Monogenic and polygenic determinants of sarcoma risk: an international genetic study

Mandy L Ballinger, David L Goode*, Isabelle Ray-Coquard, Paul A James, Gillian Mitchell, Eveline Niedermayr, Ajay Puri, Joshua D Schiffman, Gillian S Dite, Arcadi Cipponi, Robert G Maki, Andrew S Brohl, Ola Myklebost, Eva W Stratford, Susanne Lorenz, Sung-Min Ahn, Jin-Hee Ahn, Jeong Eun Kim, Sue Shanley, Victoria Beshay, Robert Lor Randall, Ian Judson, Beatrice Seddon, Ian G Campbell, Mary-Anne Young, Rajiv Sarin, Jean-Yves Blay, Seán I O'Donoghue, David M Thomas, for the International Sarcoma Kindred Study†*

Summary

Background Sarcomas are rare, phenotypically heterogeneous cancers that disproportionately affect the young. Outside rare syndromes, the nature, extent, and clinical significance of their genetic origins are not known. We aimed to investigate the genetic basis for bone and soft-tissue sarcoma seen in routine clinical practice.

Methods In this genetic study, we included 1162 patients with sarcoma from four cohorts (the International Sarcoma Kindred Study [ISKS], 966 probands; Project GENESIS, 48 probands; Asan Bio-Resource Center, 138 probands; and kConFab, ten probands), who were older than 15 years at the time of consent and had a histologically confirmed diagnosis of sarcoma, recruited from specialist sarcoma clinics without regard to family history. Detailed clinical, pathological, and pedigree information was collected, and cancer diagnoses in probands and relatives were independently verified. Targeted exon sequencing using blood (n=1114) or saliva (n=48) samples was done on 72 genes (selected due to associations with increased cancer risk) and rare variants were stratified into classes approximating the International Agency for Research on Cancer (IARC) clinical classification for genetic variation. We did a case-control rare variant burden analysis using 6545 Caucasian controls included from three cohorts (ISKS, 235 controls; LifePool, 2010 controls; and National Heart, Lung, and Blood Institute Exome Sequencing Project [ESP], 4300 controls).

Findings The median age at cancer diagnosis in 1162 sarcoma probands was 46 years (IQR 29–58), 170 (15%) of 1162 probands had multiple primary cancers, and 155 (17%) of 911 families with informative pedigrees fitted recognisable cancer syndromes. Using a case-control rare variant burden analysis, 638 (55%) of 1162 sarcoma probands bore an excess of pathogenic germline variants (combined odds ratio [OR] 1·43, 95% CI 1·24–1·64, p<0·0001), with 227 known or expected pathogenic variants occurring in 217 individuals. All classes of pathogenic variants (known, expected, or predicted) were associated with earlier age of cancer onset. In addition to *TP53*, *ATM*, *ATR*, and *BRCA2*, an unexpected excess of functionally pathogenic variants was seen in *ERCC2*. Probands were more likely than controls to have multiple pathogenic variants compared with the combined control cohort group and the LifePool control cohort (OR 2·22, 95% CI 1·57–3·14, $p=1·2\times10^{-6}$) and the cumulative burden of multiple variants correlated with earlier age at cancer diagnosis (Mantel-Cox log-rank test for trend, p=0·0032). 66 of 1162 probands carried notifiable variants following expert clinical review (those recognised to be clinically significant to health and about which patients should be advised), whereas 293 (25%) probands carried variants with potential therapeutic significance.

Interpretation About half of patients with sarcoma have putatively pathogenic monogenic and polygenic variation in known and novel cancer genes, with implications for risk management and treatment.

Funding Rainbows for Kate Foundation, Johanna Sewell Research Foundation, Australian National Health and Medical Research Council, Cancer Australia, Sarcoma UK, National Cancer Institute, Liddy Shriver Sarcoma Initiative.

Introduction

The genetic architecture of cancer risk is usually ascribed to a combination of rare variation in families with dominant inheritance patterns, and common variants with small effect sizes in the population at large. To date, less than half of the estimated heritability of even well-studied cancers is explained by rare and common variants.[1,2] The so-called missing heritability is likely due to currently unrecognised rare variants, and to unmeasured genetic interactions between both common and rare variants.[3,4] Massively parallel germline sequencing is revealing a striking new landscape of pathogenic genetic variation, which is filling in some of these gaps in cancer heritability.[5–8] Initially in high-risk[6,8,9] and increasingly in sporadic cancer populations,[5,7] genomic screens are yielding potentially explanatory rare pathogenic variants in 5–40% of cases.[10,11] These studies, typically qualitative reports of case series,[5–8] are increasingly identifying unexpected genotype–phenotype correlations and non-mendelian inheritance patterns.[10,11] The future application of statistical case-control designs developed in genome-wide association studies (GWAS)

Lancet Oncology 2016

Published Online
August 4, 2016
http://dx.doi.org/10.1016/S1470-2045(16)30147-4

See Online/Comment
http://dx.doi.org/10.1016/S1470-2045(16)30292-3

*Joint first authors

†Members provided in appendix

The Kinghorn Cancer Centre and Cancer Division (M L Ballinger PhD, A Cipponi PhD, Prof D M Thomas FRACP) **and Genomics and Epigenetics Division** (S I O'Donoghue PhD), **Garvan Institute of Medical Research, Darlinghurst, NSW, Australia; Sir Peter MacCallum Department of Oncology** (M L Ballinger, D L Goode PhD, P A James MBChB, G Mitchell PhD, E Niedermayr BSc, S Shanley MD, Prof I G Campbell PhD, M-A Young MSc), **Department of Pathology** (P A James, Prof I G Campbell), **and Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health** (G S Dite PhD), **University of Melbourne, Melbourne, VIC, Australia; Research Division** (M L Ballinger, D L Goode, E Niedermayr, Prof I G Campbell), **Familial Cancer Centre** (P A James, G Mitchell, S Shanley, M-A Young), **and Department of Pathology** (V Beshay PhD), **Peter MacCallum Cancer Centre, Melbourne, VIC, Australia; Department of Adult Medical Oncology, Centre Leon Berard, University Claude Bernard, Lyon, France** (Prof I Ray-Coquard MD, Prof J-Y Blay MD); **Department of Orthopedic Oncology, Tata Memorial Hospital, Mumbai, India** (Prof A Puri MS); **Division of Pediatric Hematology/Oncology, Department of Pediatrics** (Prof J D Schiffman MD) **and**

Sarcoma Services, Department of Orthopaedics (Prof R Lor Randall MD), **Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA; Tisch Cancer Institute, Mount Sinai Medical Center, New York, NY, USA** (Prof R G Maki MD); **Sarcoma Department, Moffitt Cancer Center, Tampa, FL, USA** (A S Brohl MD); **Department of Tumor Biology, Oslo University Hospital, Norwegian Radium Hospital, Oslo, Norway** (Prof O Myklebost DPhil, E W Stratford PhD, S Lorenz PhD); **Norwegian Cancer Genomics Consortium, Oslo, Norway** (Prof O Myklebost, E W Stratford, S Lorenz); **Gachon Institute of Genome Medicine and Science, Gachon University Gil Medical Center, Incheon, South Korea** (S-M Ahn MD); **Department of Oncology, Asan Medical Centre, Seoul, South Korea** (J-H Ahn MD, J Eun Kim MD); **Sarcoma Unit, The Royal Marsden NHS Foundation Trust, London, UK** (Prof I Judson MD); **Sarcoma Unit, University College Hospital, London, UK** (B Seddon PhD); **Cancer Genetics Unit, ACTREC, Tata Memorial Centre, Mumbai, India** (Prof R Sarin FRCR); **and Commonwealth Scientific and Industrial Research Organisation (CSIRO), Sydney, Australia** (S I O'Donoghue)

Correspondence to: Prof David M Thomas, Garvan Institute of Medical Research, The Kinghorn Cancer Centre and Cancer Division, Darlinghurst, NSW 2010, Australia
d.thomas@garvan.org.au

See Online for appendix

For **Project GENESIS** see http://healthcare.utah.edu/huntsmancancerinstitute/research/research-studies/project-genesis/

For the **EVS website** see www.evs.gs.washington.edu/EVS/

Research in context

Evidence before this study

We searched PubMed for articles published in English up to February, 2016, with the terms "genetic", "genomic", "sarcoma", "cancer", "multigene", "germline", "familial", or "heritable". Registry-based epidemiological studies have noted that sarcomas affect a young population, who have increased risk of multiple primary cancers, including sarcomas. Studies of rare families have identified genes such as *TP53*, *NF1*, and *SDHB* in syndromic risk for sarcomas in general, or for specific subtypes. Many recent studies have applied multigene panels or whole-exome analyses to cancer, often in the context of matched germline and tumour genomic panel testing. Most studies to date focus on common cancers, high-risk familial cancer subgroups, and to a lesser extent the general cancer populations. These emerging studies, largely descriptive case series lacking controls or replication sets, nonetheless consistently report a high frequency of explanatory genetic variants (up to 40%). There have been no studies in the sarcoma population, and few studies have systematically integrated genotype with familial cancer patterns.

