to the R280T (Fig

S3). There is a fourth, splice region *TP53* mutation in this case that was too far away to phase with the other mutations. The copy number state for *TP53* in this case again indicated no copy number alteration (log_2 ratio shift of 0.21). Possible explanations are therefore that there are two identically sized subclones, one of which has either one or two *TP53* mutations and the other has either two or three, depending on how the distant mutation phases (although it is important to note that the three subclone explanation is not possible in the absence of copy number alteration). Alternatively, all *TP53* mutations exist in the same cell with the splice mutation on the same allele as either the E285K/R280K or the R280T mutations. The only explanation common to all three of the scenarios described above is that the *TP53* mutations in these cases are bi-allelic.

There were other examples, albeit rare, of genes with multiple mutations, in a pattern consistent with an APOBEC signature (9). Two samples (SC001 and SC005) had more than two *TP53* mutations, SC022 had three mutations in *DDX3X*, SC002 had three *PIK3CA* mutations and SC009 had four mutations in *NOTCH1*. These clusters of mutations were mainly C>T (12/19) and C>G (4/19) but, in a TpC context rather than the CpG context that would indicate the ubiquitous cytosine deamination signature. However, numbers are small and sequencing of a larger genomic footprint would be required to capture enough mutations to carry out a formal signature analysis.

The *PIK3CA* alterations were canonical activating mutations, with the two most frequent mutations (p.E545K and p.E542K) corresponding to the two most frequent hotspots in COSMIC (Catalogue of Somatic Mutations in Cancer). *CDKN2A* was inactivated by a variety of mechanisms, the commonest being a large-scale deletion (8 cases), with nonsense mutation occurring in a single case and missense mutations in three cases. The missense mutations all occurred at mutation peaks in COSMIC, consistent with these events being under positive selective pressure and therefore inactivating.

Other recurrent changes included mutation of *SMARCA4* and *KMT2A* (4 cases each), gain of *MYC* (3 cases) and gain of *JAK2* (2 cases). Driver point mutations, as opposed to structural changes, in *KMT2A* are relatively rare, as evidenced by its point mutation profile in COSMIC, which shows a very low level of mutation spread evenly across the whole length of the gene. *KMT2A* has a relatively long coding sequence, making it more likely that mutations occur in it by chance. Given the size of our dataset we are unable to determine whether these KMT2A mutations are part of the so-called ‘long tail’ of driver mutations seen in many cancer types or if they represent the background mutation rate. The *SMARCA4* mutations are likely passengers as none of them matched the five most recurrent variants in

COSMIC. Mutations in *BRCA1* and *BRCA2* were identified in 2 cases each; all mutations were somatic missense alterations (Table S4).

Clinical correlation

Follow up data were available on 23 cases. Twelve patients died and median overall survival was 20.1 months (Fig. 3A), with no deaths occurring more than 24 months following diagnosis. The median follow-up for living patients was 55.8 months (range 6.8 – 172.6 months) with 91% followed up for at least 41.6 months. As previously described, survival was significantly better for stage I compared to stage II-IV disease (Fig. 3B; HR = 0.301; p=0.021). There was no difference in survival based on age at diagnosis (Fig. 3C), *PIK3CA* or *CDKN2A* mutation status (Fig. S4A), but there was a statistically significant association with *TP53* mutation (HR = 0.178; p=0.002; Fig. 3D). Exploratory analysis suggested that cases with bi-allelic *TP53* mutations had significantly better overall survival than those with mono-allelic mutations (HR = 0.140; p=0.029; Fig. S4B). The association of survival with stage and *TP53* mutation status remained statistically significant (p=0.047 and p=0.011 respectively) in multivariate analysis (Table 2).

