Unique mutation portraits and frequent *COL2A1* gene alteration in chondrosarcoma

Yasushi Totoki, Akihiko Yoshida, Fumie Hosoda, et al.

Genome Res. published online July 14, 2014
Access the most recent version at doi:10.1101/gr.160598.113

Title: Unique mutation portraits and frequent *COL2A1* gene alteration in chondrosarcoma

Yasushi Totoki[1], Akihiko Yoshida[2], Fumie Hosoda[1], Hiromi Nakamura[1], Natsuko Hama[1], Koichi Ogura[3], Aki Yoshida[4], Tomohiro Fujiwara[3], Yasuhito Arai[1], Junya Toguchida[5], Hitoshi Tsuda[2], Satoru Miyano[6], Akira Kawai[3], Tatsuhiro Shibata[1, 7]

[1]*Division of Cancer Genomics, National Cancer Center Research Institute, Chuo-ku, Tokyo, 104-0045, Japan.*

[2]*Division of Pathology and Clinical Laboratories,* [3]*Division of Musculoskeletal Oncology, National Cancer Center Hospital, Chuo-ku, Tokyo, 104-0045, Japan*

[4]*Department of Orthopaedic Surgery, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, 700-8558, Japan*

[5]*Department of Tissue Regeneration, Institute for Frontier Medical Sciences, Kyoto University, Kyoto, 606-8507, Japan*

[6]*Laboratory of DNA Informatics Analysis, Human Genome Center, The Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo, 108-8639, Japan*

[7]*Correspondence should be addressed to Tatsuhiro Shibata, M.D., Ph.D., Division of Cancer Genomics, National Cancer Center Research Institute, 5-1-1, Tsukiji, Chuo-ku, Tokyo, 104-0045, Japan, E-mail: tashibat@ncc.go.jp, Telephone: +81-3-3547-5201, Fax: +81-3-3248-2463*

Abstract: 185 words

Main text excluding references: 3828 words

6 Figures and 2 Tables

Supplementary materials: 14 Figures, 11 Tables and Supplementary Information

Abstract

Chondrosarcoma is the second most frequent malignant bone tumor. However, the etiological background of chondrosarcomagenesis remains largely unknown, along with detailed information on molecular alterations, including potential therapeutic targets. Massively parallel paired-end sequencing of whole genomes of ten primary chondrosarcomas revealed that the process of accumulation of somatic mutations is homogeneous irrespective of pathological subtype or the presence of *IDH1* mutation and unique among a range of cancer types, and shares significant commonalities with that of prostate cancer. Clusters of structural alterations localized within a single chromosome were observed in four cases. Combined with targeted resequencing of additional cartilaginous tumor cohort, we identified somatic alterations of the *COL2A1* gene, which encodes an essential extracellular matrix protein in chondro-skeletal development, in 19.3% of chondrosarcoma and 31.7% of enchondroma cases. Epigenetic regulators (*IDH1* and *YEATS2*) and an activin/BMP signal component (*ACVR2A*) were recurrently altered. Furthermore, a novel *FN1-ACVR2A* fusion transcript was observed in both chondrosarcoma and osteochondromatosis cases. Under the characteristic accumulative process of somatic changes as a background, molecular defects in chondrogenesis and aberrant epigenetic control are primarily causative of both benign and malignant cartilaginous tumors.

Introduction

Chondrosarcoma accounts for more than 20% of primary bone sarcomas, with an overall incidence estimated at approximately one in 200,000 (Whelan et al. 2012). The patients are mostly older than 50 years and show male dominance. There are two common subtypes, central and peripheral. Central chondrosarcoma predominates (about 80%) and arises in the medullary cavity of the long bone, while peripheral chondrosarcoma (~15%) develops from the surface of the bone (Fletcher et al. 2002, Bovée et al. 2010). Clinically, low-grade chondrosarcomas rarely metastasize, and can be managed with local resection. By contrast, high-grade chondrosarcomas often metastasize and are lethal in most cases. Since the tumor cells exist in the specific microenvironments such as low vascularity and accumulated extracellular matrix, they are largely resistant to conventional chemotherapy and radiotherapy. Therefore identification of new therapeutic targets are required in this tumor.

Benign cartilage tumors (enchondroma and osteochondroma) may progress to chondrosarcoma (Bovée et al. 2010). Mutations of the exostosin 1 (*EXT1*) and *EXT2* genes are linked to hereditary and sporadic osteochondromatosis and are also reported in chondrosarcoma (Hecht et al. 1997, Wuyts et al. 1998). *EXT1* and *EXT2* regulate proper heparan sulphate proteoglycan processing and their defects cause abnormal diffusion of hedgehog ligands (Koziel et al. 2004). Mutations in the *PTH1R* gene was also identified in enchondroma, which disrupts the IHH–PTHLH feedback loop and also induces constitutive hedgehog signaling (Hopyan et al. 2002). Consistently

transgenic mice that express *Gli2*, a downstream effector of the hedgehog signal, in chondrocytes, develop cartilaginous lesions similar to human enchondromas (Hopyan et al. 2002).

In addition to hedgehog signaling, alterations of other molecular pathways such as *TP53*/*MDM2* (Wadayama et al. 1993, Larramendy et al. 1997), *AKT1* (Schrage et al. 2009) and *CDK4*/*RB* (Asp et al. 2001) have been reported in chndrosarcomas. Recently, frequent somatic mutations in isocitrate dehydrogenase 1 (*IDH1*) and *IDH2* have been identified in both enchondroma and central chondrosarcoma (Amary et al. 2011), and somatic mosaic *IDH1*/2 mutations are associated with multiple enchondromatosis (Pansuriya et al. 2011, Amary et al. 2011). These molecular and genetic observations support a close association between benign osteocartilaginous tumors and chondrosarcoma; however, the detailed molecular events and etiological risk factors underlying chondrosarcomagenesis remain largely unknown.

Results

Whole genome sequencing (WGS) of chondrosarcoma

Massively parallel paired-end sequencing of whole genomes of ten pairs of primary chondrosarcoma and matched normal muscle tissues was performed. The cases included five central, four peripheral, and one synovial, a rare subtype, cases (Table 1). The median sequence coverage was ×33.4 for tumor and ×31.4 for normal tissue (Table 1). All peripheral cases were associated with osteochondromatosis and harbored germline *EXT1* or *EXT2* mutations (Supplementary Table S1). No *IDH1/2* or other enchondromatosis-associated gene mutations (*PTH1R*, *PTPN11*, or *ACP5*) (Hopyan et al. 2002, Bowen et al. 2011, Briggs et al. 2011) were observed in the germline genomes. In total, 44,345 somatic single nucleotide variations (ranging from 2,215 to 7,014 per genome, 1.55/Mb on average) and 4,096 small insertions/deletions (indels, ranging from 269 to 544 per genome) were identified (Figure 1A, Supplementary Table S2). The somatic point mutations included 211 non-synonymous mutations (21.1 per genome on average) and 14 indels (1.4 per genome on average) in the coding regions (Supplementary Table S3). More than 91% (78/85) of somatic substitutions and 63% (17/27) of somatic indels were validated by Sanger sequencing. These mutations were significantly enriched in membranous proteins, especially those with transmembrane receptor activity (Supplementary Table S4).

Unique somatic substitution signatures in the chondrosarcoma genomes

Analysis of genome-wide somatic mutation signatures revealed that C:G>T:A

transitions are dominant, followed by T:A>C:G, T:A>A:T, and C:G>A:T substitutions in all chondrosarcoma cases (Figure 1A). These somatic substitution signatures at CpG sites were not associated with subtype classification or the presence of *IDH1* mutations (Supplementary Figures S1 and S2). Principal component analysis of trans-cancer genome data showed that the substitution pattern in these chondrosarcomas was most similar to those of prostate (Berger et al. 2011) and liver (Fujimoto et al. 2012) cancers and chronic lymphocytic leukemia (Puente et al. 2011) (Figure 1B, permutation test; P=0.0010). A significant reduction in C:G>A:T transversions on the transcribed strand was observed in both central (CS-2T, 7T, 8T, and 9T) and peripheral (CS-1T, 3T, and 10T) cases (Figure 1C, Supplementary Figure S3), which correlated with gene expression level (Figure 1D). To explore any sequence-context dependent substitutions in the chondrosarcoma genomes, we measured the frequencies of immediate 5' and 3' nucleotides for all substitutions. This analysis revealed significant increases in C>T transitions at TpCpT, C>A transversions at ApCpA, and T>A transversions at ApTpA in all cases except CS5T (Figure 2A, Supplement Figure S4). No context-specific T>C transitions were observed. This triplet landscape differs from those of liver cancer and CLL and those caused by known etiological factors such as C>T in UV-associated melanoma (Pleasance et al. 2010a) or C>A in smoking-associated lung cancer (Pleasance et al. 2010b), but shares significant commonalities with that of prostate cancer (permutation test; P=0.0017) (Figure 2B, Supplementary Figures S5 and S6). A further context survey of the 10 nucleotides stretch in 5' and 3' direction from each somatic substitution identified a predominance of A/T around the sites of the C>A

substitutions (particularly on the 3' side), and that T was dominant at either side of the C>T substitution (Figure 3). This pattern was also observed in prostate cancer but not in melanoma and smoking-associated lung cancer (Figure 3, Supplementary Figure S7).