Added value of this study

This study builds upon a global cohort of 1162 patients with sarcoma, containing detailed phenotypic information on familial cancer burden and sarcoma pathology. One in six patients belonged to families that fitted criteria for hereditary cancer syndromes, most previously unrecognised. Another one in ten patients belonged to families considered to carry excess cancer burden, not fitting known cancer syndromes. More than half of patients with sarcoma in our study carried apparently pathogenic germline genetic variants. One in five patients harboured known or expected pathogenic variants, and one in 15 carried germline variants that seem to be clinically actionable. Specific enrichment is shown for several genes, both expected (*TP53*) and novel (*BRCA2*, *ATM*, *ATR*, and *ERCC2*). We show for the first time, to our knowledge, a measurable contribution of polygenic effects to sarcoma risk, both by rare variant burden analysis of cases and controls and age-of-onset effects. Finally, almost one in four patients carried germline genetic variants that might influence choice of therapy.

Implications of all the available evidence

An underappreciated burden of clinically important genetic variation exists among patients with sarcoma and their relatives. Familial patterns are not reliable guides as to underlying genotype, challenging the usefulness of clinical criteria in genetic testing in an era of cheap genomic panels. Integration of hereditary cancer expertise within multidisciplinary sarcoma management seems justifiable, as exists now for breast and bowel cancer.

to genomic screens will yield more quantitative insights into total cancer heritability, and particularly genetic interactions.[3,4]

Several lines of evidence suggest a strong genetic basis to sarcomas. Relative to all cancers, sarcomas disproportionately affect the young, accounting for 20% of childhood cancers and 10% of adolescent and young adult cancers.[12] Early age of diagnosis is associated with a genetic basis for many heritable diseases including heritable cancers.[10,11] Sarcoma survivors are at increased risk of second cancers,[13] while sarcomas are themselves over-represented among survivors of melanoma, breast cancer, thyroid cancer, Hodgkin's lymphoma, and leukaemias.[14] Finally, several rare genetic syndromes are associated with sarcomas, such as Li-Fraumeni syndrome which is associated with germline mutations in *TP53*.[15] Collectively, these syndromes account for a small fraction of incident cases, leaving the bulk of apparently sporadic sarcoma cases unexplained. Here, we apply quantitative methods to study, for the first time to our knowledge, rare pathogenic germline variation in the sarcoma population.

Methods

Study design and participants

The goal of this study was to characterise the individual and familial genetic determinants of sarcoma risk. We aimed to identify rare genetic variants associated with sarcoma risk through targeted exon sequencing of patients with sarcoma, assess the disease burden of these rare variants through case-control analysis, and identify the genes which contributed most to these effects.

We included 1162 sarcoma probands from four cohorts (the International Sarcoma Kindred Study [ISKS], 966 probands; Project GENESIS, 48 probands; Asan Bio-Resource Center, 138 probands; and kConFab, ten probands). We also included 6545 controls from the Caucasian subsets of three cohorts (ISKS, 235 controls; LifePool, 2010 controls; and National Heart, Lung, and Blood Institute [NHLBI] Exome Sequencing Project [ESP], 4300 controls; appendix p 11).

The LifePool control cohort consists of 2010 Australian women with no history of cancer at time of first mammogram and no subsequent diagnoses. The LifePool cohort consists of women unaffected by cancer and self-identified as Caucasian. Family history of sarcoma in these families is unknown. The second control set were cancer-free individuals (non-blood relatives of probands) recruited to the ISKS. The NHLBI ESP Exome Variant Server (EVS) is a database of exome sequencing results collected from multiple studies. Variant and coverage data from the European American samples (within the EVS population) were downloaded from the EVS website on Nov 14, 2014. The database is publicly available and it is unknown if it contains individuals with cancer or a family history of cancer including sarcoma. In general, EVS samples were selected on the basis of non-cancer phenotypes, such as

traits related to cardiovascular disease (LDL cholesterol and blood pressure), early-onset myocardial infarction and early-onset stroke, and non-malignant lung disease. They were not ascertained on the basis of a personal or family history of cancer. The EVS represents a useful independent control set of whole-exome sequence data on a Caucasian population unselected for cancer phenotypes.

We used 317 patients with sarcoma for validation of variants (Norwegian Sarcoma Consortium [NoSarc], 93 individuals; and The Cancer Genome Atlas [TCGA], 224 individuals; appendix p 11).

Between July, 2009, and March, 2015, patients with a sarcoma (aged >15 years at the time of consent) were eligible for participation in the ISKS cohort irrespective of family history of cancer with the exception of nine sarcoma probands who were recruited into the kConFab cohort on the basis of a high incidence of breast cancer in family members. Other than family history, the eligibility criteria for all case and validation cohorts was the same; the three control cohorts were all ascertained for different reasons, but for the purpose of the use as controls in our study they were all selected because they represented ethnically matched (Caucasian) cancer-free controls. At the time of recruitment, probands were either undergoing or had completed treatment. All sarcoma diagnoses were confirmed by expert sarcoma pathologists at the recruiting centres. Dates of recruitment and further clinical details were not available for the Asan, Project GENESIS, kConFab, NoSarc, or TCGA samples, or for the control cohorts. All participants provided written informed consent according to local requirements, and parental consent was also obtained for probands aged 16–17 years in all cohorts.

Procedures

Either blood or saliva samples were obtained at the time of study enrolment and in all case cohorts could be taken before or after chemotherapy. Probands provided family history information, while medical history and treatment records were obtained for each proband when possible. If dates of birth and death were unavailable, estimates were made using a 25-year generation time and expected lifespan of 70 years, as described previously.[16] All reported cancer diagnoses in probands and relatives were independently verified by reference to medical records, cancer registries, or death certificates. When verification was unavailable, the age at diagnosis was estimated as described previously.[16] Study questionnaires containing demographic, medical, epidemiological, and psychosocial information were completed, including personal history of cancer or past exposure to known risk factors for sarcoma such as ionising radiation. Pedigree information was not collected for the Asan Bio-Resource Center participants. Ethnicity was inferred when information was unavailable.

Proband pedigrees were assessed according to recognised clinical criteria (appendix p 4) and cancer risk to relatives was estimated (appendix p 4). Pedigrees not conforming to recognised criteria but displaying any of the following features were considered clinically suspicious: more than half of first-degree relatives with cancer at any age; three or more cancers per proband at any age; one or more sarcomas in first-degree relatives of a sarcoma proband; more than one sarcoma per proband at any age; more than two cancers per proband younger than 50 years; and average age of cancer onset in first-degree relatives younger than 50 years when at least two cancers were reported or at least two connective tissue tumours were reported per proband.

