CANCER GENETICS AND EPIGENETICS

The landscape of genetic aberrations in myxofibrosarcoma

Yasuhide Takeuchi1,2,3,4 | Kenichi Yoshida1 | Adriane Halik5 | Annegret Kunitz5
Hiromichi Suzuki1 | Nobuyuki Kakiuchi1,2 | Yusuke Shiozawa1
Akira Yokoyama1,6 | Yoshikage Inoue1,2 | Tomonori Hirano1,2
Tetsuichi Yoshizato1 | Kosuke Aoki1 | Yoichi Fujii1,2 | Yusuhito Nannya1,2
Hideki Makishima1,2 | Berit Maria Pfitzner7 | Lars Bullinger5,8,9
Masahiro Hirata3 | Keita Jinnouchi3 | Yuichi Shiraiishi10 | Kenichi Chiba10
Hiroko Tanaka11 | Satoru Miyano11 | Takeshi Okamoto12 | Hironori Haga3 | Seishi Ogawa1,2,13 | Frederik Damm5,8,9

1Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto
2Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Kyoto
3Department of Diagnostic Pathology, Kyoto University Hospital, Kyoto, Japan
4Japan Society for the Promotion of Science for Young Scientists, Tokyo, Japan
5Department of Hematology, Oncology, and Cancer Immunology, Charité-Universitätsmedizin Berlin, Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health, Berlin, Germany
6Department of Therapeutic Oncology, Graduate School of Medicine, Kyoto University, Kyoto, Japan
7Charité-Universitätsmedizin Berlin, Institut für Pathologie, Berlin, Germany
8German Cancer Consortium (DKTK), Berlin, Germany
9German Cancer Research Center (DKFZ), Heidelberg, Germany
10Center for Cancer Genomic and Advanced Therapeutics, National Cancer Center, Tokyo, Japan
11M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan
12Department of Orthopaedic Surgery, Kyoto University Hospital, Kyoto, Japan
13Department of Medicine, Centre for Haematology and Regenerative Medicine, Karolinska Institute, Stockholm

Correspondence
Seishi Ogawa, Pathology and Tumor Biology, Kyoto University, Yoshida Konoe cho, Sakyo ku, 606 8501, Kyoto.
Email: sogawa-tky@umin.ac.jp
Frederik Damm, Department of Hematology, Oncology, and Cancer Immunology, Charité-Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany.
Email: frederik.damm@charite.de

Funding information
This work was supported by Grant in Aid for

Abstract
Myxofibrosarcoma (MFS) is a rare subtype of sarcoma, whose genetic basis is poorly understood. We analyzed 69 MFS cases using whole-genome (WGS), whole-exome (WES) and/or targeted-sequencing (TS). Newly sequenced genomic data were combined with additional deposited 116 MFS samples. WGS identified a high number of structural variations (SVs) per tumor most frequently affecting the TP53 and RB1 loci. 40% of tumors showed a BRCAness-associated mutation signature, and evidence of chromothripsis was found in all cases. Most frequently mutated/copy number altered

Abbreviations: CNAs, copy number alterations; DNA, deoxyribonucleic acid; DSB, double-strand break; FFPE, formalin-fixed paraffin-embedded; IARC, International Agency for Research on Cancer; IHC, immunohistochemical staining; LMS, leiomyosarcoma; MFS, myxofibrosarcoma; PARP, polyadenosine diphosphate-ribose polymerase; SNVs, single nucleotide variations; STS, soft tissue sarcoma; SV, structural variation; TLs, telomere lengths; TS, targeted-sequencing; UPS, undifferentiated pleomorphic sarcoma; WES, whole-exome sequencing; WGS, whole-genome sequencing.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. International Journal of Cancer published by John Wiley & Sons Ltd on behalf of UICC.