Structural alterations in chondrosarcoma

We determined copy number and structural alterations by analyzing sequence depth and paired sequence reads (Supplementary Tables S5 and S6). We predicted 350 rearrangements in 10 chondrosarcomas. Validation analysis by genomic PCR and Sanger sequencing of randomly selected breakpoints verified more than 98% of predictions (67/68) as somatic. No recurrent rearrangements were detected. Remarkably complex rearrangement, a cluster of structural alterations localized within a single chromosome, was observed in four cases (Supplementary Figures S8 and S9A-C). Massive rearrangements were involved in CS2T with amplification in the short arm of chromosome 5 including the *TERT* gene (Figures 4A and 4B). Interstitial deletion of exons 2 and 3 and a resulting premature stop codon in the *WNK2* tumor suppressor gene (Hong C et al 2007), which encodes a negative regulator of the MEK/ERK pathway (Moniz et al. 2007), was detected in CS6T (Figure 4C).

Novel driver genes in chondrosarcoma

To explore driving alterations in chondrosarcomagenesis, we calculated the expected number of somatic non-synonymous and splice site substitutions, coding indels and rearrangements with the adjustment for background mutation rate and gene length in

each gene, and identified five recurrently altered genes (*IDH1*, *TP53*, *ACVR2A*, *COL2A1* and *YEATS2*) with a false discovery rate of <1% (Supplementary Tables S7). Among these, recurrent mutations of the *IDH1* and *TP53* genes have been previously reported in chondrosarcoma (Amary et al. 2011, Wadayama et al. 1993). To validate mutation frequency in a larger number of cases, we performed target exon resequencing of three potential new driver genes (*COL2A1*, *YEATS2*, and *ACVR2A*) and *IDH1/2* in additional 47 chondrosarcoma with 19 corresponding adjacent non-tumor tissues, and 41 enchondroma samples (Figure 5, Supplementary Tables S8 and S9). *COL2A1* encodes the alpha-1 chain of type II collagen, which is enriched in cartilage (Cheah et al. 1985) and was mutated in eleven chondrosarcoma (19.3%) and thirteen enchondroma (31.7%) cases. Structural alterations (intragenic tandem duplication and deletion) of the *COL2A1* gene were also detected. *YEATS2* alterations including one nonsense mutation and intra-chromosomal inversion were detected in seven chondrosarcoma (12.3%) and one enchondroma (2.4%) cases. *YEATS2* is a scaffolding subunit of the nuclear acetyltransferase complex which targets histone H3 and represses transcription of target genes (Wang et al. 2008), and a nonsense mutation (p.W1229*) disrupts the regulatory histone fold module (Figure 5B). They are significantly more frequent in higher grade (grade 2 and 3) cases than grade 1 chondrosarcoma/enchondroma cases (P=0.023) and tend to be mutually exclusive to *COL2A1* mutation (P=0.07). *ACVR2A* mutations were observed in four chondrosarcoma (7%) and three enchondroma (7.3%) cases. *IDH1* mutations at codon 132 were observed in twelve chondrosarcoma (21.1%) and five enchondroma (12.2%) cases, and were not associated with any other mutations. The

presence of *COL2A1* or *IDH1* mutation was not associated with patients' prognosis (Supplementary Figure S10).

Novel fusion gene in chondrosarcoma

Four in-frame fusion transcripts were predicted by structural rearrangements (Table 2). Among these, two (*ACVR2A-FN1* and *FN1-ACVR2A*) were expressed and validated in the synovial CS6T case (Figure 6A-C, Supplementary Figure S11). Paired-end whole transcriptome sequencing demonstrated that expression of the *ACVR2A* gene transcript increased sharply, more than 10-fold, in the fused exons to the *FN1* gene (Figure 6D), without any overt change in *FN1* gene expression at the fusion junction (Supplementary Figure S12). This suggests that promoter activity of the *FN1* gene conferred increased *FN1-ACVR2A* transcript in the tumor. RNA sequencing further revealed that expression of the *ACVR2A* gene increased more than 25-fold in the *FN1-ACVR2A* fusion positive case (CS6T) compared to that of other samples (CS1T and CS7T) (Supplementary Table S10). Since this fusion gene was detected in a chondrosarcoma with features suggestive of preexisting synovial osteochondromatosis (Figure 6A), we further surveyed a range of cartilaginous tumors including extra-medullary ones (17 chondrosarcomas, 7 synovial chondromatoses, and 13 osteochondromas) by RT-PCR, and detected expression of the *FN1-ACVR2A*, but not the *ACVR2A-FN1*, fusion transcript in a case of osteochondromatosis (Figure 6E, Supplementary Figure S13).

Discussion

The present study revealed comprehensive landscapes of genetic alterations in the chondrosarcoma genome that include global mutational signatures, structural alterations including copy number changes and rearrangements, and new driver genes, by whole genome sequencing (WGS) approach.

Somatic mutation signatures are affected by both environmental carcinogen exposures and defects in DNA repair systems (Stratton et al. 2009, Nik-Zainal et al 2012). Previous WGS analyses of small cell lung cancer and melanoma cell lines showed an intimate association between specific carcinogens (smoking or UV exposure) and the patterns of base substitutions (Pleasance et al. 2010a, 2010b). Furthermore WGS of breast cancers have reported that *BRCA1*-mutated breast cancers showed a unique mutational signature (Nik-Zainal et al. 2012). Our analysis showed that C:G>T:A transitions with significant transcription-coupled repair are broadly observed in chondrosarcoma, which is also characteristic of smoking-associated lung cancer (Pleasance et al 2010b, Govindan et al. 2012). Nonetheless, further mutation trait analysis demonstrated that the mutagenesis process in chondrosarcoma seems to be distinct from mutagenesis induced by known carcinogens, including tobacco smoking. The presence of an *IDH1* mutation is associated with epigenetic alterations such as CpG hypermethylation in tumors, including enchondroma (Turcan et al. 2012, Pansuriya et al. 2011). It could be possible that frequent CpG methylation induces chance of somatic C to T transition by spontaneous deamination of 5-methylcytosine. However, the genome

sequencing described here revealed that the *IDH1* mutation exerts no significant effect on the somatic mutation frequency at CpG sites during chondrosarcomagenesis.

Our analysis uncovered the unexpected similarity of the mutation portraits (C>T transitions at TpCpT, C>A transversions at ApCpA, and T>A transversions at ApTpA) of chondrosarcoma and prostate cancer, both of which frequently occur in older and male patients. Nine out of ten cases showed similar mutational signatures and these nine cases harbors distinct set of somatic mutations, suggesting that this signature could be caused by common etiological factors but not affected by somatic deficiency in the DNA repair system. The etiological risk factors including genetic susceptibility for chondrosarcoma remain unknown partly because of the small number of patients, but our observation implies that aging, hormonal status, or dietary factors, all of which are suggestive etiological factors of prostate cancer (Hsing et al. 2006), or genetic risk factors, may be associated with this sarcoma.

WGS of ten cases followed by further validation of a larger cohort identified few common alterations (*COL2A1* and *IDH1*) and other rarer events, the latter of which include progression-associated changes such as *YEATS2*, in chodrosarcomagenesis. Especially *COL2A1*, which is rarely mutated in other tumor types, emerged as a new frequently altered gene in chondrosarcoma. The pattern of *COL2A1* mutation in this study is different from that reported in the previous study (Tarpey et al. 2013). A preponderance of missense mutation has been observed in the present study whereas

truncating mutations were predominant in the previous report. This discrepancy could be due to the difference in ethnicity (Japanese and Caucasian) or unknown carcinogenesis background of each cohort, and larger collection of samples by the international collaboration will be required to explore this. Germline *COL2A1* mutations are associated with a number of chondro-skeletal malformation syndromes, so-called type II collagenopathies (Spranger et al. 1994, Nishimura et al. 2005), including spondyloepiphyseal dysplasia congenita, achondrogenesis type II, and osteoarthritis associated with chondrodysplasia. *Col2a1* mutant mice demonstrated growth plate disorganization with reduced elaborate collagen fibrils (Esapa et al. 2012). However, no clinical association between these congenital diseases and cartilaginous tumors has been reported so far. The presence of frequent *COL2A1* as well as *IDH1/2* mutations in both chondrosarcoma and enchondroma demonstrated by the present and previous studies (Amary et al. 2011) supports a model of progression from enchondroma to chondrosarcoma. Our analysis also identified that aberrations in the epigenetic regulators play important roles in cartilaginous tumors. The presence of an *IDH1* mutation is associated with CpG hypermethylation in tumors, including enchondroma (Turcan et al. 2012, Pansuriya et al. 2011). *YEATS2* is a scaffolding subunit of the nuclear acetyltransferase complex and harbors a histone-like module which interacts with TATA-binding protein and negatively regulates gene transcription (Wang et al. 2008). Somatic mutations in the *YEATS2* gene have also been reported in lung (3.8%), colorectal (1.8%) and endometrial (1.8%) cancers in COSMIC database (Forbes et al. 2011).