Targeted exon sequencing (HaloPlex Enrichment System, Agilent, Santa Clara, CA, USA) was undertaken on 72 genes, selected because of associations with increased cancer risk (appendix pp 4, 12–17, 27–29). Sequence alignment, variant calling and mapping, and functional assessment was performed and Exome Aggregation Consortium data were used to filter out common segregating variants (appendix pp 5–6). Rare variants were stratified using an automated algorithm into three classes approximating the International Agency for Research on Cancer (IARC) 3–5 clinical classification for genetic variation (appendix p 6).[18] Briefly, class 5 variants represented disease-causing mutations in the Human Genetic Mutation Database, class 4 variants were predicted to result in a frameshift, premature stop or affect an essential splice site or initiation codon, and class 3 variants comprised missense variants predicted in silico to be deleterious. Class 1 and 2 variants represent known benign and likely benign variants and were excluded from the analysis. We performed a principal component analysis to control for genetic ancestry in the LifePool cohort. The LifePool samples were sequenced with a separate batch of HaloPlex gene panel reagents that included 58 genes in common with the ISKS capture. The ISKS controls were assayed with the same set of HaloPlex reagents as the sarcoma cases. The EVS, TCGA, and NoSarc datasets were generated using whole-exome DNA sequencing, using either Roche/Nimblegen or Agilent capture reagents.

Functionality of *ERCC2* variants was assessed using a range of bioinformatics approaches and, for selected variants, an in-vitro cisplatin sensitivity assay (appendix pp 8).

Statistical analysis

To identify total and intragenic enrichment in rare variants, a case-control rare variant burden analysis was performed using the Caucasian datasets as controls (from LifePool, ISKS, and EVS), and data from TCGA and NoSarc were used as independent replication sets (appendix pp 3, 11, 26). Odds ratios (ORs) and p values reported for rare variant burden analysis were obtained

from one-sided Fisher's exact tests to compare total burden of rare deleterious variants relative to rare synonymous variants in cases and controls. p values were combined across comparisons using Kost's method and ORs combined using Mantel-Haenszel's method (appendix pp 6, 7). Risk to relatives was calculated as described (appendix p 4). The standardised incidence ratio in risk-to-relatives analyses was estimated by comparing the number of affected first-degree relatives of the case probands with the number expected to be affected using the sex-specific and age-specific population incidence, as described previously (appendix p 4).[17] Australian population incidences for all cancers combined (except non-melanoma skin cancer) were obtained from the Australian Institute of Health and Welfare. These data include annual cancer incidence for 1982 to 2011 specific for sex and age (in 5-year groups). Living unaffected relatives were censored from the date of proband questionnaire completion (at the earliest) or 100 years, while deceased unaffected relatives were censored at age at death. To deal with incomplete or incorrect self-reported ancestry, we conducted principal component analysis using all segregating single nucleotide variants (ie, called in two or more samples) covered by our gene capture panel, enabling removal of samples that were clear outliers in our population with respect to genetic ancestry. We did a multivariate regression using a Cox proportional hazards model that included genotype, sarcoma subtype, sex, and previous chemotherapy or radiotherapy. We did an analysis of the relationship between carriage of class 3–5 variants and multiple primary cancers. Two independent control sets were used (appendix p 11), one assayed on the same set of HaloPlex capture reagents as the sarcoma cases (n=235), and the other was assayed on an independent set of HaloPlex capture reagents (LifePool, n=2010). Rare variant burden analyses were limited to Caucasian probands and controls (to eliminate ethnicity as a confounding factor), after removal of outliers following a principal components analysis for ethnic stratification (n=848).

We deemed p values less than 0·05 to be significant. The Gehan-Breslow-Wilcoxon method was used to compare tumour-free survival (for all cohorts, from birth to date of diagnosis in years) between groups, and the Mantel-Haenszel method to generate hazard ratios (HRs) and 95% CIs. The log-rank test for trend was used for estimating the cumulative effect of class 3 variant burden on age at cancer diagnosis (appendix pp 6, 7). No individuals were censored in time-to-tumour onset analyses. Analyses were done with GraphPadPrism 6 and R version 3.02.

For the **data appendix** see http://www.garvan.org.au/research/cancer/novel-monogenic-and-polygenic-determinants-of-sarcoma-risk.xlsx/view

Role of the funding source

The funders had no role in the study design, collection, analysis, or interpretation of the data or writing of the manuscript. MLB and DLG had full access to all the raw data. The corresponding author had full access to all of the data and the final responsibility for the decision to submit for publication.

Results

Of 1162 patients with sarcoma included in our analysis, most self-identified as Caucasian (853 [73%]), whereas 151 (13%) reported east Asian ancestry and 94 (8%) reported southeast Asian ancestry (table). The median age at first cancer diagnosis was 46 years (IQR 29–58, range 1 month to 93 years; table, appendix p 25), and at first sarcoma diagnosis was 47 years (IQR 29–60, range 3–93 years). The youngest age at which the first cancer was diagnosed in the cohort was 4 months (retinoblastoma) and the oldest was 93 years (undifferentiated pleomorphic sarcoma). Among probands with multiple cancers, sarcomas arose an average of 12 years (SD 10) after a first cancer in 99 (58%) of 170 probands, most commonly after breast cancer (30 [3%]). 32 sarcoma subtypes were observed (table).

First-degree relatives of sarcoma probands had an increased burden of cancer (654 diagnoses of cancer in 3978 first-degree relatives of sarcoma probands; standardised incidence ratio [SIR] compared with the Australian population 1·09, 95% CI 1·01–1·18), specifically sarcomas (2·65, 1·6–4·4), brain tumours (1·62, 1·02–2·57), breast cancer (1·52, 1·25–1·84), and melanoma (1·39, 1·09–1·78). On blinded expert review of 911 informative families (those with informative pedigrees—ie, sufficient information for assessment of family history of cancer), 155 (17%) fulfilled criteria for recognised cancer syndromes including Li-Fraumeni syndrome, hereditary colorectal cancer, familial melanoma, and hereditary breast or ovarian cancer (table). An additional 87 families carried excess or unusual patterns of cancer in first-degree relatives (186 cancers in 499 first-degree relatives; SIR 1·83, 95% CI 1·55–2·15), similar with that seen in recognised syndromes. Only 30 (5%) of 579 Australian families reported referral for genetic counselling. Data about genetic counselling was only available for Australian ISKS families.

Of 1162 sequencing samples from the sarcoma cohorts at enrolment, 1114 (96%) were from blood and 48 (4%) were from saliva. 759 (65%) were taken before and 344 (30%) were taken after chemotherapy, and 58 (5%) were unknown; source and timing of samples was not available for the validation cohort. All probands were genotyped successfully; targeted exon sequencing identified a total of 956 class 3–5 rare variants in 638 (55%) of 1162 probands, with 127 class 5 (known disease causing) variants in 122 individuals, 100 class 4 variants in 95 individuals, and 729 class 3 variants in 529 individuals (some probands carried more than one variant; data appendix). The median age at first cancer diagnosis was significantly younger in probands carrying class 4 or 5 variants compared with those

	Probands (n=1162)
Sex	
Male	586 (50%)
Female	576 (50%)
Age at diagnosis, years	
First cancer	46 (29–58)
Sarcoma	47 (29–60)
Number with multiple primary cancers	170 (15%)
Two primary cancers	128 (11%)
Three primary cancers	32 (3%)
Four or more primary cancers	10 (1%)
Pedigree classification	
No syndrome	669 (58%)
Classic or Chompret Li-Fraumeni syndrome	116 (10%)
Hereditary colorectal cancer	14 (1%)
Familial melanoma	9 (1%)
Hereditary breast or ovarian cancer	6 (<1%)
Clinically suspicious*	87 (7%)
Other†	10 (1%)
Uninformative pedigrees	251 (22%)
Sarcoma subtypes	
Bone sarcoma	348 (30%)
Ewing sarcoma or primitive neuroectodermal tumour	134 (12%)
Osteosarcoma	124 (11%)
Chondrosarcoma	88 (9%)
Other	2 (<1%)
Soft-tissue sarcoma‡	830 (70%)
Undifferentiated pleomorphic sarcoma	205 (18%)
Leiomyosarcoma	132 (11%)
Well differentiated or dedifferentiated liposarcoma	94 (9%)
Synovial sarcoma	68 (6%)
Myxoid liposarcoma	51 (4%)
Gastrointestinal stromal tumour	44 (4%)
Angiosarcoma	28 (2%)
Malignant peripheral nerve sheath tumour	28 (2%)
Liposarcoma not otherwise specified	24 (2%)
Dermatofibrosarcoma protuberans	22 (2%)
Aggressive fibromatosis	22 (2%)
Endometrial stromal sarcoma	20 (2%)
Epithelioid sarcoma	13 (1%)
Solitary fibrous tumour	12 (1%)
Pleomorphic liposarcoma	12 (1%)
Other	55 (5%)
Ethnicity	
Caucasian	853 (73%)
East Asian	151 (13%)
Southeast Asian	94 (8%)
Other	22 (2%)
Unknown	42 (4%)