INTRODUCTION

Soft tissue sarcomas (STS) are a diverse group of tumors with remarkable histologic diversity leading to more than 50 recognized subtypes. Identification of subtype-specific translocations, including SS18-SSX in synovial sarcoma, FUS-DDIT3 in myxoid/round cell liposarcoma and BCOR-CCNB3 in Ewing-like and undifferentiated sarcomas, has revolutionized the diagnostics of sarcoma and has provided new insight into oncogenesis. In addition, discovery of activating mutations in KIT or PDGFR in gastrointestinal stromal tumors led to routine application of tyrosine kinase inhibitors for these sarcomas, highlighting the great value of genetics for both diagnostics and targeted treatment approaches. Another group of sarcomas are characterized not only by a recurring, tumor-specific genetic alteration, but also by complex karyotypes that are characteristic of severe genetic and chromosomal instability. Most common among this latter group of STS is myxofibrosarcoma (MFS), which typically occurs in the lower extremities.

In the current study, fresh-frozen tumor and normal tissues/ peripheral blood were provided by Charité Universitätsmedizin Berlin (17 cases), and formalin-fixed paraffin-embedded (FFPE) tumor and normal tissues were provided by Charité Universitätsmedizin Berlin (25 cases) and Kyoto University Hospital (27 cases). Clinical information is summarized in Table S1. The diagnosis of 69 cases with MFS from current study was confirmed by pathological review according to WHO and Mentzel et al.

Here, we conducted a large-scale integrated genetic study of 185 MFSs, including 69 cases from current study and 116 cases from external data sets and identified recurrent driver genes, remarkable complex structural variations and high intratumor heterogeneity in MFS.
the tumors. In two cases with multiregional sampling, tumor specimens were annotated as myxoid or solid based on the pathological features that myxoid areas harbored extracellular myxoid matrix with more than 50% of the area, meanwhile solid areas harbored less than 10%. All tumor samples in the discovery set \( n = 44 \) with paired nontumor DNA were from primary tumors. Among the 17 cases with fresh frozen tumor tissue in the discovery cohort, 5 cases with abundant tumor/normal DNA were further analyzed by WGS. Additional 26 cases without matched nontumor DNA were enrolled to our study, among which one case (UPN056) was later removed because the tumor located in abdominal area, and copy number analysis revealed high amplification in both MDM2 and CDK4, both of which are typical findings of liposarcoma.\(^{11}\) A total of 69 samples including 44 samples from the discovery cohort and 25 additional samples were analyzed as validation cohort. The patient demographics and tumor characteristics of all cases in our study \( n = 69 \) are summarized in Table S1. On 53 tumor samples with available FFPE tissue, immunohistochemical staining (IHC) for TP53 (DO-7, DAKO) was performed. A positive case was indicated by an H score of 50 to 300 and a negative case by an H score of 0 to 49.\(^{12}\) Macrodissection of thin-sliced unstained tissues on glass slides was performed on FFPE tumor to ensure sample tumor content. DNA was extracted from FFPE tissue, fresh frozen tissue or fresh peripheral blood using a DNeasy Tissue Kit, a Gentra Puregene Blood Kit or a GeneRead DNA FFPE Kit (all from QIAGEN), according to the manufacturer’s protocols, as previously described.\(^{13,14}\) DNA was quantified using a Qubit Fluorometer (Life Technologies, Carlsbad, California). Sequence data from 17 MFSs from the TCGA dataset\(^{7}\) and 99 MFSs from a previous report\(^{7}\) were enrolled to our study.

### 2.2 WES, targeted deep sequencing, WGS and confirmatory deep sequencing

For WES, paired tumor (fresh frozen and FFPE tissues) and germline (tumor-free tissue or peripheral blood) samples were collected from all patients. WES libraries were prepared from 100 to 500 ng of genomic DNA and generated using a SureSelect Human All Exon V5 kit/V5 + UTR kit (Agilent Technologies), followed by massively parallel sequencing of enriched exon fragments on a HiSeq 2500 with 125-base pair paired-end mode, as previously described.\(^{13}\) Following the manufacturers’ protocols, libraries from FFPE samples were prepared using a KAPA Hyper Prep Kit (Kapa Biosystems) or a SureSelect Target Enrichment System (Agilent Technologies). The mean coverage of WES was 100.3×, and in average 90.6% of coding regions were analyzed at a depth of more than 20×.