No disease-specific common fusion gene was detected in our cohort of chondrosarcomas. However, analysis of a synovial subtype identified recurrent *FN1-ACVR2A* fusion transcript, which further suggests a molecular link between chondrosarcoma and osteochondromatosis. *ACVR2A* encodes a membranous serine/threonine protein kinase which functions as a receptor for the activin and bone morphogenetic proteins 4 and 6 (Donaldson et al. 1992) and is involved in skeletal development (Matzuk et al. 1995). *FN1-ALK* fusion transcript has been identified as a potential therapeutic target in ovarian cancer (Ren et al. 2012).

Collectively our analysis demonstrated that molecular defects in chondrocyte differentiation and epigenetic regulators are synergistically causative of both benign and malignant cartilaginous tumors (Supplementary Figure S14). Recent study reported the potential efficacy of an *IDH1* inhibitor for treating *IDH1*-mutated tumors (Rohle et al. 2013). Aberrant *FN1-ACVR2A* gene fusion and recurrent *ACVR2A* mutations also warrant small-molecular inhibitors targeting activin receptor kinases (Harrison et al. 2005) as potential therapeutic modalities against a subset of chondrosarcoma.

Methods

Clinical samples

The clinical and pathological features of 10 patients and their tumors are shown in Table S1. Tissue samples were provided by the National Cancer Center Biobank, Japan. High-molecular weight genomic DNA and RNA were extracted from fresh frozen tumor specimens and non-cancerous muscle tissue. The study protocol was in agreement with the Ethical Committee of the National Cancer Center, Tokyo, Japan.

Whole-genome and transcriptome sequencing

We prepared 300–500 bp (base pairs) insert libraries from 3 μg of genomic DNA and 150–200 bp insert libraries from 2 μg of total RNA using the TruSeq DNA Sample Preparation Kit and mRNA-Seq Sample Preparation Kit (Illumina), respectively. The libraries were subjected to paired-end sequencing of 100 bp on the HiSeq 2000 (Illumina), according to the manufacturer's instructions.

Detection of somatic point mutations and short indels

Paired-end reads were aligned to the human reference genome (GRCh37) using the Burrows-Wheeler Aligner (BWA) (Li and Durbin 2009a). Probable PCR duplications, in which paired-end reads aligned to the same genomic positions, were removed using SAMtools (Li et al. 2009b) and a program developed in-house. To find somatic point mutations and short indels, SAMtools was applied with stringent confidence filtering conditions we developed (Totoki et al. 2011). The details of our filtering conditions are

described in the Supplementary Information.

Validation of candidate driver genes in an additional cohort

To validate the mutation frequencies of *COL2A1*, *YEATS2*, and *ACVR2A* with recurrent mutations in benign and malignant cartilaginous tumors, we amplified all protein-coding exons of those genes using formalin-fixed, paraffin-embedded (FFPE) DNA from 47 chondrosarcoma, 19 corresponding adjacent non-tumor tissue, and 41 enchondroma samples. Six sequencing libraries were prepared from 115 amplicon mixtures of pooled DNA from three chondrosarcomas, pooled normal DNA, and pooled DNA from two enchondromas. The 115 amplicons covered a total of 20 kb of coding regions of the three genes. The six libraries were subjected to paired-end sequencing of 100 bp using an Illumina GA IIx sequencer. Paired-end reads were aligned to the human reference genome (GRCh37) using BWA, and somatic mutations were called using SAMtools (Li et al. 2009b) and programs developed in-house. All candidate 137 SNVs and 13 short indels for *ACVR2A*, *COL2A1* and *YEATS2* and the mutational hot spots for *IDH1* and *IDH2* were further verified in individual case by MassARRAY system (Sequenom). The primer sets, which include a pair of amplicon primers and an extension primer for each SNV, were designed using the MassARRAY Designer software (Sequenom) (Supplementary Table S11). The details of our filtering conditions of the mutation call and the verification by MassARRAY system are described in the Supplementary Information.

Fusion gene validation by RT-PCR and sequencing

Total RNA was reverse-transcribed to cDNA using SuperScript III (Invitrogen). cDNA was subjected to PCR amplification using Ex Taq (Takara Bio). The PCR products were directly sequenced, in both directions, by Sanger sequencing using the BigDye terminator kit (Applied Biosystems).

Data Access

Sequence and mutation/indel data have been submitted to the European Genome-phenome Archive (EGA; https://www.ebi.ac.uk/ega/) under accession number EGAS00001000505.

Acknowledgements

This study was supported by the Program for Promotion of Fundamental Studies in Health Sciences from the National Institute of Biomedical Innovation (NIBIO), National Cancer Center Research and Development Funds (23-A-8 and 23-A-10), Grant in Aid for Scientific Research B (22390296) and Health Labor Sciences Research Grant. National Cancer Center Biobank is supported by the National Cancer Center Research and Development Fund, Japan. We thank Dr. Toshifumi Ozaki, Department of Orthopaedic Surgery, Okayama University Medical School for samples, Drs. S. Hosokawa and H. Iwama, Sequenom Inc. for MassARRAY data analysis and Ms. T. Urushidate, N. Okada, A. Kokubu, H. Shimizu and S. Ohashi for technical assistance. The supercomputing resource SHIROKANE was provided by the Human Genome Center at The University of Tokyo.

Author Contributions:

Conception and design: YT, AY, AK, TS

Provision of study materials or patients: AY, KO, AY, TF, HT

Data analysis and interpretations: YT, AY, FH, HN, NH, YA, JT, SM, TS

Manuscript writing: YT, AY, AK, TS

Authors' disclosure of potential conflicts of interest

The authors declare no conflicts of interest.

Figure legends

Figure 1 Somatic mutation signature of chondrosarcoma

(A) Number of somatic substitutions and indels in 10 chondrosarcoma cases (upper panel). Percentage of six somatic substitutions in each case (lower panel). (B) Principal component analysis of six substitution patterns in various cancer genomes including chondrosarcoma. Red dots indicate each chondrosarcoma case from the present study. (C) Number of somatic substitutions on transcribed (blue) and untranscribed (red) strands. Significant reduction of C>A substitution was noted on the transcribed strand. **, P<0.01; ****, P<0.0001. (D) Negative correlation between number of C>A substitutions on transcribed strand and gene expression level. Blue, C>A substitution; Solid line, Untranscribed strand; Dotted line, Transcribed strand.

Figure 2 Somatic mutation portraits in the chondrosarcoma genome

(A) Frequencies of 96 mutation portraits (combination of immediate 5' and 3' bases with six substitutions) in peripheral (CS1T) and central (CS2T and CS5T) cases. Sixteen triplet sequence patterns (mutation with immediate 5' and 3' nucleotides) for each substitution are indicated by different color columns. (B) Frequencies of 96 mutation portraits in chondrosarcoma (average), prostate cancer (average), UV-associated melanoma, and smoking-associated lung cancer genomes.

Figure 3 Characteristic base contexts at C>A and C>T mutations in chondrosarcoma

Sequence logos of consensus surrounding sequences enriched at C>A (A) and C>T (B) mutations in chondrosarcoma (top left), prostate cancer (bottom left), melanoma (top right), and smoking-associated small-cell lung cancer (bottom right). The X-axis indicates base position with respect to the mutation site (center, mutation; left, 5' of the mutation; right, 3' of the mutation). The Y-axis indicates the information content at each position in the sequence.