Data are n (%) or median (IQR). *More than half of first-degree relatives with cancer at any age, three or more cancers per proband at any age, one or more sarcoma in first-degree relatives of a sarcoma proband, more than one sarcoma per proband at any age, more than two cancers per proband younger than 50 years, average age of cancer onset in first-degree relatives younger than 50 years where two or more cancers are reported, or two or more connective-tissue tumours per proband. †Familial papillary thyroid cancer, hereditary paraganglioma syndrome, multiple schwannomatosis, neurofibromatosis type 1 and 2, Gorlin syndrome, multiple endocrine neoplasia type 1, retinoblastoma (two cases), and McCune-Albright syndrome. ‡Values for sarcoma subtypes sum to more than 1162 because 13 individuals had more than one sarcoma.

Table: **Clinical and pedigree details of individuals participating in this study**

carrying no variants (43 years [IQR 26–56, 95% CI 37–46] *vs* 50 years [32–60, 46–51]; p=0·0010; figure 1A), and in those carrying class 3 variants compared with those carrying no variants (45 years [IQR 28–56, 95% CI 42–46] *vs* 50 years [31–60, 46–51]; p=0·0014; figure 1B). The median age of first cancer diagnosis was 44 years (IQR 27–56, 95% CI 41–46) for those carrying any variant (p=0·00020 *vs* probands without variants), and 38 years (IQR 20–52, 95% CI 31–44) for 81 individuals carrying class 4 or 5 variants in 24 dominant cancer genes based on those considered notifiable by the American College of Medical Genetics (ACMG; p<0·0001 when compared with probands without variants; notifiable variants are those recognised to be clinically significant to health and about which patients should be advised).[19] The results of the multivariate regression analysis are shown in the appendix (pp 24, 32). The effect of genotype on age of cancer diagnosis remained significant in this analysis, even after correcting for sarcoma subtype, previous therapy, or sex.

The burden of putatively pathogenic rare variation was assayed by rare variant burden analysis, comparing probands with cancer-free controls (appendix pp 18–22). We noted consistent enrichment of pathogenic variants of all types in cases compared with controls. The OR of having any class 3–5 variant in probands in the ISKS cohort was significantly greater compared with each control group independently or when combined (figure 1C). The effect was more marked for class 4 or 5 variants than for class 3 variants, and persisted after excluding cases with a previous history of chemotherapy or radiotherapy, with an OR for combined risk in all class 3–5 carriers of 1·36 (95% CI 1·20–1·54), only marginally lower than the OR of 1·43 (95% CI 1·24–1·66) for the entire cohort (appendix pp 7, 31). These findings were replicated in an independent test set composed of 317 sarcoma probands from TCGA and NoSarc (figure 1C, appendix p 23). The enrichment in class 3–5 variants in patients with sarcoma compared with controls remained significant after exclusion of genes already associated with increased risk for sarcomas (ie, *TP53*, *RB1*, *NF1*,

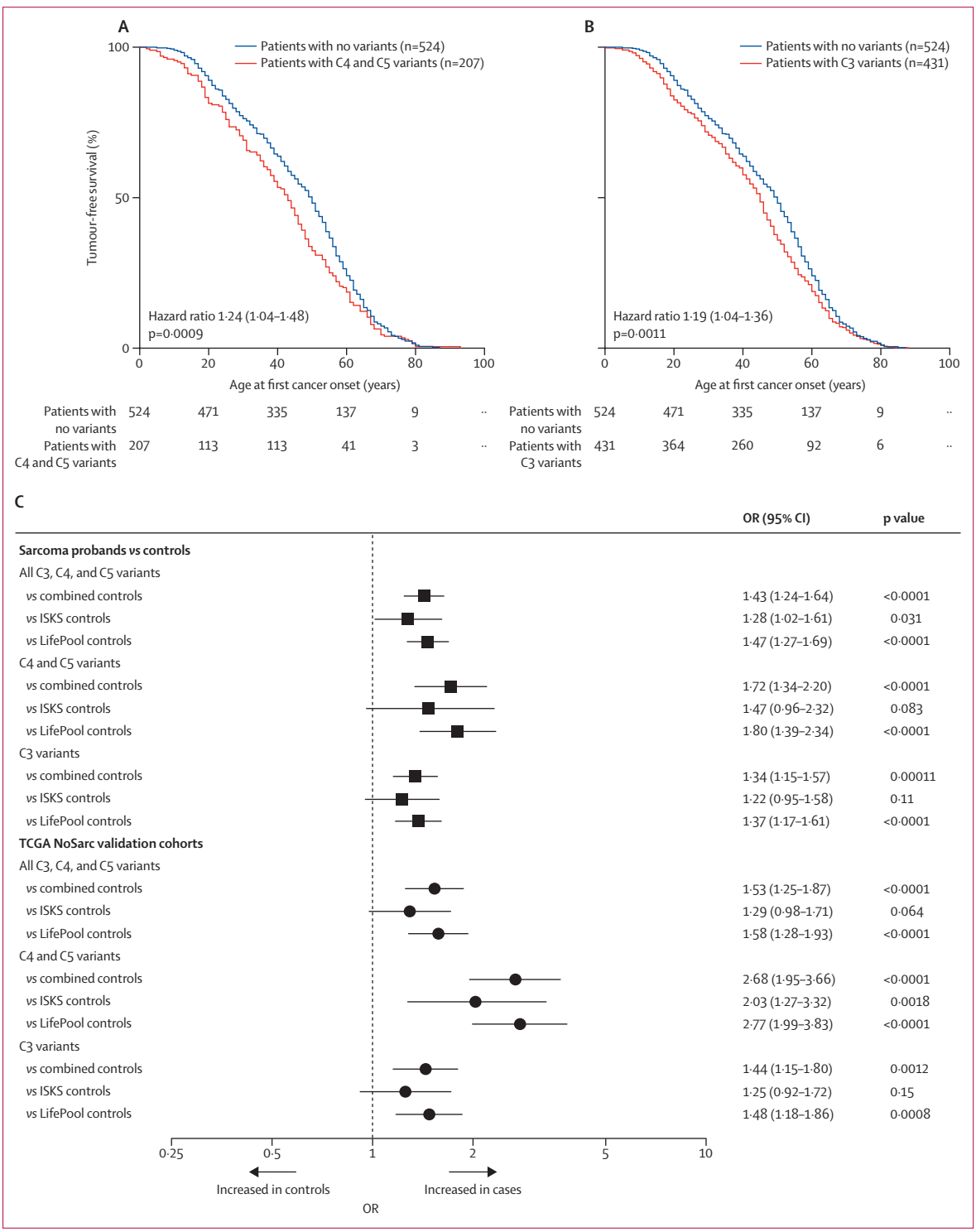

Figure 1: **Monogenic rare variant burden analysis**
(A) Kaplan-Meier tumour-free survival in sarcoma probands who have class 4 or 5 variants versus patients with no variants. (B) Kaplan-Meier tumour-free survival for patients with class 3 variants only versus patients with no variants. (C) Rare variant burden analysis in the Caucasian subset of sarcoma cases (top panel) or TCGA and NoSarc cases (bottom panel) with two control populations. Data are odds ratios and 95% CIs for each comparison. C3=class 3. C4=class 4. C5=class 5. ISKS=International Sarcoma Kindred Study. TCGA=The Cancer Genome Atlas. OR=odds ratio.