Zygosity analysis

In order to determine the stage of oogenesis that gave rise to the MCT, we analysed SNPs in the 151 sequenced genes for heterozygosity, and for a change in zygosity between normal tissue and MCT. All cases had some homozygous SNPs, but analysis of non-homozygous SNPs revealed four groups (Table 3, Fig. S5). Group A (10/21, 48%) showed the same levels of heterozygosity in normal and MCT, suggesting that they arise from cells prior to crossover (e.g. primordial germ cells or oogonia), whilst group B (3/21 cases, 14%) showed substantially fewer heterozygous positions in the MCT compared to normal. This is consistent with development after prophase I in meiosis I, when chromosomal crossover occurs. Group C (3/21,14%) had SNPs at the expected 0:1 and 0.5:0.5 ratios for diploid cells, but also had SNPs with other allelic ratios. SC011 and SC002, for example, showed SNPs at an allelic ratio of approximately 0.3:0.7. Given the absence of any copy number alterations in these samples, this would indicate a tetraploid MCT with some regions of the genome in a 2:2 maternal:paternal ratio and some in a 1:3 ratio. SC013 showed an additional allelic ratio of approximately 0.2:0.8, consistent either with some regions being at 6N in a 1:5 ratio between the maternal and paternal chromosomes or contamination of a group B-type profile with normal cells. SC013 did show some regions of copy number change, making it difficult to eliminate this as an explanation of the altered SNP allele

frequencies. However, in the other two MCTs with copy number alterations, there were insufficient SNPs within the limited regions of copy number change to impact on the peaks of SNP allele frequencies (Fig. S1B, Fig. S6, Table 3). Group D (6/21, 29%) did not have any SNPs in a 0.5:0.5 ratio but instead each showed a single allelic ratio consistent with a polyploid genome with unequal numbers of maternal and paternal chromosomes. For example, SC001, SC006 and SC030 all appeared to be pentaploid in a 2:3 ratio, while SC022, SC007 and SC029 were triploid (1:2 ratio), tetraploid (1:3 ratio) and hexaploid (1:5 ratio) respectively.

Discussion

We believe this to be the first description of the mutational landscape of squamous cell carcinoma (SCC) arising within mature cystic teratoma of the ovary (MCT, also known as dermoid cyst and benign cystic teratoma). The data presented here show that MCT are genomically bland, but the associated SCC have a high mutation burden characterised by a high frequency of mutations in *TP53, CDKN2A* and *PIK3CA*. *TP53* mutation is an early event, and likely bi-allelic *TP53* mutations are seen frequently. Although previous individual case reports describe high risk HPV in an ovarian SCC (13), all 25 SCC analysed here were negative for high/intermediate risk HPV genomes. There are few data on recurrent genomic abnormalities in malignant ovarian germ cell tumours (GCT) but SCC arising in MCT differ greatly from malignant testicular GCT, which have low mutational burden (0.9/MB) and are universally *TP53* wild-type (14).

Although we found mutations in genes frequently mutated in both lung and skin SCC (*TP53* and *CDKN2A*), these two tumour types also contain frequent mutations in *NOTCH1, 2* and 3 (15, 16). We only identified *NOTCH1* mutations in two of our 25 cases (8%), one of which had only a synonymous change. However, our sequencing panel did not include other *NOTCH* genes. The rates of both *PIK3CA* and *CDKN2A* abnormality in our samples are potentially higher than in lung and skin SCC, but, overall SCC arising in MCT has features in common with other non-HPV SCC. A larger sample size would be required to make more definite statements about relative mutational frequency. The overall mutational burden estimate for the SCC was high at 10.2 mutations per MB of sequenced DNA. Assessment of mutation rates derived from cancer gene panels may over-estimate the genome-wide rate as these panels include only genes that are known to be under selective pressure. Chalmers et al recently compared mutational burden estimates from whole exome sequencing with targeted panels, demonstrating close correlation between the two (10). They also suggested that sequencing of approximately 1.1Mb of coding genome could accurately estimate overall mutational burden, with significant variance only below 0.5 Mb. Our data were derived from approximately 0.8Mb, and thus may be an overestimate, but are largely in line for data from other SCC types, including NSCLC.

Bi-allelic *TP53* mutations have been identified in skin SCC (17). Exome sequencing of eight skin SCC identified *TP53* abnormalities in seven. Four of these showed LOH at chromosome 17p, but the 3 remaining cases lacked LOH and contained two or more distinct *TP53* mutations, presumed to be biallelic. We are not aware of other descriptions of bi-

allelic alterations in other SCC, nor of any previous potential correlation between bi-allelic *TP53* mutations and improved outcome.