For targeted deep sequencing (TS), tumor (fresh frozen and FFPE tissues) and germline (tumor-free tissue or peripheral blood) samples were collected. Libraries were prepared from 100 to 500 ng of genomic DNA and generated using a SureSelect custom kit (Agilent Technologies), followed by massively parallel sequencing of enriched exon fragments on a HiSeq 2500 with 125-base pair paired-end mode, as previously described.\(^{13,15}\) Following the manufacturers’ protocols, libraries were prepared using a SureSelect Low Input Automated Target Enrichment System (Agilent Technologies). The mean sequencing coverage across the targeted loci was 277.0×, with 96.8% achieving coverage >20×. For WGS, libraries were prepared from 300 ng of genomic DNA and generated using the TruSeq Nano DNA Library Preparation kit (Illumina) according to the manufacturers’ instructions. The mean coverage of WGS was 52.9× in tumor samples and 36.1× in control samples. WES and WGS sequencing statistics are summarized in Table S2. Mutations were called using the in-house pipeline Genomon2 (version 2.5.2) and EBCall, as previously described.\(^{13}\) Validation of mutations identified by WES were performed by amplicon-based deep sequencing and targeted deep sequencing, as previously described.\(^{16}\) The overall validation rate was 92.1% for fresh samples (35/38 mutations) and 91.4% for FFPE samples (127/139 mutations). Validation of mutations identified by WGS were performed by amplicon-based deep sequencing, TS and WES, as previously described.\(^{16}\) The overall validation rate was 95.4% (267/280 mutations).

Detection of structural variations was performed by Genomon2. Briefly, Genomon2 uses information from chimeric reads (containing breakpoints) and discordant read pairs. For each candidate structural variation, it realigns reads to the assembled contig sequence containing the structural variation breakpoint (variant sequence). Fisher’s exact test was performed to compare the proportion of read pairs aligned to variant sequences relative to reference sequences in tumor vs matched normal samples. Putative structural variations were manually curated and further filtered by removing those with (a) Fisher’s exact P-value >.01; (b) <4 supporting reads in tumor samples; (c) a VAF in tumor samples <0.1; (d) a VAF in matched normal samples ≥0.02 or (e) length between breakpoints <1000 bp. Identified structural variations were confirmed by PCR and/or Sanger sequencing, in which 24 of 26 (92.3%) variations were validated.

Significantly mutated genes or driver genes, were investigated by analyzing WES data based on MutSigCV (https://www.genepattern.org/modules/docs/MutSigCV) and dNdSCV (https://github.com/im3Sanger/dndscv) with their default setting.

Copy number alterations (CNAs) were evaluated from WES/targeted deep sequencing data using our in-house pipeline, “CNACS.”\(^{16}\) GISTIC 2.0 (https://www.genepattern.org/modules/docs/GISTIC_2.0) was used to calculate significant CN gains and losses. Amplification and deletions were annotated if copy number of gene of interest is 2 or ~2 in the all_thresholded.by_genes.txt file obtained from GISTIC 2.0 analysis, respectively. CNAs in WGS were analyzed by Patchwork (https://patchwork.r-forge.r-project.org/). Mutations, SVs and CNAs detected in WGS were integrated into circos plot (http://circos.ca/). Analysis on chromothripsis were done by svtools (https://github.com/PapenfussLab/sv_tools), ShatterSeek (https://github.com/parklab/ShatterSeek) and CTLPScanner (http://callab.labshare.cn/CTLPScanner/about.php). On the analysis of multi-regional/time-point samples, phylogenetic trees were illustrated by MEGAX (https://www.megasoftware.net/). Clonal compositions were analyzed using PyClone (v.0.13.0) (https://github.com/Roth-Lab/pyclone) as previously described.\(^{16}\)
Mutational signatures are analyzed by using whole genome sequencing data. We determined the contribution of known mutational processes reported by Alexandrov et al.\textsuperscript{17} using the SigProfiler (https://github.com/AlexandrovLab/SigProfilerExtractor). Kataegis is defined by the significant accumulation of the number of mutations, of which adjusted P-value of the frequency of the mutations in 10 kb per 10 Mb calculated based on the Poisson distribution are $10^{-6}$ or less. Mutational signatures in the kataegis, defined by $q$-value is $10^{-6}$ or less, were calculated by R package deconstructSigs (https://github.com/raerose01/deconstructSigs).