Figure 4 Structural alterations in chondrosarcoma

(A) Complex rearrangements in CS2T indicated by Circos plot (Krzywinski et al. 2009). Structural alterations (Red line, deletion; Green line, inversion; Blue line, tandem duplication; Purple line, translocation) are shown in the inner circle. Copy number changes (green, copy gain/amplification; red, copy loss) are shown in the outer circle. Thick blue lines indicate regions of localized accumulation of structural alterations. (B) Massive structural alterations in the short arm of chromosome 5 in CS2T. The location of the *TERT* gene is indicated by an arrow. (C) Interstitial deletion in the *WNK2* gene in CS6T. Schematic representation of exon 2-3 deletion in the *WNK2* gene and its protein product (top). Validation of the predicted deletion by genomic PCR and RT-PCR (lower left). Sanger sequencing validation of the exon 2-3 deleted transcript (lower right).

Figure 5 Mutation landscape of chondrosarcoma

(A) Enchondroma and chondrosarcoma samples were displayed with the presence of the *IDH1/2*, *COL2A1*, *YEATS2*, and *ACVR2A* genes, tumor subtype and grade (G1-G3).

Q value for mutated genes was shown on the right. (B) Distribution of mutations and rearrangements in the coding regions of *COL2A1*, *YEATS2*, and *ACVR2A* genes. Black and blue arrowheads indicate missense mutations in chondrosarcoma and enchondroma, respectively. Red and green arrowhead indicates non-sense mutation and the position of intrachromosomal inversion. Asterisks indicate mutations verified as somatically acquired by sequencing the corresponding normal genome. Colored boxes indicate functional domains in each molecule (green, signal peptide; blown, von Willebrand factor type C domain; orange, collagen triple helix repeat; purple, fibrillar collagen C-terminal domain; blue, Coiled-coil domain; light green, *YEATS* family domain; gray, histone fold domain; yellow, activin types I and II receptor domain; black, transmembrane domain; red, serine/threonine protein kinase domain).

Figure 6 *FN1-ACVR2A* fusion gene in chondrosarcoma

(A) Synovial chondrosarcoma case. Left panel, radiograph revealed a calcified soft tissue mass surrounding the knee joint without bone involvement. Right lower panel, this hypercellular myxoid tumor diffusely infiltrated periarticular and subcutaneous tissues dissecting through fat and collagen (scale bar indicates 500 μm). Right lower panel, the tumor focally resembled synovial chondromatosis (scale bar indicates 1 mm). The tumor later metastasized to the inguinal lymph node. (B) Schematic presentation of an intrachromosomal inversion in chromosome 2q (left) and two in-frame fusion proteins generated by this rearrangement. (C)

Validation of fusion transcripts (left, *ACVR2A-FN1*, and right, *FN1-ACVR2A*) by RT-PCR and Sanger sequencing. (D) Estimation of *ACVR2A* gene expression at nucleotide resolution by counting RNA sequencing reads. Note the sharp increase in *AVCR2A* expression after the fusion point. (E) Left, histology of multiple osteochondroma case. This tibial tumor is composed of exophytic bony growth capped by benign hyaline cartilage. Scale bar indicates 1mm. Right, validation of the *FN1-ACVR2A* fusion transcript by RT-PCR and Sanger sequencing in an osteochondromatosis case (1804T).

Table 1

Summary of clinical data and somatic alterations of 10 chondrosarcoma cases analyzed by whole-genome sequencing.

Table 2

Inframe fusion genes detected in chondrosarcoma

References

Amary MF, Bacsi K, Maggiani F, Damato S, Halai D, Berisha F, Pollock R, O'Donnell P, Grigoriadis A, Diss T, *et al.* 2011a. IDH1 and IDH2 mutations are frequent events in central chondrosarcoma and central and periosteal chondromas but not in other mesenchymal tumours. *J Patho* **224**: 334-343.

Amary MF, Damato S, Halai D, Eskandarpour M, Berisha F, Bonar F, McCarthy S, Fantin VR, Straley KS, Lobo S, *et al.* 2011b. Ollier disease and Maffucci syndrome are caused by somatic mosaic mutations of IDH1 and IDH2. *Nat Genet* **43**: 1262-1265.

Asp J, Inerot S, Block JA, Lindahl A. 2001. Alterations in the regulatory pathway involving p16, pRb and cdk4 in human chondrosarcoma. *J Orthop Res* **19**: 149-154.

Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K, Sivachenko AY, Sboner A, Esgueva R, Pflueger D, Sougnez C, *et al.* 2011. The genomic complexity of primary human prostate cancer. *Nature* **470**: 214-220.

Bovée JV, Hogendoorn PC, Wunder JS, Alman BA. 2010. Cartilage tumours and bone development: molecular pathology and possible therapeutic targets. *Nat Rev Cancer* **10**: 481-488.

Bowen ME, Boyden ED, Holm IA, Campos-Xavier B, Bonafé L, Superti-Furga A, Ikegawa S, Cormier-Daire V, Bovée JV, Pansuriya TC, *et al*. 2011. Loss-of-function mutations in PTPN11 cause metachondromatosis, but not Ollier disease or Maffucci syndrome. *PLoS Genet* **7**: e1002050.

Briggs TA, Rice GI, Daly S, Urquhart J, Gornall H, Bader-Meunier B, Baskar K, Baskar S, Baudouin V, Beresford MW, *et al*. 2011. Tartrate-resistant acid phosphatase deficiency causes a bone dysplasia with autoimmunity and a type I interferon expression signature. *Nat Genet* **43**: 127-131.

Cheah KS, Stoker NG, Griffin JR, Grosveld FG, Solomon E. 1985. Identification and characterization of the human type II collagen gene (COL2A1). *Proc Natl Acad Sci U S A* **82**: 2555-2559.

Donaldson CJ, Mathews LS, Vale WW. 1992. Molecular cloning and binding properties of the human type II activin receptor. *Biochem Biophys Res Commun* **184**: 310-316.

Esapa CT, Hough TA, Testori S, Head RA, Crane EA, Chan CP, Evans H, Bassett JH, Tylzanowski P, McNally EG, *et al*. 2012. A mouse model for spondyloepiphyseal dysplasia congenita with secondary osteoarthritis due to a Col2a1 mutation. *J Bone Miner Res* **27**: 413-428.

Fletcher CDM, Unni K, Mertens F. 2002. Pathology and genetics of tumours of soft tissues and bone. IARC Press, Lyon.

Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, *et al*. 2011. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* **39(Database issue)**: D945-950.

Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, Aoki M, Hosono N, Kubo M, Miya F, *et al.* 2012. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in

chromatin regulators. *Nat Genet* **44**: 760-764.

Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, Maher CA, Fulton R, Fulton L, Wallis J, *et al*. 2012. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**: 1121-1134.

Harrison CA, Gray PC, Vale WW, Robertson DM. 2005. Antagonists of activin signaling: mechanisms and potential biological applications. *Trends Endocrinol Metab* **16**: 73-78.

Hecht JT, Hogue D, Wang Y, Blanton SH, Wagner M, Strong LC, Raskind W, Hansen MF, Wells D. 1997. Hereditary multiple exostoses (EXT): mutational studies of familial EXT1 cases and EXT-associated malignancies. *Am J Hum Genet* **60**: 80-86.

Hong C, Moorefield KS, Jun P, Aldape KD, Kharbanda S, Phillips HS, Costello JF. 2007. Epigenome scans and cancer genome sequencing converge on WNK2, a kinase-independent suppressor of cell growth. *Proc Natl Acad Sci U S A* **104**: 10974-10979.

Hopyan S, Gokgoz N, Poon R, Gensure RC, Yu C, Cole WG, Bell RS, Jüppner H, Andrulis IL, Wunder JS, *et al*. 2002. A mutant PTH/PTHrP type I receptor in enchondromatosis. *Nat Genet* **30**: 306-310.

Hsing AW, Chokkalingam AP. 2006. Prostate cancer epidemiology. *Front Biosci* **11**: 1388-1413.

Koziel L, Kunath M, Kelly OG, Vortkamp A. 2004. Ext1-dependent heparan sulfate regulates the range of Ihh signaling during endochondral ossification. *Dev Cell* **6**: 801-813.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res* **19**:1639-1645.

Larramendy ML, Tarkkanen M, Valle J, Kivioja AH, Ervasti H, Karaharju E, Salmivalli T, Elomaa I, Knuutila S. 1997. Gains, losses, and amplifications of DNA sequences evaluated by comparative genomic hybridization in chondrosarcomas. *Am J Pathol* **150**: 685-691.

Li H, Durbin R. 2009a. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754-1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Matth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009b. The Sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078-2079.

Matzuk MM, Kumar TR, Bradley A. 1995. Different phenotypes for mice deficient in either activins or activin receptor type II. *Nature* **374**: 356-360.

Meijer D, de Jong D, Pansuriya TC, van den Akker BE, Picci P, Szuhai K, Bovée JV. 2012. Genetic characterization of mesenchymal, clear cell, and dedifferentiated chondrosarcoma. *Genes Chromosomes Cancer* **51**: 899-909.