SDHA, *SDHB*, *SDHC*, and *SDHD*), suggesting that previously unrecognised genes contribute to risk for sarcoma (OR 1·36 [95% CI 1·20–1·54] for all class 3–5 variants *vs* 1·43 [1·24–1·64] for class 3 variants; appendix pp 7, 30).

We next considered which individual genes out of the 72 genes on the panel contributed to the overall enrichment in deleterious variants seen in sarcoma cases compared with controls (figure 2, appendix pp 19–22). Significant enrichment for pathogenic variation was observed in *TP53* and also genes implicated in DNA damage sensing (*ATM*, *ATR*) and homologous recombination (*BRCA2*) in probands versus controls. 23 patients included in the rare variant burden analysis had variants in *ERCC2*, a helicase involved in nucleotide excision repair associated with the autosomal recessive cancer-prone syndrome xeroderma pigmentosum type D.[20] Mutations in *ERCC2* affect DNA binding, DNA damage sensing, helices activity, or basal transcription, and increase sensitivity to cisplatin.[21] 12 patients with *ERCC2* variants had class 4 or 5 variants and 11 patients had class 3 variants. Two additional Caucasian probands had the pathogenic Leu461Val variant of *ERCC2*, linked to xeroderma pigmentosum type D and trichothiodystrophy. However, the *ERCC2* Leu461Val was excluded from the rare variant burden analysis because it is present at 1·3% minor allele frequency among the south Asian population, despite a minor allele frequency of 0·11% in non-Finnish Caucasian populations (the threshold for other variants is ≤1% in the control population from Exome Aggregation Consortium). One of 12 patients with class 4 or 5 variants had an osteosarcoma, three patients had primitive neuroectodermal tumours, one patient had a chondrosarcoma, and the remainder (seven patients) had sarcomas of various kinds. No pattern was observed suggesting an association of *ERCC2* variants with specific sarcoma subtypes, although the numbers are too small to make any meaningful associations. A full discussion of all *ERCC2* class 3–5 variants found to be associated with sarcomas in our study is provided in the appendix, including experimental confirmation of loss-of-function for nine of 12 variants in an assay for cisplatin sensitivity (appendix pp 9–10, 33).

In a pooled analysis of all sarcoma probands, 240 probands carried multiple variants (2–6 per individual; median 2 [IQR 2–3]), suggesting a polygenic contribution to sarcoma risk. To distinguish polygenic from monogenic effects due specifically to multiple class 4 or 5 variants, we restricted subsequent analyses to class 3 variants. Progressively earlier age of first cancer diagnosis correlated with increasing cumulative burden of class 3 variants (figure 3A). Using rare variant burden analysis, we noted that it was significantly more likely for sarcoma probands to have multiple pathogenic variants compared with the combined control cohort group and the LifePool control cohort; although this difference was not significant in the ISKS sarcoma cohort compared with the ISKS controls due to the small size of the ISKS

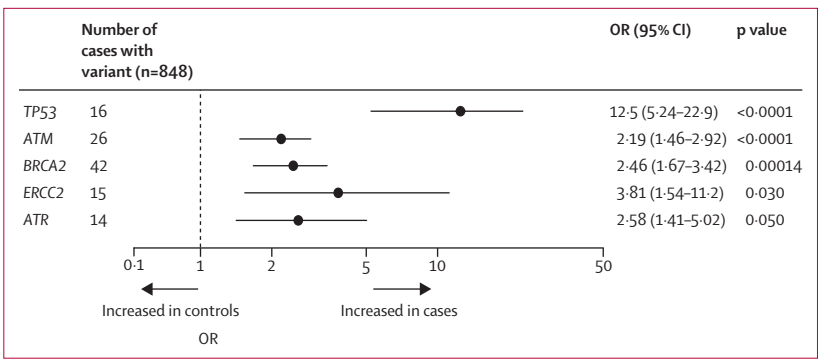

Figure 2: **Rare variant burden analysis for *TP53*, *ATM*, *BRCA2*, *ERCC2*, and *ATR***
Comparisons were made between the Caucasian subset of the sarcoma cases with the combined controls from the ISKS, LifePool, and Exome Variant Server datasets. Data are odds ratios and 95% CIs for each comparison. ISKS=International Sarcoma Kindred Study. OR=odds ratio.

control group, the direction and magnitude of the change was consistent with that seen in comparisons with LifePool controls (figure 3B). These results were replicated in TCGA and NoSarc samples (figure 3B). Since polygenic inheritance patterns should more closely resemble recessive than dominant effects, we compared the incidence of cancer in first-degree relatives of patients with multiple class 3 variants (62 cancers in 563 individuals) with that in patients with class 4 or 5 variants in autosomal dominant cancer genes (68 cancers in 361 individuals). The standardised incidence of cancer was lower (SIR compared with the Australian population 1·21, 95% CI 0·65–2·25, p=0·64) in first-degree relatives of patients carrying multiple class 3 variants compared with that in first-degree relatives of patients with class 4 or 5 variants in autosomal dominant genes (1·63, 1·25–2·11, p=0·00071). The strongest polygenic effects on age at first cancer diagnosis were noted among 34 patients who had either biallelic class 3–5 variants in one gene or a class 4 or 5 variant in two or more genes. Patients who carried biallelic class 3–5 variants in one gene, or class 4 or 5 variants in two or more genes, had a median age of first cancer diagnosis of 25 years (IQR 17–47, 95% CI 18–45), compared with patients who did not have any variants who had a median age of first cancer diagnosis of 50 years (IQR 32–60, 95% CI 46–51); p<0·0001). Patients who had variants in *TP53* had a similarly early age of first cancer diagnosis (32 years [IQR 22–44, 95% CI 24–38], figure 3C).

ACMG guidelines recommend the return of information to the patient on known or expected pathogenic mutations in 20 autosomal dominant tumour suppressor genes.[19] 61 (5%) of 1162 individuals carried a class 4 or 5 mutation in *APC* (six individuals), the mismatch repair genes *MLH1*, *MSH2*, *MSH6*, and *PMS2* (11 individuals), *BRCA1* or *BRCA2* (28 individuals), *TP53* (12 individuals), *TSC2* (three individuals), *SDHB* (two individuals), *RB1* (one individual), and *PTEN* (one individual); three probands carried two variants each. A further 19 (2%) of the 1162 probands carried class

For the **eviQ guidelines** see http://www.eviq.org.au

4 or 5 variants in *PTCH1* (three individuals), *PALB2* (five individuals), *CDH1* (six individuals), and *NF1* (five individuals; data appendix). *BRCA1/2*, *TP53*, *TSC2*, *SDHB*, *RB1*, and *PTEN* are all ACMG genes, whereas *CDH1*, *NF1*, *PTCH1*, and *PALB2* are at present not on the ACMG list of notifiable genes but are widely considered by the clinical community to be significant to health (see, for example, the eviQ guidelines for the management of cancer genetics guidelines). After expert review (SS, PAJ, VB, GM, and JDS), 66 of the 81 variants were considered to be notifiable (ie, it was deemed clinically appropriate to inform patients carrying that of their status because of implications for future health; data appendix).