Telomere lengths (TLs) were calculated using TelSeq software (https://github.com/zd1/telseq) from WES data, which was validated for the use on the WES data in the original article. The relative TL ratio was defined as tumor telomere length/normal telomere length ($tTL/nTL$), that is, the $tTL$ divided by the $nTL$, corresponding to the pair-matched TL ratio information.
The methods used which 11 and 41 from a previ-
ously known as driver genes in MFS
or positively selected in MFS were investigated using MutSigCV and
dNdSCV with a
5.2.3 Computational analysis
Statistical analyses were performed using R (http://www.R-project.org). Multiple significance testing was adjusted according to the method described by Hochberg and Benjamini. The methods used for the statistical analyses are described in detail in Section 3.

3 Results
3.1 Whole-genome sequencing identified a median of 495 structural variations
Five MFS cases were evaluated by whole-genome sequencing (WGS) (Figure 1). All tumors were obtained from the primary site with a mean of 52.9× sequencing depth, by which 96.4% of coding regions were analyzed with a depth >20× on average. A total of 26 716 somatic mutations, including 24 974 single nucleotide variations (SNVs) and 1742 insertion-deletions (indels), were detected with a median of 5149 (1.6/Mb) SNVs and 328 (0.1/Mb) indels per case. According to an NMF-based decomposition, “clock-like” signatures, namely SBS1, 5 and 40, were predominant, explaining 57.9% of all SNVs. In two cases (ID UPN015 and UPN017; Figure 2), a substantial fraction of SNVs were assigned to SBS3, which is implicated in defective homologous recombination-based DNA damage repair, typically caused by defective BRCA1 or BRCA2. While none of these two cases carried pathogenic somatic/germline variants in BRCA1 or BRCA2, we identified bi-allelic RAD51D deletion (UPN015) and CHEK1-involving structural variations (SVs) (UPN017), which might have contributed to increased SBS3 mutations in these two cases (Figure 1). APOBEC-related signatures (SBS2/13) were detected in three cases (Figure 2).

In total, we identified 2298 somatic SVs (median: 495 per case), consisting of 434 deletions, 662 inversions, 284 tandem duplications and 918 translocations (Figure 1). Recurrent SVs most frequently affected the TP53 (n = 4) and RB1 loci (n = 3) (Figure S2A). In particular, SVs with and without concomitant SNV accounted for the biallelic alteration of TP53 in three cases carrying intrachromosomal (n = 2) and/or interchromosomal (n = 1) deletions, while it was explained by a combination of a SNV plus an indel in the remaining two cases. Notably, chromothripsis was present in all five samples, with the number of affected chromosomes per tumor ranging from 5 to 10 with a median of 7 (Figure S2B), where chromosomes 2 and 17 were most commonly affected (n = 4, respectively), followed by chromosomes 3, 7 and 13 (n = 3, respectively). Chromothripsis was often accompanied with local hypermutation, so called “kataegis” (Figure 1). Indeed, mutation signature analysis of mutations within the kataegis showed a predominance of APOBEC-related signatures 13 and 2 (38% and 10.0%, respectively), pointing to central role of APOBEC-mediated local hypermutation within the chromothripsis regions.