Moniz S, Veríssimo F, Matos P, Brazão R, Silva E, Kotelevets L, Chastre E, Gespach C, Jordan P. 2007. Protein kinase WNK2 inhibits cell proliferation by negatively modulating the activation of MEK1/ERK1/2. *Oncogene* **26**: 6071-6081.

Nishimura G, Haga N, Kitoh H, Tanaka Y, Sonoda T, Kitamura M, Shirahama S, Itoh T, Nakashima E, Ohashi H, *et al.* 2005. The phenotypic spectrum of COL2A1

mutations. *Hum. Mutat* **26**: 36-43.

Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, *et al*. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**: 979-993.

Pansuriya TC, van Eijk R, d'Adamo P, van Ruler MA, Kuijjer ML, Oosting J, Cleton-Jansen AM, van Oosterwijk JG, Verbeke SL, Meijer D, *et al.* 2011. Somatic mosaic IDH1 and IDH2 mutations are associated with enchondroma and spindle cell hemangioma in Ollier disease and Maffucci syndrome. *Nat. Genet* **43**: 1256-1261.

Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, Lin ML, Beare D, Lau KW, Greenman C, *et al.* 2010a. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463:** 184-190.

Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordóñez GR, Bignell GR, *et al.* 2010b. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463**: 191-196.

Puente XS, Pinyol M, Quesada V, Conde L, Ordóñez GR, Villamor N, Escaramis G, Jares P, Beà S, González-Díaz M, *et al*. 2011. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475**: 101-105.

Ren H, Tan ZP, Zhu X, Crosby K, Haack H, Ren JM, Beausoleil S, Moritz A, Innocenti G, Rush J, *et al*. 2012. Identification of anaplastic lymphoma kinase as a potential therapeutic target in ovarian cancer. *Cancer Res* **72**: 3312-3123.

Rohle D, Popovici-Muller J, Palaskas N, Turcan S, Grommes C, Campos C, Tsoi J, Clark O, Oldrini B, Komisopoulou E, *et al*. 2013. An inhibitor of mutant IDH1

delays growth and promotes differentiation of glioma cells. *Science* **340**: 626-630.

Schrage YM, Briaire-de Bruijn IH, de Miranda NF, van Oosterwijk J, Taminiau AH, van Wezel T, Hogendoorn PC, Bovée JV. 2009. Kinome profiling of chondrosarcoma reveals SRC-pathway activity and dasatinib as option for treatment. *Cancer Res* **69**: 6216-6222.

Spranger J, Winterpacht A, Zabel B. 1994. The type II collagenopathies: a spectrum of chondrodysplasias. *Eur. J. Pediatr* **153**: 56-65.

Stratton MR, Campbell PJ, Futreal PA. 2009. The cancer genome. *Nature* **458**: 719-724.

Tarpey PS, Behjati S, Cooke SL, Loo PV, Wedge DC, Pillay N, Marshall J, Meara SO, Davies H, Zainal SN, *et al*. 2013. Frequent mutation of the major cartilage collagen gene COL2A1 in chondrosarcoma. *Nat. Genet* **45**: 923-926.

Totoki Y, Tatsuno K, Yamamoto S, Arai Y, Hosoda F, Ishikawa S, Tsutsumi S, Sonoda K, Totsuka H, Shirakihara T, *et al*. 2011. High-resolution characterization of a hepatocellular carcinoma genome. *Nat Genet* **43**: 464-469.

Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, Campos C, Fabius AW, Lu C, Ward PS *et al.* 2012. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483**: 479-483.

Wadayama B, Toguchida J, Yamaguchi T, Sasaki MS, Yamamuro T. 1993. p53 expression and its relationship to DNA alterations in bone and soft tissue sarcomas. *Br J Cancer* **68**: 1134-1139.

Wang YL, Faiola F, Xu M, Pan S, Martinez E. 2008. Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold

module that interacts with the TATA-binding protein. *J Biol Chem* **283**: 33808-33815.

Whelan J, McTiernan A, Cooper N, Wong YK, Francis M, Vernon S, Strauss SJ. 2012. Incidence and survival of malignant bone sarcomas in England 1979-2007. *Int J Cancer* **131**: E508-517.

Wuyts W, Van Hul W, De Boulle K, Hendrickx J, Bakker E, Vanhoenacker F, Mollica F, Lüdecke HJ, Sayli BS, Pazzaglia UE, *et al.*1998. Mutations in the EXT1 and EXT2 genes in hereditary multiple exostoses. *Am J Hum Genet* **62**: 346-354.

Table 1

	CS01	CS02	CS03	CS04	CS05	CS06	CS07	CS08	CS09	CS10
Age	14	32	58	43	62	86	50	58	54	33
Gender	F	M	M	M	F	F	F	F	F	M
Tumor Site	Rib	Pelvis	Pelvis	Pelvis	Humerus	Knee	Rib	Pelvis	Scapula	Tibia
Subtype	Peripheral	Central	Peripheral	Peripheral	Central	Synovial	Central	Central	Central	Peripheral
Tumor Grade	1	2	1	1	3	2	2	2	3	1
Clinical Background	Multiple osteochondromatosis		Multiple osteochondromatosis	Multiple osteochondromatosis						Multiple osteochondromatosis
Average squence depth (Tumor)	42.6	34.2	27.6	30.8	33.4	37.7	37.2	37.2	29	24.7
Average sequence depth (Normal)	32.7	31.6	25.4	32.9	32.8	32.7	27.6	31.6	35	31.6
Number of somatic substitutions	3,521	5,005	6,128	2,215	5,043	3,705	4,100	5,326	2,288	7,014
Number of nonsynonymous substitution	10	26	37	7	26	11	29	26	12	27
Rearrangement break points	66	60	45	44	24	23	1	40	10	37

Table 2

Fusion gene	Sample	Rearrangement Type	Chromosom 1	Position 1	Cromosome 2	Position 2
PDE1C - MACC1	CS01T	deletion	chr7	20,201,367	chr7	32,263,702
LIPI - FAM176C	CS03T	inversion	chr21	15,528,857	chr21	33,858,432
ACVR2A - FN1	CS06T	inversion	chr2	148,646,674	chr2	216,289,052
FN1 - ACVR2A	CS06T	inversion	chr2	148,646,754	chr2	216,289,134