There remains considerable debate as to the origin of MCT. Originally thought to represent parthenogenic activation of oocytes (ie embryological development in the absence of a male gamete) at the end of meiosis I (18), cytogenetic analyses later suggested the existence of five separate MCT groups, arising from distinct stages of meiosis (2). More recent short tandem repeat analyses generated conflicting data, supporting origins in oogonia (19) or primary oocytes that have escaped meiotic arrest (3).

Our study was not designed to identify cell of origin of MCT, but to analyse whether there were obvious SCC precursor mutations present in MCT. These we did not find - indeed, the MCT had few mutations or copy number alterations. However, we did identify four separate zygosity states in the 22 MCT samples from 21 cases. Groups A and B appear to have arisen primordial germ cells/oogonia and primary oocytes respectively. Group C contained some SNPs with altered allelic ratios. In the absence any detectable copy number alterations, one potential mechanism for generating such a state would be fusion of a cell that has undergone meiosis I but not completed meiosis II with a somatic cell, such that regions that had crossed over would end up in a 1:3 ratio while the rest of the genome would be 2:2. Alternatively this profile would be consistent with a mixture of a group B profile and contaminating normal cells. Group D, which lacked any SNPs in a 0.5:0.5 ratio but instead in each case showed a single allelic ratio consistent with a polyploid genome with unequal numbers of maternal and paternal chromosomes, may arise due to errors in meiosis I, when sets of homologous chromosomes should be separated. Given that only a single allelic ratio is present in these cells it seems unlikely that crossover has occurred.

Analysis of any rare cancer is logistically problematic, and we acknowledge that there are potential shortcomings in our study. Identifying rare cancer samples is challenging, and we acquired the original 31 samples only by interrogating the archives of four large UK gynaecological cancer centres, with limited clinical data available. However, the clinical outcomes we describe are in line with a large systematic review (6), which suggests that our patients are broadly representative of this patient population. SCC arising within MCT has the added disadvantage that it is almost invariably an unexpected pathological diagnosis in an apparently benign tumour - there are no proven imaging characteristics to differentiate a benign MCT from one containing SCC (20, 21). We therefore had to work with archival formalin-fixed paraffin-embedded material up to 17 years old. Finally, our sequencing panel,

although containing 151 genes, did not cover the whole exome, and so there may be critical mutations that we have missed.

Despite these issues, this first systematic genomic evaluation of ovarian SCC arising in MCT shows clearly that they are similar to other non-HPV SCC, especially NSCLC, but with distinct features, including bi-allelic *TP53* mutations. Further studies will be required to address the question of MCT cell of origin and to understand what triggers transformation of the MCT into SCC. However, our data suggest that patients with MCT-SCC could be included in trials of SCC-specific therapy and may, like lung SCC (22), derive benefit from immune checkpoint inhibition.

Acknowledgements

This Scottish Genomes Partnership is funded by the Chief Scientist Office of the Scottish Government Health Directorates (grant reference SGP/1) and The Medical Research Council Whole Genome Sequencing for Health and Wealth Initiative. Additional funding was provided by the Medical Research Council (the Glasgow Molecular Pathology Node, grant reference MR/N005813/1), Cancer Research UK (grant references A15973 [IMcN] and A17263 [AVB]), the Wellcome Trust (grant reference 103721/Z/14/Z [AVB]) and the Beatson Cancer Charity (grant reference 15-16-051 [IMcN]). We thank Craig Nourse for technical assistance, and Kymab for licencing GeneCN to us.