3.2 The landscape of genetic alterations in MFS analyzed by WES
Next, we conducted WES of 44 paired tumor and normal samples. All tumors were obtained from the primary site with a mean of 100.3× sequencing depth, by which 90.6% of coding regions were analyzed with a depth >20× on average. In addition, by integrating WES data from 17 MFS samples from the TCGA database and 41 from a previous report, we investigated a total of 102 MFS samples. We detected a total of 7330 somatic mutations with a mean of 50 mutations per case, consisting of 2027 synonymous, 4476 missense, 286 nonsense, 2 stop-loss and 113 splice site SNVs as well as 356 frameshift indels and 70 in-frame indels (Table S3). Driver genes significantly mutated or positively selected in MFS were investigated using MutSigCV and dNdSCV with a P-value <.01 (Tables S4 and S5). Besides genes previously known as driver genes in MFS such as TP53 (25/102, 24.5%), RB1 (23/102, 22.5%), ATRX (15/102, 14.7%), NFI (9/102, 8.9%) and TET2 (5/102, 4.9%), we also identified several novel candidates for driver genes including MUC17 (25/102, 24.5%), COL6A3 (6/102, 5.9%), NLRP4 (6/102, 5.9%), SLC12A5 (6/102, 5.9%), VCAN (6/102, 5.9%), WDR87 (6/102, 5.9%) and ZNF680 (5/102, 4.9%), each being
mutated in at least five cases and with P-value <.01 in dNdSCV or MutSigCV.

Next, WES data were investigated for CNAs. Using GISTIC 2.0, we identified significant focal amplifications (q-value <0.01) and copy number losses (homozygous and heterozygous deletions) (q-value <0.25) at 35 loci (13 amplifications and 22 copy number losses), including known oncogenes or oncogenic lesions [DRG2, COPS3, YAP1, JUN, AXL, SETDB1, TFDP1, 1q21.3 and PRKD1].

intron regions of driver genes (such as 3A oncogenic function of tumour suppressors (TP53, CDKN2A, CDKN2B, RB1, HDLBP, FANCA, AKT3, SENP6, FH, BRD7, SNTG2, C16orf78 and MYT1L)] (Figure 3A). Focal amplifications of CD274 encoding PD-L1 were identified in four cases (3.9%).

Combining SNV and CNA data, we identified genes recurrently affected by mutations and CNAs that have not been reported previously in MFS. For example, FLG were recurrently affected in 24 cases (23.5%) by SNVs (n = 12) (11.8%, P-value and q-value by dNdSCV were .036 and 1, respectively) and amplifications (n = 14) (13.7%, q-value by GISTIC 2.0 was >0.2). FLG encodes filaggrin, an intermediate filament-associated protein that aggregates keratin intermediate filaments in mammalian epidermis. While germline mutations leading to loss of FLG function are known to be strong genetic risk factors for allergic diseases, such as atopic dermatitis, oncogenic function of filaggrin has not yet been previously reported. Alterations in ZNF780A were also recurrently detected (23 cases, 22.5%), including mutations in nine cases (8.8%, P-value and q-value by dNdSCV were .17 and 1, respectively), amplifications in 11 patients (10.8%, q-value by GISTIC 2.0 was 1.7 × 10^-7) and tandem duplications in four tumors (3.9%).

3.3 Detailed analysis on genetic alterations in MFS (n = 185) identified mutually exclusive alterations in CDKN2A/B and HDLBP

Next, to confirm the findings on WES and further interrogate additional driver alterations involved in MFS pathogenesis, we analyzed 69 MFS cases, including the 44 cases from our WES discovery cohort, by TS of coding and promoter regions of 173 selected genes, including those identified by WES and those identified by a review of the relevant literature. Intronic regions of driver genes (such as TP53, RB1 and ATRX) and a copy number backbone were also integrated to sensitively detect SVs and CNAs. The mean sequencing depth across the targeted loci was 277×, which cover 96.8% of the target regions with >20× depth. Using this approach, we identified 549 nonsilent somatic mutations in 141 genes with a median of six SNVs per tumor. In 43 cases analyzed by both WES and TS, TS detected an additional 141 mutations, validated 80.2% (89/111) of the mutations detected by WES (Table S6). Combining 17 MFS cases from TCGA database and 99 from previous report, we analyzed a total of 185 MFS cases. In accordance with the results from WES, FLG and ZNF780A were shown to be recurrently mutated (22.8% and 20.5%, respectively) (Figure 3B).