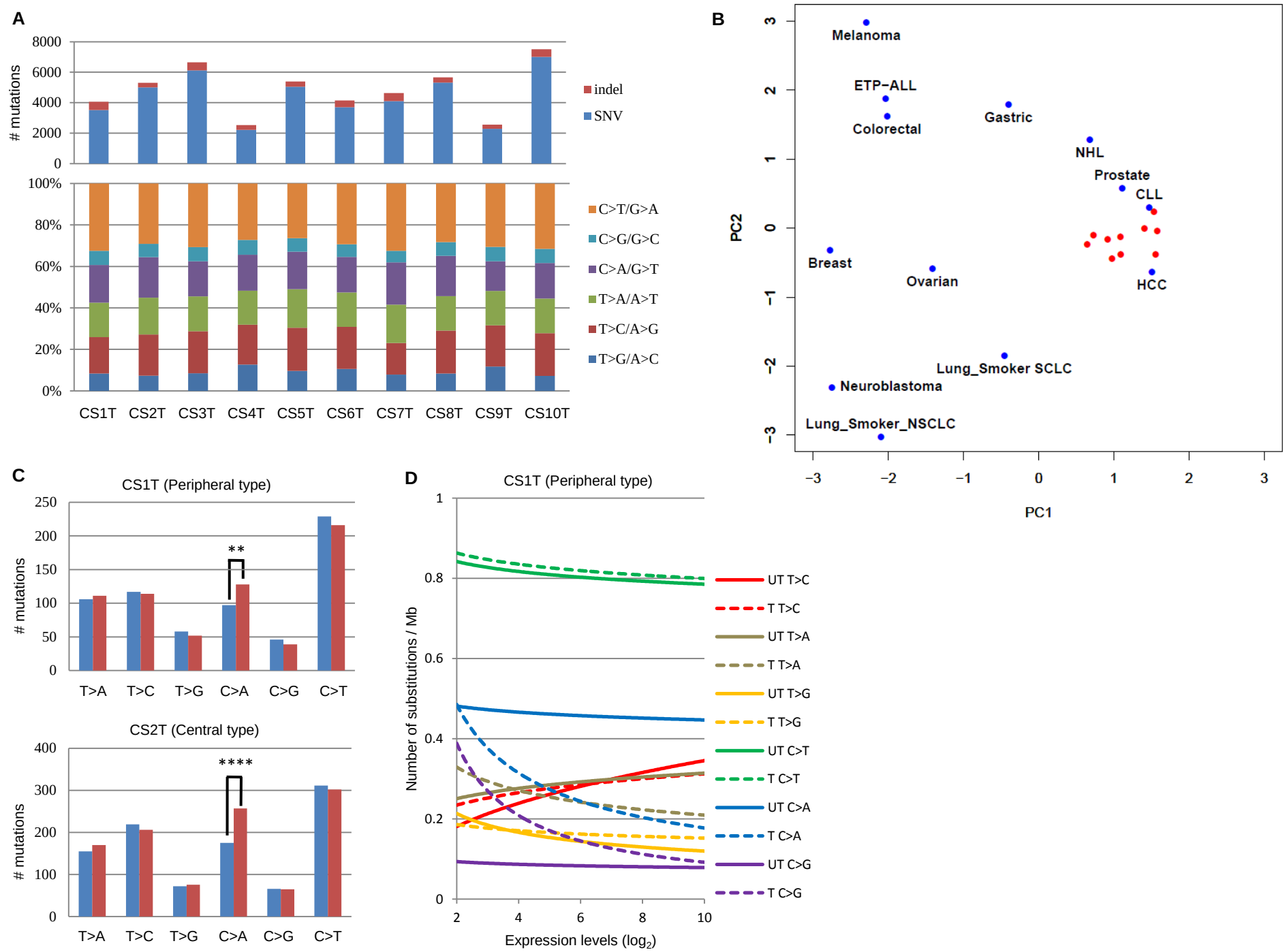
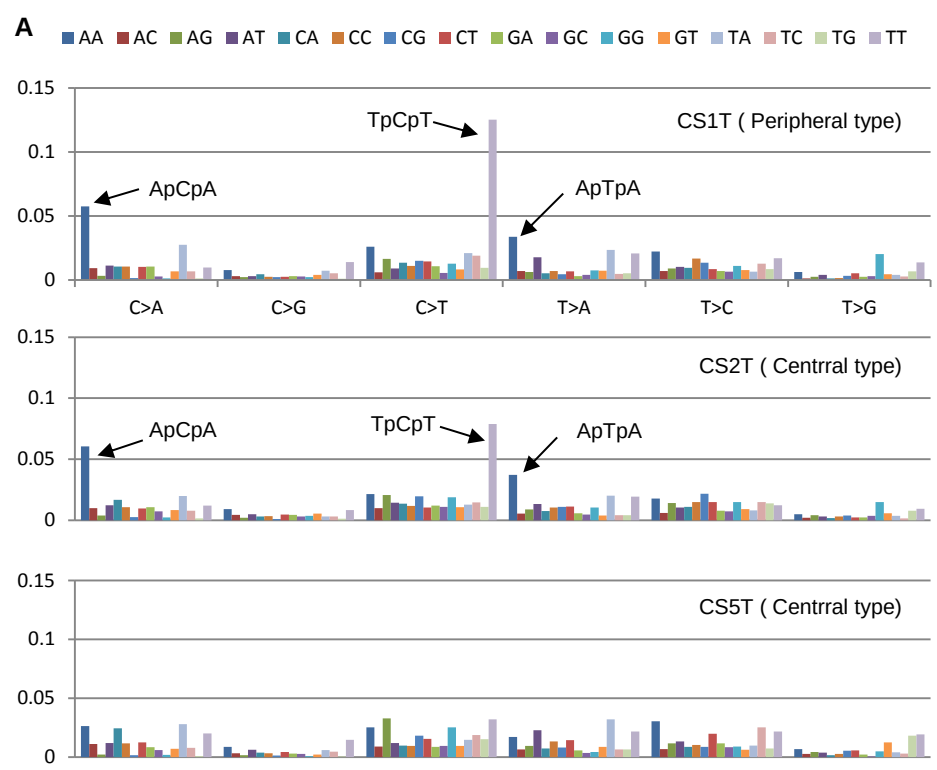

A
mutations
indel
SNV
C>T/G>A
C>G/G>C
C>A/G>T
T>A/A>T
T>C/A>G
T>G/A>C
CS1T CS2T CS3T CS4T CS5T CS6T CS7T CS8T CS9T CS10T
B
Melanoma
ETP−ALL
Colorectal
Gastric
NHL
Prostate
CLL
Breast
Ovarian
HCC
Lung_Smoker SCLC
Neuroblastoma
Lung_Smoker_NSCLC
PC1
PC2
C
CS1T (Peripheral type)
CS2T (Central type)
**
****
T>A T>C T>G C>A C>G C>T
D
CS1T (Peripheral type)
Number of substitutions / Mb
Expression levels (log2)
UT T>C
T T>C
UT T>A
T T>A
UT T>G
T T>G
UT C>T
T C>T
UT C>A
T C>A
UT C>G
T C>G

Figure 1


A
AA AC AG AT CA CC CG CT GA GC GG GT TA TC TG TT
CS1T (Peripheral type)
TpCpT
ApCpA
ApTpA
CS2T (Centrral type)
CS5T (Centrral type)
0.15
0.1
0.05
0
C>A
C>G
C>T
T>A
T>C
T>G

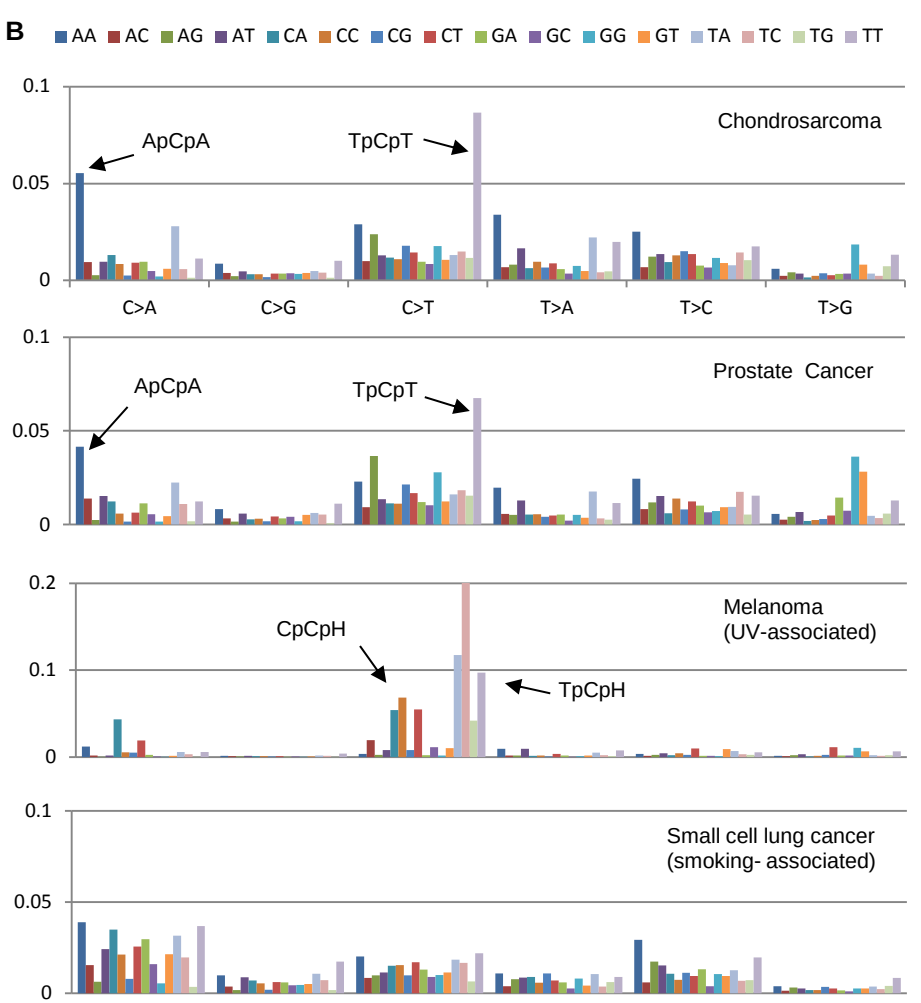

B
AA AC AG AT CA CC CG CT GA GC GG GT TA TC TG TT
Chondrosarcoma
ApCpA
TpCpT
Prostate Cancer
Melanoma
(UV-associated)
CpCpH
TpCpH
Small cell lung cancer
(smoking- associated)
0.2
0.1
0.05
0
C>A
C>G
C>T
T>A
T>C
T>G