References

1. Ulbright TM. Germ cell tumors of the gonads: a selective review emphasizing problems in differential diagnosis, newly appreciated, and controversial issues. Mod Pathol. 2005;18 Suppl 2:S61-79.
2. Surti U, Hoffner L, Chakravarti A, Ferrell RE. Genetics and biology of human ovarian teratomas. I. Cytogenetic analysis and mechanism of origin. American journal of human genetics. 1990;47:635-43.
3. Kaku H, Usui H, Qu J, Shozu M. Mature cystic teratomas arise from meiotic oocytes, but not from pre-meiotic oogonia. Genes, chromosomes & cancer. 2016;55:355-64.
4. Zhu HL, Zou ZN, Lin PX, Li WX, Huang YE, Shi XX, et al. Malignant transformation rate and p53, and p16 expression in teratomatous skin of ovarian mature cystic teratoma. Asian Pacific journal of cancer prevention : APJCP. 2015;16:1165-8.
5. Westhoff C, Pike M, Vessey M. Benign ovarian teratomas: a population-based case-control study. Br J Cancer. 1988;58:93-8.
6. Hackethal A, Brueggmann D, Bohlmann MK, Franke FE, Tinneberg HR, Munstedt K. Squamous-cell carcinoma in mature cystic teratoma of the ovary: systematic review and analysis of published data. Lancet Oncol. 2008;9:1173-80.
7. Glasspool RM, Gonzalez Martin A, Millan D, Lorusso D, Avall-Lundqvist E, Hurteau JA, et al. Gynecologic Cancer InterGroup (GCIG) consensus review for squamous cell carcinoma of the ovary. Int J Gynecol Cancer. 2014;24:S26-9.
8. TCGA. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489:519-25.
9. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013;500:415-21.
10. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome medicine. 2017;9:34.
11. TCGA. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015;517:576-82.
12. Pickering CR, Zhou JH, Lee JJ, Drummond JA, Peng SA, Saade RE, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. Clin Cancer Res. 2014;20:6582-92.
13. Verguts J, Amant F, Moerman P, Vergote I. HPV induced ovarian squamous cell carcinoma: case report and review of the literature. Arch Gynecol Obstet. 2007;276:285-9.
14. Taylor-Weiner A, Zack T, O'Donnell E, Guerriero JL, Bernard B, Reddy A, et al. Genomic evolution and chemoresistance in germ-cell tumours. Nature. 2016;540:114-8.
15. Wang NJ, Sanborn Z, Arnett KL, Bayston LJ, Liao W, Proby CM, et al. Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. Proc Natl Acad Sci U S A. 2011;108:17761-6.
16. Li YY, Hanna GJ, Laga AC, Haddad RI, Lorch JH, Hammerman PS. Genomic analysis of metastatic cutaneous squamous cell carcinoma. Clin Cancer Res. 2015;21:1447-56.
17. Durinck S, Ho C, Wang NJ, Liao W, Jakkula LR, Collisson EA, et al. Temporal dissection of tumorigenesis in primary cancers. Cancer Discov. 2011;1:137-43.
18. Linder D, McCaw BK, Hecht F. Parthenogenic origin of benign ovarian teratomas. N Engl J Med. 1975;292:63-6.
19. Wang WC, Lai YC. Genetic analysis results of mature cystic teratomas of the ovary in Taiwan disagree with the previous origin theory of this tumor. Human pathology. 2016;52:128-35.

20. Kido A, Togashi K, Konishi I, Kataoka ML, Koyama T, Ueda H, et al. Dermoid cysts of the ovary with malignant transformation: MR appearance. AJR American journal of roentgenology. 1999;172:445-9.
21. Futagami M, Yokoyama Y, Mizukami H, Shigeto T, Mizunuma H. Can malignant transformation in mature cystic teratoma be preoperatively predicted? Eur J Gynaecol Oncol. 2012;33:662-5.
22. Brahmer J, Reckamp KL, Baas P, Crino L, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med. 2015;373:123-35.
23. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401-4.
24. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science signaling. 2013;6:pl1.

Tables

Table 1

Summary of patient characteristics and samples. Stage is based upon FIGO classification at the time of diagnosis. Samples: N, normal; M, Mature Cystic Teratoma; S, Squamous Cell Carcinoma. OS event: 0 = alive, 1 = dead. NA - data not available

Table 2

Multivariate analysis of overall survival. HR = Hazard Ratio; UL = Upper limit (95% confidence interval); LL = Lower limit (95% confidence interval)

Table 3

Zygosity summary of the MCT samples. Columns 5 and 6 indicates the percentage of all SNPs that were heterozygote in MCT and normal tissue respectively for groups where the MCT showed the same set of frequency states as the normal ie 0. 0.5 and 1. For groups C and D, the metric of percentage heterozygous SNPs is not relevant given the altered SNP frequencies in the MCT compared to the normal, and thus the analysis was not performed, as indicated by dashes (-). NA; sample not available

Figure legends

Figure 1. Sample profile

Flow of patients and samples in this study.