Of 911 evaluable probands, those with a recognisable familial syndrome (n=155) were more likely to carry a class 4 or 5 variant in ACMG genes than were those without (n=756; OR 4·63 [95% CI 2·63–8·12], p<0·0001). However, only 13 (20%) of 66 probands with informative pedigrees carrying class 4 or 5 variants in autosomal dominant ACMG genes had a cognate syndrome, and only 25 (16%) of 155 with a recognisable syndrome were associated with ACMG class 4 or 5 variants, although not always concordant with the observed syndrome. Of ten informative pedigrees of probands with class 4 or 5 variants in *TP53*, seven met Chompret or classic criteria for Li-Fraumeni syndrome (appendix p 3).[22,23] 11 of 1162 probands had class 3–5 variants in *NF1* and five of these patients developed malignant peripheral nerve sheath tumours or gastrointestinal stromal tumours. Four patients with variants of *SDHA* or *SDHB* developed gastrointestinal stromal tumours (two with class 4 or 5 *SDHA* mutations and two with class 4 or 5 *SDHB* mutations). Among 33 probands with class 4 or 5 variants in *BRCA1* (nine individuals), *BRCA2* (19 individuals), or *PALB2* (five individuals), none met Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm thresholds for *BRCA1/2* testing. However, two patients with *BRCA2* variants and two patients with *PALB2* variants developed both breast cancer and sarcoma, and seven patients (three *PALB2*, four *BRCA2*) had first-degree relatives with breast or prostate cancer, typically under 60 years of age. Among 14 informative families of probands with class 4 or 5 variants in hereditary colorectal cancer genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*, or *APC*), four met formal criteria for familial adenomatous polyposis or Lynch syndrome.

Recessive and compound heterozygous effects were also observed in our study. Three individuals carried biallelic variants in *RECQL4* (two carried the same

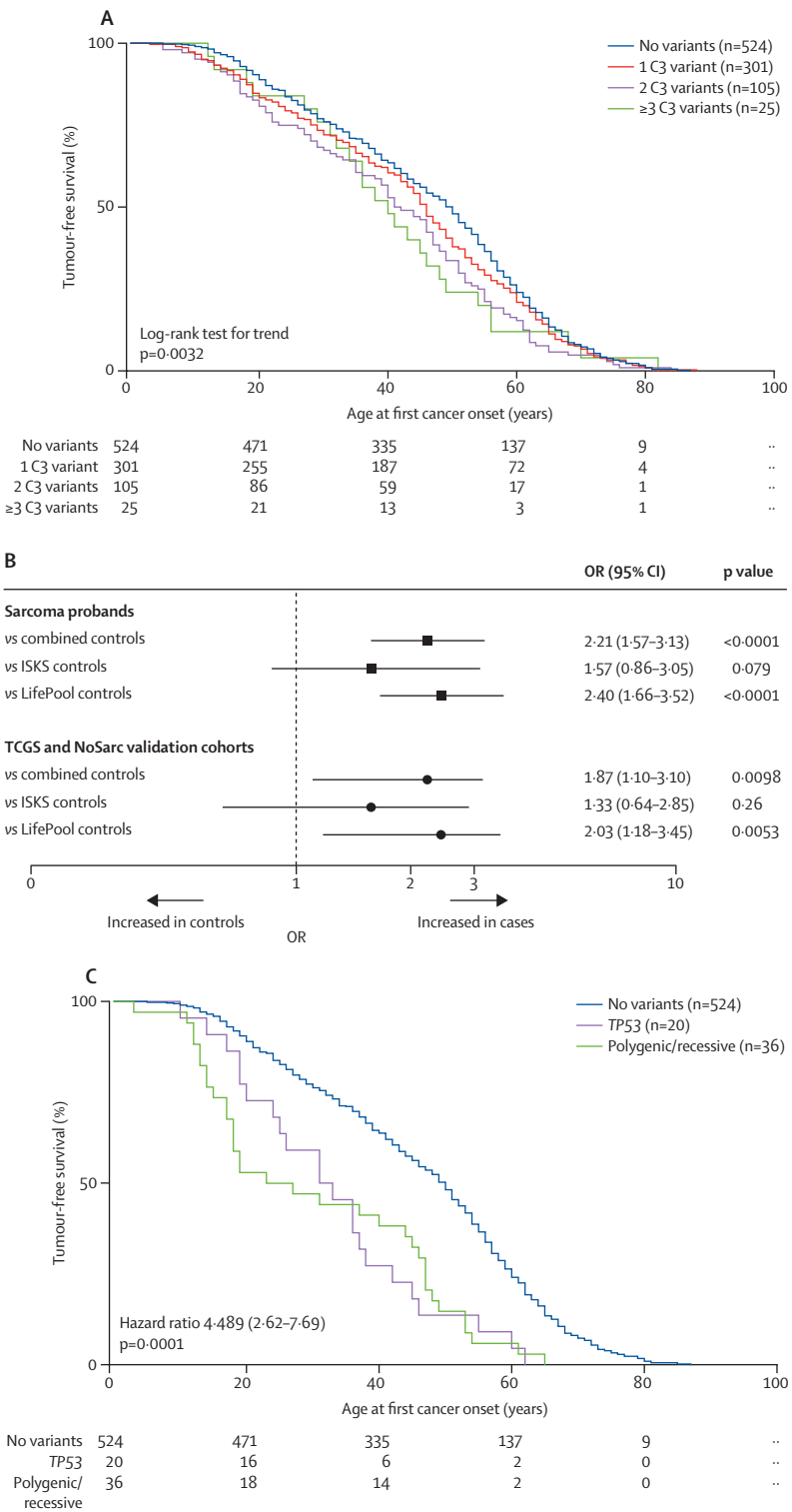

Figure 3: **Polygenic rare variant burden analysis in sarcoma probands**
(A) Tumour-free survival in patients who had one, two, or three or more class 3 variants compared with sarcoma probands who did not carry any variants. (B) Rare variant burden analysis of class 3 variant burden in the Caucasian sarcoma population and TCGA and NoSarc with two control populations. (C) Tumour-free survival in sarcoma probands carrying biallelic class 3–5 variants in one gene, or heterozygous class 4 or 5 variants in two or more genes, versus class 3–5 variants in *TP53*, or no variants. C4=class 4. C5=class 5. ISKS=International Sarcoma Kindred Study. TCGA=The Cancer Genome Atlas. OR=odds ratio.

two *RECQL4* variants and another proband carried a completely different combination of two *RECQL4* variants), which are associated with Rothmund-Thomson syndrome.[24] These individuals developed osteosarcoma at 15 years, dermatofibrosarcoma protuberans at 17 years, and leiomyosarcoma at 53 years, respectively. 12 probands carried homozygous (five individuals) or digenic (seven individuals) loss-of-function variants in genes associated with Fanconi anaemia. The homozygote individuals included a patient with a *FANCM* variant who developed a primative neuroectodermal tumour at age 11 years, a patient with a variant of *FANCG* who developed an endometrial stromal sarcoma at age 27 years, a patient with a variant of *FANCI* who developed a pleomorphic liposarcoma at age 48 years, another patient with a variant of *FANCI* who developed a synovial sarcoma at age 53 years, and a patient with a variant of *FANCL* who developed a myxoid liposarcoma tumour at age 40 years. We observed an association between germline variants in *FANC* genes and sarcomas associated with somatic translocations (eg, Ewing sarcoma, synovial sarcoma, and myxoid liposarcoma) compared with sarcomas with complex (eg, undifferentiated pleomorphic, leiomyosarcoma, and osteosarcoma) or simple (eg, well-differentiated liposarcoma) genotypes (134 probands with germline *FANC* variants, of whom 38 had translocation-associated sarcomas; OR 1·94, 95% CI 1·2–2·94, p=0·00060). We also observed potential modifier effects (where the potential for malignancy may be increased), exemplified by *EXT1* and *EXT2* variants. Mutations in these genes are typically associated with multiple benign osteochondromas, but of 13 individuals with mutations in *EXT1* or *EXT2* who carried a second mutation in one or more of *PMS2* (two), *WRN* (two), *RECQL4* (three), *MLH3*, *FANCA*, *EXT2* itself, *FANCL*, *BUB1B*, *XPC*, *TP53*, *BRCA1*, *BRCA2*, *ERCC4*, and *WT1*, eight developed osteosarcoma or chondrosarcoma (bone malignancies) at a median age of 31 years (range 8–61 years).