Figure 2

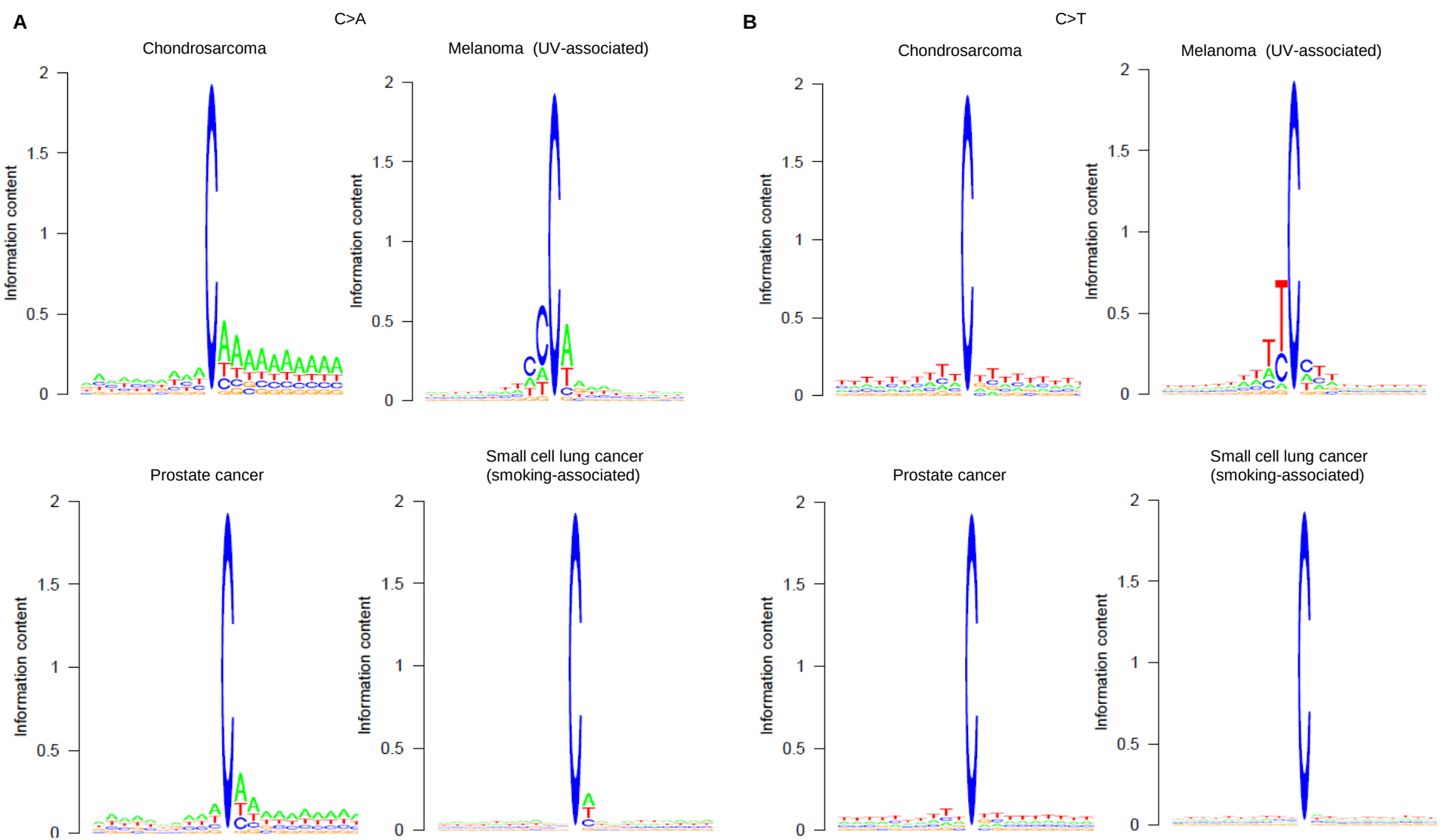

A
C>A
Chondrosarcoma
Melanoma (UV-associated)
Prostate cancer
Small cell lung cancer
(smoking-associated)
B
C>T
Chondrosarcoma
Melanoma (UV-associated)
Prostate cancer
Small cell lung cancer
(smoking-associated)
Information content
0
0.5
1
1.5
2

Figure 3

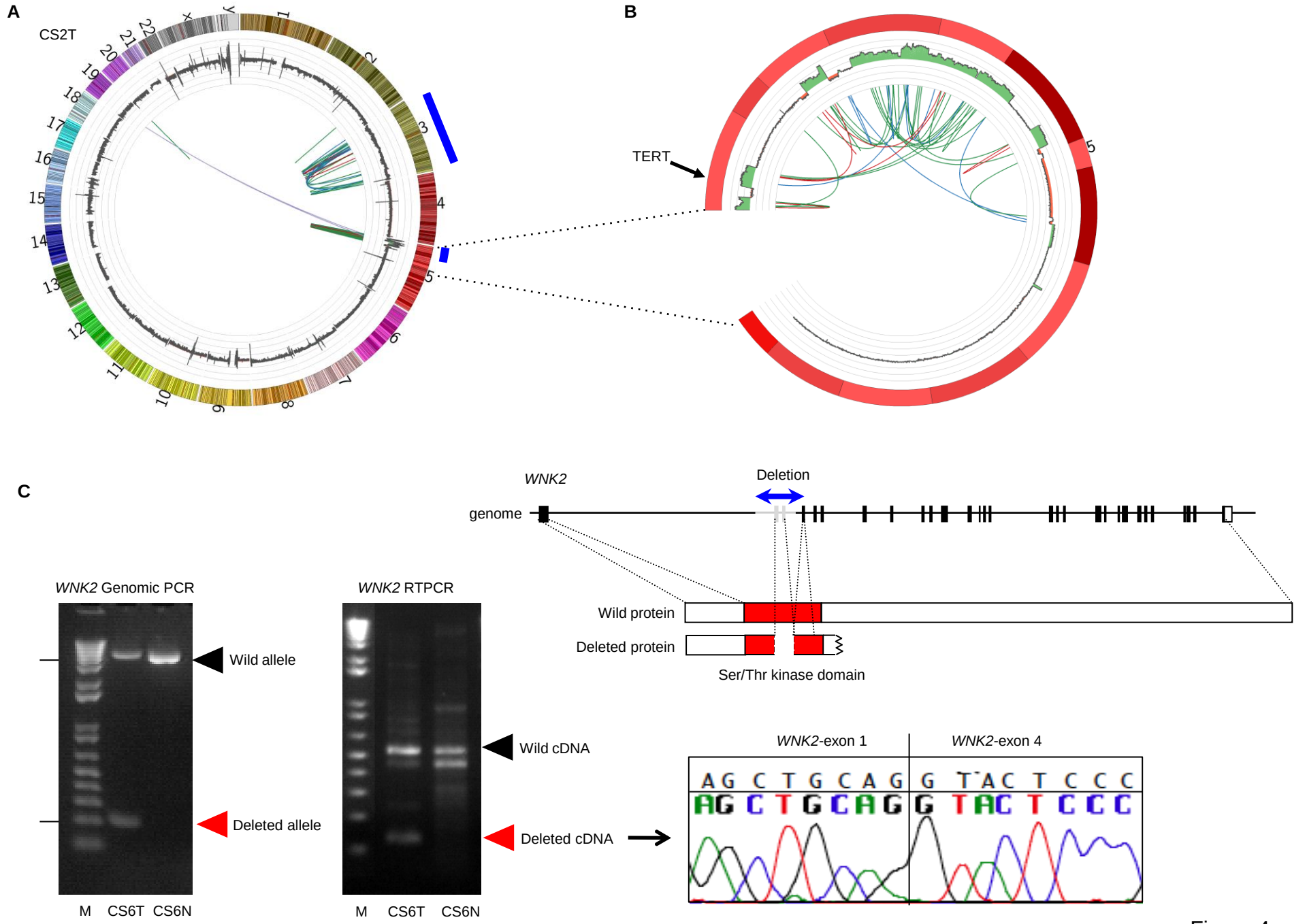

A
CS2T
B
TERT
C
WNK2
Deletion
genome
Wild protein
Deleted protein
Ser/Thr kinase domain
WNK2 Genomic PCR
Wild allele
Deleted allele
M CS6T CS6N
WNK2 RTPCR
Wild cDNA
Deleted cDNA
M CS6T CS6N
WNK2-exon 1
WNK2-exon 4
AG C T G C A G G T AC T C C C