The incidence and distribution of driver mutations and CNAs are presented in Figure 4. Although infrequent, actionable hotspot mutations were identified in KRAS, BRAF and PIK3CA (Figure S3).

For more sensitive detection of TP53 mutations, 52 tumor samples were further analyzed by IHC for TP53 protein expression using FFPE samples. Immunostaining was negative in 30 cases, the majority of which harbored nonsense/frameshift mutations (n = 7), biallelic deletions (n = 8) or intra/intercromosomal SVs (n = 2). By contrast, 22 cases (42.3%) showed strong staining of TP53 in IHC suggestive of expression of mutant TP53. In fact, 10 (45%) of 22 had somatic TP53 mutations (Figures 3B and 4), including missense mutations (n = 9) cases and one frameshift deletion (NM_000546: p.S314fs), for which the stabilized TP53 is expected on the basis of the registered data in the International Agency for Research on Cancer (IARC) TP53 Mutation Database (http://p53.iarc.fr/). In wild-type state, p53 protein is rapidly degraded by ubiquitination, which results in negative/heterogeneous staining on IHC. No genomic alterations in TP53 were found in 12 cases despite strong TP53 immunostaining in IHC, in which therefore activated, rather than inactivated, TP53 functions are implicated. In general, TP53 alterations were frequently biallelic events (89/96, 92.7%) most often resulting from a combination of a SNV plus a CNA (n = 54), followed by biallelic deletions (n = 33). As expected, TP53 alterations were almost mutually exclusive with MDM2 amplifications (P-value = .020, Fisher’s exact test).

We also identified recurrent mutations in regulators of the telomeres in 51 cases (27.6%), which included ATRX, SP100, DAXX and RBL2 (Figure S4). ATRX cooperates functionally with DAXX, whose mutations are associated with altered telomeres. While genetic DAXX alterations have not yet been reported in MFSs, we identified two cases with mutations and seven cases with copy number alterations. Deletion of RBL2 showed a significant positive correlation (P-value: .016, by unequal variances t-test) with relative telomere length, which was evaluated by the ratio of tumor telomere length to normal telomere length and calculated by TelSeq (Figure S5). RBL2, which is also known as Rb-related p130 protein, is reported to suppress telomerase-independent telomere lengthening and, in LMS, deletions of RBL2 were reported to be correlated with alternative shortening of telomeres (ALT) activity. These data indicate that frequent genetic alterations in genes of telomere maintenance may accelerate the chromosomal instability in MFS.

Genetic alterations in HDLBP (24/127 cases, 18.9%, 22 biallelic deletions and 2 missense mutations) were mutually exclusive (P-value = .007, Fisher’s exact test) with biallelic deletions or loss of function mutations in CDKN2A/B (39/127 case, 30.7%). HDLBP encodes Vigilin, which is among the largest RNA-binding proteins known to date and has been implicated in the induction of double-stranded RNA heterochromatin formation. It represents an emerging cancer pathway, particularly in a mouse model of myxofibrosarcoma. Comprehensive genomic analysis of adult soft tissue sarcoma (n = 206, The Cancer Genome Atlas Research Network) also showed this pattern of mutual exclusivity between biallelic deletion of CDKN2A/B and HDLBP, suggesting functional redundancy of CDKN2A and HDLBP or synthetic lethality in between sarcomagenesis.
FIGURE 3  Significant copy number alterations, distribution and mutational profile of TP53, FLG and ZNF780 genes. (A) Recurrent focal gains (left) and losses (right) analyzed by GISTIC2.0 in WES of 102 cases were plotted. The dashed line indicates the cut-off for significance ($q = 0.25$). For each $q$-value peak, putative gene targets are listed. (B) Lollipop plot showing the distribution and mutation profile in TP53, FLG and ZNF780. The types and status of the somatic alterations are shown by the indicated colors. On the vertical axis the frequency of appearance of each mutation is shown. On the horizontal axis the amino acid position of each mutation is shown [Color figure can be viewed at wileyonlinelibrary.com]
Frequent deletions of genes implicated in DNA double-strand break (DSB) repair have recently been reported for LMS.