We also investigated whether patients had mutations in genes known to predict response to available targeted therapies and found that 293 (25%) of 1162 patients had mutations in genes which can predict response to available therapies, or to therapies that are currently in development. In particular, 124 patients had *BRCA1* and *BRCA2* mutations that predict responses to poly ADP-ribose polymerase inhibitors,[25] *PTCH1* mutations that predict response to hedgehog pathway inhibitors,[26] and *TSC1* or *TSC2* mutations that predict responses to mTOR inhibitors.[27] 142 patients had mutations in members of the homologous recombination pathway which might predict response to established therapies; 32 patients had mutations in genes for which agents are in clinical development (eg, *IDH1*, *IDH2*, *APC*; n=32) and 46 individuals had mutations in mismatch repair genes, which could increase sensitivity to immune checkpoint inhibitors.[28]

Discussion

In this genetic analysis of individuals with sarcoma, and consistent with recent studies,[10,11] we observed a large, clinically significant, and under-recognised burden of genetic risk. The excess risk lies in both classic monogenic and previously unrecognised polygenic rare variation. In many cases, there was poor concordance between cancer phenotype or familial pattern and the underlying genetic variation. A substantial fraction of the cohort carried variants that could affect risk counselling, management, or drug therapy. Notably, few Australian probands had been referred for genetic counselling, also noted in a recent paediatric cancer study in which only 12 (16%) of 75 potential beneficiaries had been referred for genetic counselling.[11] The burden of likely pathogenic (class 4 or 5) variants in 24 autosomal dominant cancer genes based on the ACMG gene list (noted in 81 [7%] of the 1162 probands in our study) accords with the proportion reported in sarcoma cases tested as part of an institutional personalised medicine programme.[10] Some of the excess burden lies in known pan-sarcoma genes, such as *TP53*, but the contribution of *BRCA2*, *ATM*, *ATR*, and in particular *ERCC2* was less predictable. Both *ATM* and *ATR* encode important sensors of DNA damage upstream of *TP53* itself, and a case report identified *BRCA2* mutations in families meeting Chompret criteria for Li-Fraumeni syndrome.[29] Mutations in *TP53*, *ATM*, and *ATR* are associated with sensitivity to ionising radiation, the strongest environmental risk factor for sarcoma. *ERCC2* is a helicase functioning in base excision repair associated with increased cancer risk in the autosomal recessive disorder xeroderma pigmentosum type D.[30] Several lines of evidence from this study suggest *ERCC2* might be a new sarcoma susceptibility gene. In addition to case-control rare variant burden analysis and supporting in-vitro data, the proportion of probably pathogenic (class 4 or 5) to total variant burden is higher for *ERCC2* than for any gene other than *TP53*. Many of these variants affect domains crucial to the function of *ERCC2*, and mutations in this gene have been clinically linked by others[31] with trichothiodystrophy, cranio-oculofacial syndrome, or xeroderma pigmentosum type D. We and others have shown that mutations in *ERCC2* enhance the sensitivity of cells to cisplatin.[21] Like radiation, cisplatin is commonly used in the treatment of osteosarcomas.

The biological contribution of polygenic variation to cancer risk is novel and significant. Although cancer genetics has traditionally focused on single-gene disorders, recent genomic studies have suggested that roughly 1% of individuals in the normal population carry more than one presumed pathogenic variant,[10] although biological evidence supporting polygenic effects has been lacking. Measured by age of cancer onset and numbers of individuals affected, class 4 or 5 polygenic or recessive effects seemed to be at least similar to those of *TP53*, the strongest monogenic driver

For the **BOADICEA algorithm** see http://ccge.medschl.cam.ac.uk/boadicea

of sarcoma risk. Polygenic effects might account for atypical genotype–phenotype associations increasingly seen in genomic studies.[3,10,11] In a recent study, only 41% of participants carrying pathogenic variants in cancer genes had concordant diagnoses, similar to our observations.[10] Previously invisible so-called modifier effects[32] are suggested by high frequency of early-onset bone sarcomas in *EXT1/2* mutation carriers who also bear mutations in cancer genes such as *TP53* or *PMS2*, reminiscent of malignant chondroid transformation in transgenic mice combining loss-of-function alleles in *EXT1* with *TRP53* or *CDKN2A*.[33] Polygenic effects could also account for young-onset translocation-associated sarcoma subtypes not characterised by dominant mendelian inheritance patterns, such as synovial sarcomas, Ewing sarcomas, and myxoid liposarcomas.

This study, despite being to our knowledge the largest to date, is not powered to identify specific genetic or pathway interactions underlying the polygenic effects that are measurable in aggregate. We also focus on genetic associations common to sarcomas as a group, and although we observe qualitative associations with some sarcoma subtypes, our analysis is not powered to examine differences between different sarcoma subtypes. The focus on rare variation omits the contribution of common alleles relevant to more common cancer types,[34] and targeted gene panels cannot identify novel genes or non-coding or structural variation, and therefore probably underestimate the role that genetics play in disease causation. In this study we also did not consider the contribution of de-novo or mosaic variants to the total burden of pathogenic variation. Our algorithmic variant classification, designed primarily for rare variant burden analysis, does not correlate well with the clinical classification recommended by IARC.[18] The large group of class 3 variants in our study corresponds to variants classified by IARC as variants of uncertain or indeterminate significance. Although class 4 corresponds to likely pathogenic variants, class 5 includes some variants that do not meet current clinical criteria for pathogenicity. We note, however, that the clinical significance of any individual variant will increasingly depend on genetic context, and specifically on the presence of complementary variants. Despite measurably affecting age of cancer onset and enrichment at the population level, clinical review is vital in defining what actions should be taken clinically, should individuals be found to have genetic variants associated with sarcoma.

This study represents an important first step in mapping the heritability of sarcoma in human beings. There are recognised risk management strategies for individuals with several hereditary cancer syndromes, most commonly breast and bowel cancer. Sarcoma families found to be carrying high-risk genetic variants might benefit from surveillance and prevention strategies. Our findings suggest a genetic basis for susceptibility to the carcinogenic effects of ionising radiation, responsiveness to cisplatin-based therapy, and a rational basis for novel therapies either inside or outside clinical trials.[25,26] Ironically, many filtering pipelines for somatic tumour panels subtract potentially actionable germline variation from variation detected in tumour material (so that somatic—ie, tumour specific—changes are more apparent). The frequency of potentially actionable monogenic and polygenic germline variants in patients affected by sarcoma warrants attention as personalised medicine evolves, with particular relevance to other young-onset cancers.

For the **TCGA Research Network** see http://cancergenome.nih.gov

For the **Norwegian Sarcoma Consortium** see www.NoSarC.org

For the **National Comprehensive Cancer Network Guidelines for Detection, Prevention and Risk-Reduction** see http://www.nccn.org

Contributors

DMT is principal investigator of the International Sarcoma Kindred Study (ISKS) and conceived this study. IR-C, PAJ, GM, M-AY, AP, JDS, RGM, S-MA, J-HA, JEK, RLR, IJ, BS, RS, and J-YB are ISKS co-investigators and collected data. DMT, MLB, DLG, EN, GSD, AC, and SIO'D performed literatures searches, collected, analysed, interpreted and generated data, constructed figures and tables, and wrote the first draft. ASB, OM, EWS, SL, and IGC contributed data. PAJ, GM, JDS, SS, VB, and M-AY clinically classified pedigrees and variants. All authors contributed to and approved the final version of the manuscript.

Declaration of interests

GM is on an advisory board for AstraZeneca around genetic predisposition to ovarian cancer and the role of poly ADP-ribose polymerase inhibitors in the treatment of BRCA associated cancer. The other authors declare no competing interests.