Figure 2. Summary of genomic alterations

A. Summary of frequently altered genes across the SCC and MCT samples (23, 24). Cases with bi-allelic *TP53* mutation are marked *.
B. Schematic representation of p53 showing protein domains. (Green, transactivation domain [TAD]; red, DNA binding domain [DBD]; blue, tetramerisation domain [TMD]) with lollipops showing positions and counts of identified mutations. Mutation type is indicated by circle fill: green, nonsynonymous; black, loss of function (including nonsense, splicing and frameshift); red, inframe indel/synonymous.

Figure 3. Survival analyses

A. Median overall survival for the whole cohort was 20.1 months
B. Overall survival was significantly better for FIGO stage I compared to FIGO stage II-IV (HR 0.301; p=0.021)
C. Age has no impact upon survival - overall survival by age at diagnosis (<median vs >median).
D. Overall survival was significantly better for *TP53* mutant cases compared to *TP53* wild-type cases (HR 0.178; p=0.002)

Table 1

Study No	Age at Diagnosis (years)	Time since diagnosis (months)	OS Event	Stage	Samples
SC001	78.0	44.5	0	IA	N M S
SC002	60.0	13.4	1	IBB	M S
SC005	35.0	NA	NA	IC2	N M S
SC006	43.5	55.5	0	IIIC	N M S
SC007	53.6	55.8	0	IA	N M S
SC008	51.0	64.6	0	IA	N M S
SC009	61.0	123.8	0	IA	N S
SC010	65.0	18.9	1	IIIC	N M S
SC011	41.0	11.6	1	IIIB	N M S
SC012	65.0	2.6	1	IIB	N M S
SC013	58.0	20.1	1	IC2	N M S
SC014	29.0	24.2	1	IB	N M Sx2
SC015	50.7	12.7	1	IIIC	N M S
SC019	37.1	41.6	0	IA	N M S
SC020	56.2	2.6	1	IC	N M S
SC021	45.7	95.5	0	IA	N M S
SC022	50.4	172.6	0	IA	N M Sx2
SC023	31.3	83.3	0	IA	N S
SC024	86.2	3.8	1	III	N S
SC025	50.8	6.6	1	IIIC	N M Sx2
SC027	68.0	42.1	0	II	N S
SC028	46.0	NA	NA	III	N M S
SC029	78.4	5.5	1	IC2	N M S
SC030	25.6	6.8	0	IA	N M Sx3
SC031	73.1	7.2	1	IIA	N Mx2 Sx2

Table 2

Variable	Category	N (events)	HR	LL	UL	P-value
TP53	WT (ref)	4 (3)	1	-	-	0.011
	Mono	12 (8)	0.015	0.001	0.253	
	BI	7 (1)	0.077	0.01	0.602	
Stage	I (ref)	13 (4)	1	-	-	0.047
	II-IV	10 (8)	3.802	1.017	14.285	

Table 3

Group	Study No.	Non-homozygous allele frequency (MCT)	Non-homozygous allele frequency (Normal)	Heterozygous SNPs MCT (%)	Heterozygous SNPs Normal (%)	Copy Number profile
A	SC005	0.5	0.5	39.9	39.9	
A	SC010	0.5	NA	39.4	NA	
A	SC012	0.5	0.5	38.6	38.6	
A	SC015	0.5	0.5	34.0	38.0	
A	SC020	0.5	0.5	43.5	43.3	
A	SC021	0.5	0.5	33.6	33.6	
A	SC025	0.5	0.5	30.4	29.6	Amplification of MYC and EGFR
A	SC028	0.5	0.5	33.4	33.4	
A	SC031_MCT1	0.5	0.5	31.1	31.1	
A	SC031_MCT2	0.5	0.5	31.1	31.1	
B	SC008	0.5	0.5	23.6	41.9	
B	SC014	0.5	0.5	18.9	37.5	
B	SC019	0.5	0.5	17.6	36.1	
C	SC002	0.3; 0.5; 0.7	NA	-	NA	
C	SC011	0.3; 0.5; 0.7	0.5	-	43.7	
C	SC013	0.2; 0.5; 0.8	NA	-	NA	Some low-level changes
D	SC001	0.4; 0.6	0.4; 0.6	-	-	
D	SC006	0.4; 0.6	0.5	-	37.1	One low-level gain
D	SC007	0.25; 0.75	0.5	-	41.1	
D	SC022	0.3; 0.7	0.5	-	38.4	
D	SC029	0.2; 0.8	0.5	-	38.4	
D	SC030	0.4; 0.6	0.5	-	36.1	