Figure 4

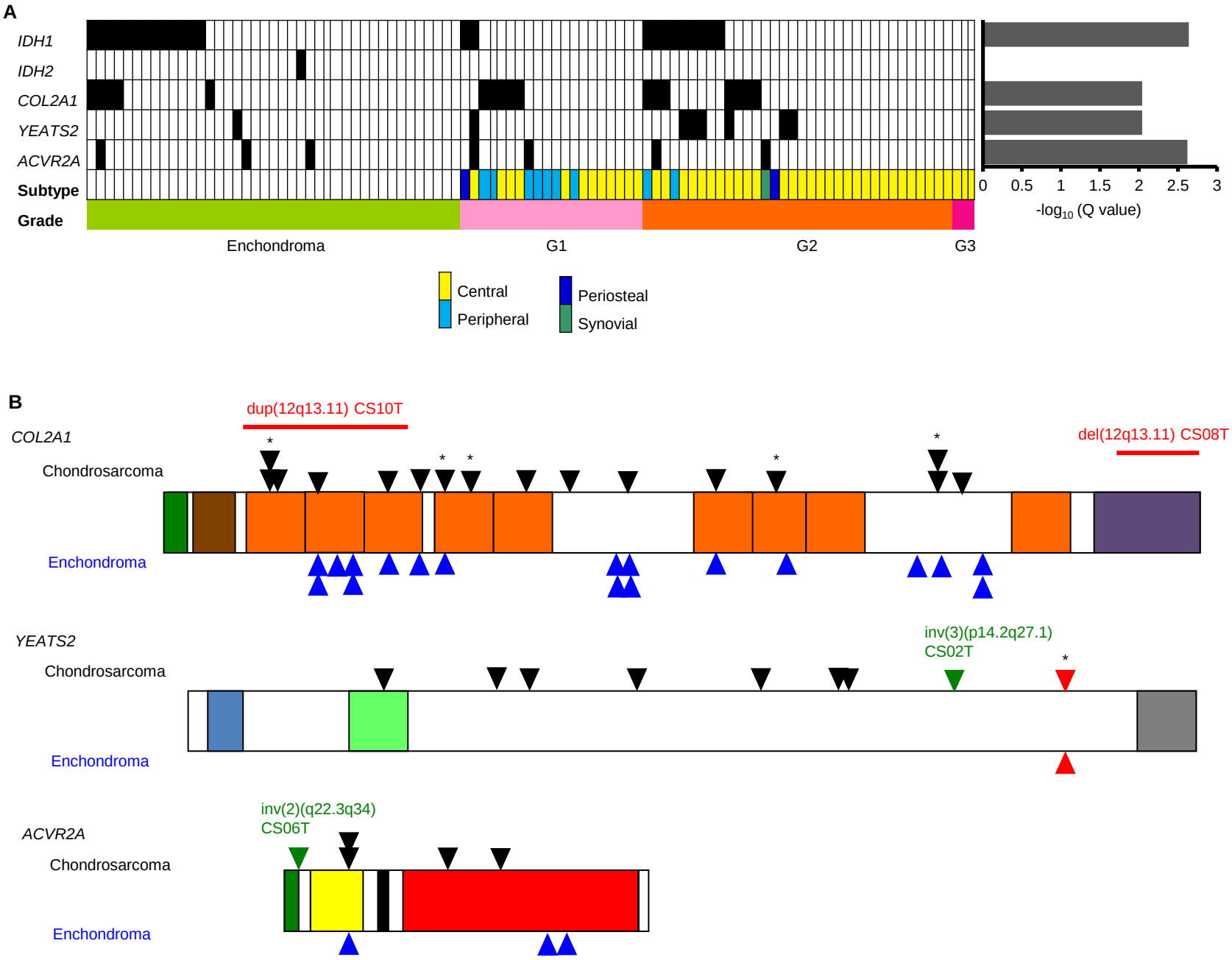

A
IDH1
IDH2
COL2A1
YEATS2
ACVR2A
Subtype
Grade
Enchondroma
G1
G2
G3
0
0.5
1
1.5
2
2.5
3
-log10 (Q value)
Central
Peripheral
Periosteal
Synovial
B
dup(12q13.11) CS10T
del(12q13.11) CS08T
COL2A1
Chondrosarcoma
Enchondroma
YEATS2
inv(3)(p14.2q27.1)
CS02T
Chondrosarcoma
Enchondroma
inv(2)(q22.3q34)
CS06T
ACVR2A
Chondrosarcoma
Enchondroma

Figure 5

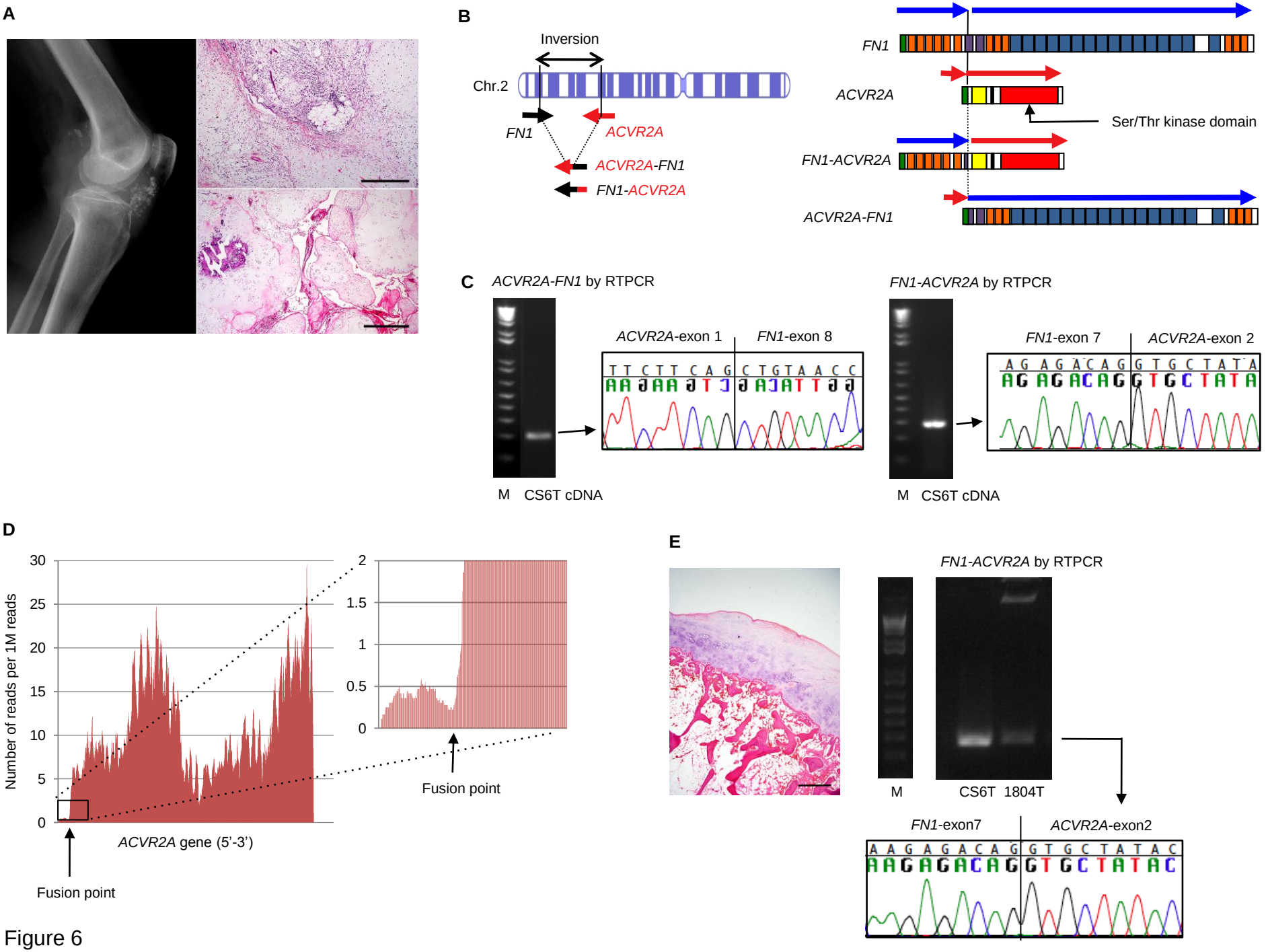

A
B
Inversion
Chr.2
FN1
ACVR2A
ACVR2A-FN1
FN1-ACVR2A
FN1
ACVR2A
Ser/Thr kinase domain
FN1-ACVR2A
ACVR2A-FN1
C
ACVR2A-FN1 by RTPCR
ACVR2A-exon 1
FN1-exon 8
M
CS6T cDNA
FN1-ACVR2A by RTPCR
FN1-exon 7
ACVR2A-exon 2
M
CS6T cDNA
D
Number of reads per 1M reads
ACVR2A gene (5'-3')
Fusion point
Fusion point
E
FN1-ACVR2A by RTPCR
M
CS6T 1804T
FN1-exon7
ACVR2A-exon2

Figure 6