Acknowledgments

We thank the patients and their families for generous participation in research studies. We acknowledge the support of ISKS collaborators and staff. For a portion of our analysis, we used data generated by TCGA Research Network. Some biospecimens and data used in this study were provided by Asan Bio-Resource Center, Korea Biobank Network (2015-1[91]). We acknowledge the Victorian, South Australian, and Queensland cancer registries for verification of reported cancers. This study was supported by grants from the Rainbows for Kate Foundation, the Johanna Sewell Research Foundation, Australian National Health and Medical Research Council (APP1004017), Cancer Australia (APP1067094), Sarcoma UK, NCI (1R01CA161780–01) and the Liddy Shriver Sarcoma Initiative. Authors have received the following support: DMT, NHMRC Principal Research Fellowship (APP1104364); DLG, NHMRC Early Career Fellowship (APP1052904); IR-C and J-YB, LYric (DGOS-INCA 4664), NetSARC (INCA), RREPS (INCA), LabEx DEvweCAN (ANR-10-LABX-0061), Eurosarc (FP7–278742), Ligue de L'Ain contre le Cancer and la Fondation ARC, patient advocacy group InfoSarcome & DAM'S; ASB, Pediatric Cancer Foundation (fastercure.org); RGM, Quad W Foundation; JDS, Primary Children's Hospital (PCH) Pediatric Cancer Research Program, funded by the PCH Foundation and the Intermountain Healthcare Foundation; S-MA, Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (HI14C0198); IGC, NH&MRC Principal Research Fellowship (APP1041975); BS, National Institute for Health Research University College London Hospitals Biomedical Research Centre. The Norwegian Sarcoma Consortium (NoSarC) was supported by grants from the Norwegian Research Council and the Norwegian Cancer Society.

References

1 Mucci LA, Hjelmborg JB, Harris JR, et al. Familial risk and heritability of cancer among twins in Nordic countries. *JAMA* 2016; **315:** 68–76.

2 Lu Y, Ek WE, Whiteman D, et al. Most common 'sporadic' cancers have a significant germline genetic component. *Hum Mol Genet* 2014; **23:** 6112–18.

3 Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 2012; **109:** 1193–98.

4 Zuk O, Schaffner SF, Samocha K, et al. Searching for missing heritability: designing rare variant association studies. *Proc Natl Acad Sci USA* 2014; **111:** e455–64.

5 Couch FJ, Hart SN, Sharma P, et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. *J Clin Oncol* 2015; **33:** 304–11.

6 Kurian AW, Hare EE, Mills MA, et al. Clinical evaluation of a multiple-gene sequencing panel for hereditary cancer risk assessment. *J Clin Oncol* 2014; **32:** 2001–09.

7 Maxwell KN, Wubbenhorst B, D'Andrea K, et al. Prevalence of mutations in a panel of breast cancer susceptibility genes in BRCA1/2-negative patients with early-onset breast cancer. *Genet Med* 2015; **17:** 630–38.

8 Tung N, Battelli C, Allen B, et al. Frequency of mutations in individuals with breast cancer referred for BRCA1 and BRCA2 testing using next-generation sequencing with a 25-gene panel. *Cancer* 2015; **121:** 25–33.

9 LaDuca H, Stuenkel AJ, Dolinsky JS, et al. Utilization of multigene panels in hereditary cancer predisposition testing: analysis of more than 2,000 patients. *Genet Med* 2014; **16:** 830–37.

10 Schrader KA, Cheng DT, Joseph V, et al. Germline variants in targeted tumor sequencing using matched normal DNA. *JAMA Oncol* 2016; **2:** 104–11.

11 Zhang J, Walsh MF, Wu G, et al. Germline mutations in predisposition genes in pediatric cancer. *N Engl J Med* 2015; **373:** 2336–46.

12 Bleyer A, Barr R, Hayes-Lattin B, et al. The distinctive biology of cancer in adolescents and young adults. *Nat Rev Cancer* 2008; **8:** 288–98.

13 Fernebro J, Bladstrom A, Rydholm A, et al. Increased risk of malignancies in a population-based study of 818 soft-tissue sarcoma patients. *Br J Cancer* 2006; **95:** 986–90.

14 Levi F, Randimbison L, Maspoli-Conconi M, Blanc-Moya R, La Vecchia C. Incidence of second sarcomas: a cancer registry-based study. *Cancer Causes Control* 2014; **25:** 473–77.

15 Li FP, Fraumeni JF Jr, Mulvihill JJ, et al. A cancer family syndrome in twenty-four kindreds. *Cancer Res* 1988; **48:** 5358–62.

16 Downing ME, Dite GS, Ballinger ML, for the International Sarcoma Kindred Study Consortium. An increased incidence of Hodgkin's lymphoma in patients with adult-onset sarcoma. *Clin Sarcoma Res* 2012; **2:** 1.

17 Dite GS, Whittemore AS, Knight JA, et al. Increased cancer risks for relatives of very early-onset breast cancer cases with and without *BRCA1* and *BRCA2* mutations. *Br J Cancer* 2010; **103:** 1103–08.

18 Plon SE, Eccles DM, Easton D, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 2008; **29:** 1282–91.

19 Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 2013; **15:** 565–74.

20 Arrand JE, Bone NM, Johnson RT. Molecular cloning and characterization of a mammalian excision repair gene that partially restores UV resistance to xeroderma pigmentosum complementation group D cells. *Proc Natl Acad Sci USA* 1989; **86:** 6997–7001.

21 Van Allen EM, Mouw KW, Kim P, et al. Somatic *ERCC2* mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. *Cancer Discov* 2014; **4:** 1140–53.

22 Li FP, Fraumeni JF Jr, Mulvihill JJ, et al. A cancer family syndrome *Cancer Res* 1988; **48:** 5358–62.

23 Bougeard G, Renaux-Petel M, Flaman JM, et al. Revisiting Li-Fraumeni syndrome from *TP53* mutation carriers. *J Clin Oncol* 2015; **33:** 2345–52.

24 Veith S, Mangerich A. RecQ helicases and *PARP1* team up in maintaining genome integrity. *Ageing Res Rev* 2015; **23:** 12–28.

25 Kaufman B, Shapira-Frommer R, Schmutzler RK, et al. Olaparib monotherapy in patients with advanced cancer and a germline *BRCA1/2* mutation. *J Clin Oncol* 2015; **33:** 244–50.

26 Von Hoff DD, LoRusso PM, Rudin CM, et al. Inhibition of the hedgehog pathway in advanced basal-cell carcinoma. *N Engl J Med* 2009; **361:** 1164–72.

27 Bissler JJ, McCormack FX, Young LR, et al. Sirolimus for angiomyolipoma in tuberous sclerosis complex or lymphangioleiomyomatosis. *N Engl J Med* 2008; **358:** 140–51.

28 Le DT, Uram JN, Wang H, et al. PD-1 Blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015; **372:** 2509–20.

29 Evans DG, Wu CL, Birch JM. BRCA2: a cause of Li-Fraumeni-like syndrome. *J Med Genet* 2008; **45:** 62–63.

30 Broughton BC, Berneburg M, Fawcett H, et al. Two individuals with features of both xeroderma pigmentosum and trichothiodystrophy highlight the complexity of the clinical outcomes of mutations in the *XPD* gene. *Hum Mol Genet* 2001; **10:** 2539–47.

31 Fan L, Fuss JO, Cheng QJ, et al. XPD helicase structures and activities: insights into the cancer and aging phenotypes from *XPD* mutations. *Cell* 2008; **133:** 789–800.

32 Grant RC, Selander I, Connor AA, et al. Prevalence of germline mutations in cancer predisposition genes in patients with pancreatic cancer. *Gastroenterol* 2015; **148:** 556–64.

33 de Andrea CE, Zhu JF, Jin H, Bovee JV, Jones KB. Cell cycle deregulation and mosaic loss of Ext1 drive peripheral chondrosarcomagenesis in the mouse and reveal an intrinsic cilia deficiency. *J Pathol* 2015; **236:** 210–18.

34 Sawyer S, Mitchell G, McKinley J, et al. A role for common genomic variants in the assessment of familial breast cancer. *J Clin Oncol* 2012; **30:** 4330–36.

35 Villani A, Tabori U, Schiffman J, et al. Biochemical and imaging surveillance in germline *TP53* mutation carriers with Li-Fraumeni syndrome: a prospective observational study. *Lancet Oncol* 2011; **12:** 559–67.