Figure 1

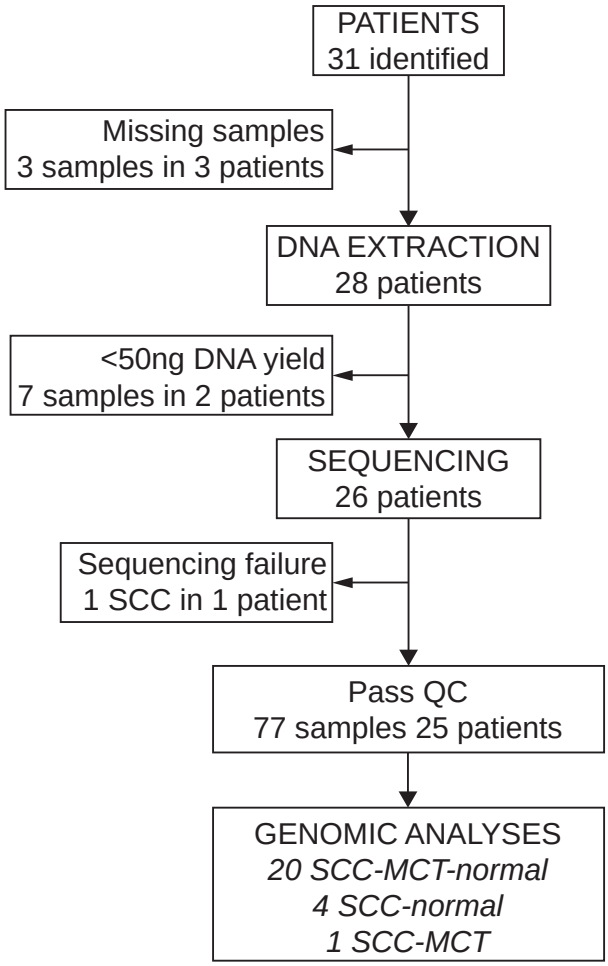
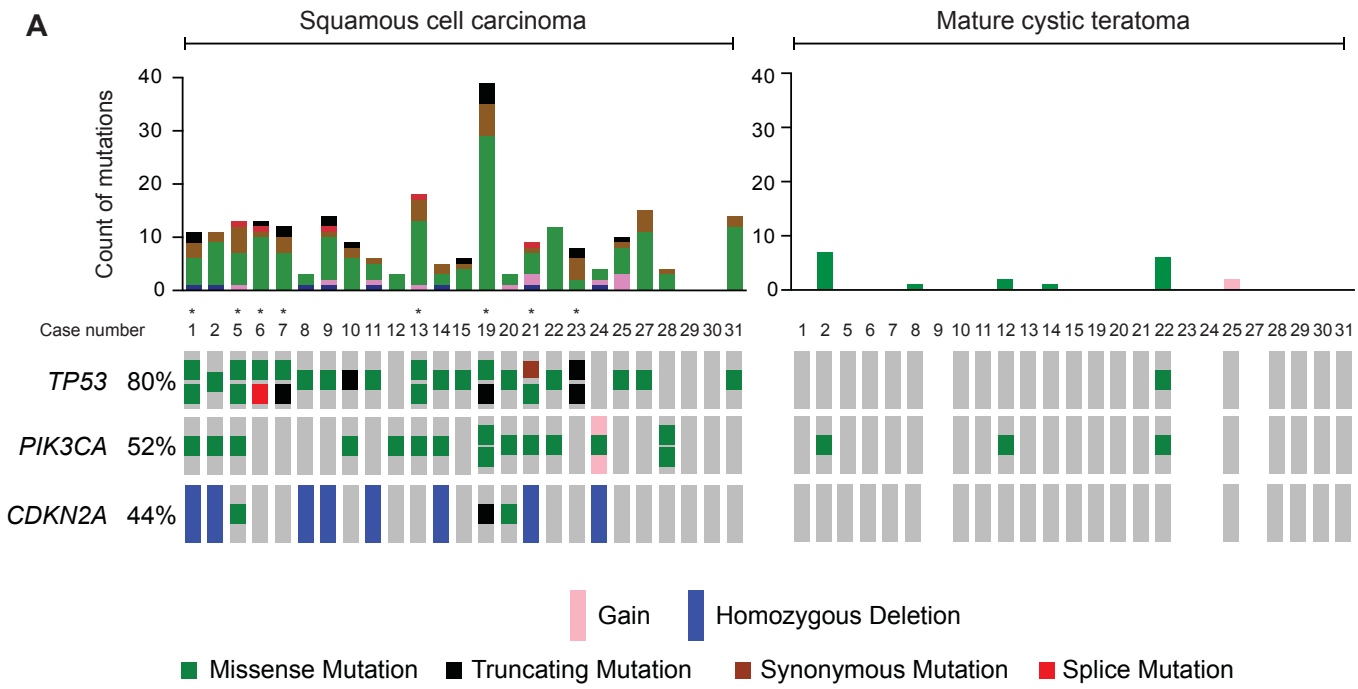

PATIENTS
31 identified
Missing samples
3 samples in 3 patients
DNA EXTRACTION
28 patients
<50ng DNA yield
7 samples in 2 patients
SEQUENCING
26 patients
Sequencing failure
1 SCC in 1 patient
Pass QC
77 samples 25 patients
GENOMIC ANALYSES
20 SCC-MCT-normal
4 SCC-normal
1 SCC-MCT

Figure 2


A
Squamous cell carcinoma
Mature cystic teratoma
Count of mutations
40
30
20
10
0
Case number
1 2 5 6 7 8 9 10 11 12 13 14 15 19 20 21 22 23 24 25 27 28 29 30 31
1 2 5 6 7 8 9 10 11 12 13 14 15 19 20 21 22 23 24 25 27 28 29 30 31
TP53 80%
PIK3CA 52%
CDKN2A 44%
Gain
Homozygous Deletion
Missense Mutation
Truncating Mutation
Synonymous Mutation
Splice Mutation

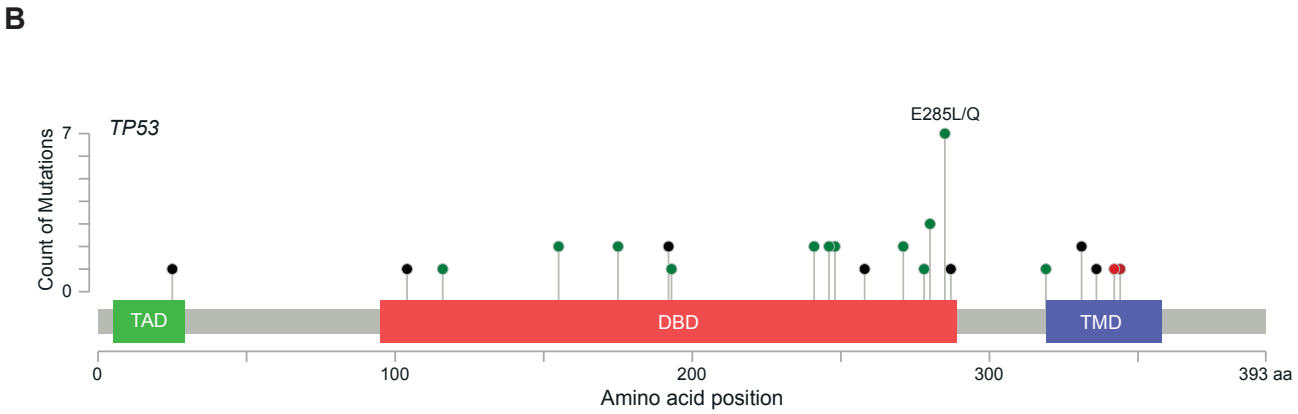
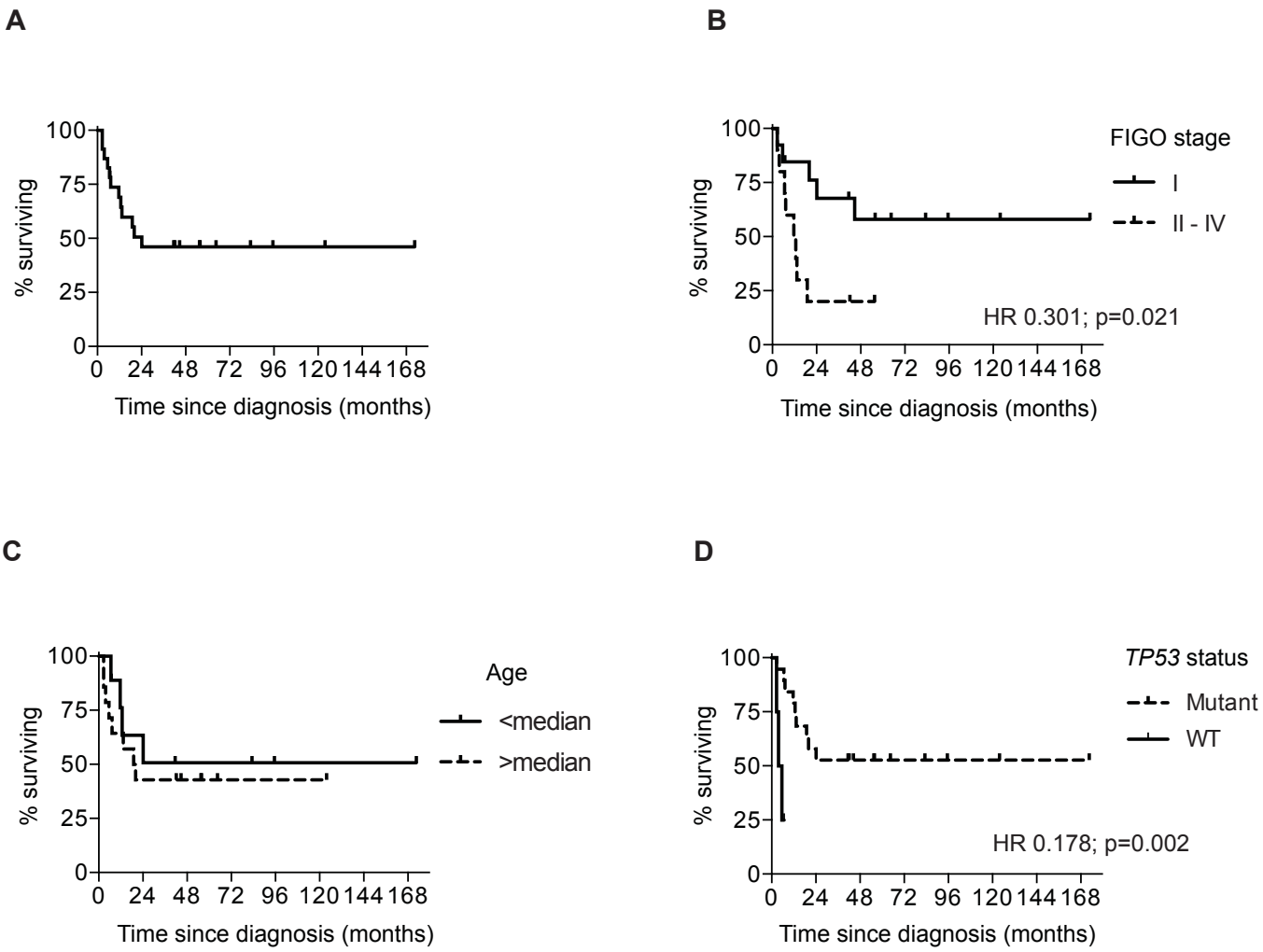

B
TP53
Count of Mutations
7
0
E285L/Q
TAD
DBD
TMD
0
100
200
300
393 aa
Amino acid position

Figure 3


A
B
C
D
% surviving
Time since diagnosis (months)
0
25
50
75
100
0 24 48 72 96 120 144 168
FIGO stage
I
II - IV
HR 0.301; p=0.021
Age
<median
>median
TP53 status
Mutant
WT
HR 0.178; p=0.002