In our cohort, alterations in these DSB repair genes, including *ATR*, *FANCA*, *BRCA1* and *BRCA2*, were found in 59 of 185 cases (31.9%), which were less frequent than reported for LMS (73.5%) (Figure S6). Amplification in *VGLL3* and *MDM2* were significantly enriched in the samples without alterations in DSB repair genes (P-value: .047 and .042, respectively, Fisher’s exact test). Alterations in *TP53*, *RB1*, *HDLBP*, *FANCA* and *PTEN*, and strong staining in *TP53* IHC were significantly enriched in the samples with alterations in DSB repair genes (P-value: 3.9 x 10^{-3}, 2.4 x 10^{-2}, 7.9 x 10^{-4}, 4.7 x 10^{-2}, 6.9 x 10^{-3} and 4.3 x 10^{-7}, respectively, Fisher’s exact test).

![Figure 4](wileyonlinelibrary.com)

**FIGURE 4** Landscape of somatic genetic alterations in MFS. Based on the combined results from our study and previous studies, the spectrum of somatic mutations, somatic structural variations, copy number alterations and staining in IHC in MFS are shown. Each column represents an individual case, and each row represents the indicated gene with alteration. The types and status of the somatic alterations are shown by the indicated colors. The frequency of each genetic alteration is shown on the right. IHC, immunohistochemistry; InDel, insertion-deletion. [Color figure can be viewed at wileyonlinelibrary.com]
Survival analysis was performed in 63 cases having sufficient clinical data. During a median of 3.2 years of follow-up period, only four patients died, indicating a good prognosis for patients having a resectable, early-stage tumor. Only alterations in RB1 were marginally associated with poor prognosis ($P$-value = 0.046). No other genetic alterations or clinical features (surgical margin, histological grade) were significantly associated with patients’ prognosis (Figure S7). In addition, we did not identify any significant association between genetic alterations and clinical features (age, gender, tumor location) or tumor characteristics (tumor depth, grade, stage).

3.4 | Whole-exome sequencing with multiregional and/or multi-time-point sampling proved increasing genetic alterations with clinical course of MFS

To elucidate the intratumor heterogeneity and the clonal evolution in MFSs, we performed WES from samples obtained from multiple tumor sites at different time points (Figures 5, S8 and Data S1). For this purpose, we enrolled six patients from whom 16 samples, including seven diagnostic and nine relapse samples, were collected. Macrodissection of thin-sliced unstained FFPE tissues on glass slides was performed to ensure high tumor content. In total 530 (mean: 75.7, median: 35, range:
14–345) and 907 (mean: 100.8, median: 41, range: 19–622) mutations were identified in seven diagnostic and nine relapse samples, respectively. Analysis using a generalized linear mixed model revealed that the number of mutations was significantly larger at relapse (1.2 times on average, \( P\)-value = .01). This observation holds true, even when excluding those samples from a tumor (UPN044) showing significantly higher number of mutations than the other samples (Figure 5A). Including a patient receiving radiation therapy before disease recurrence (UPN023) (Figure S8), no significant changes in mutational signatures were observed between primary and relapsed tumors. All samples carried at least one alteration in TP53, CDKN2A/B or RB1, all of which were observed in more than 25% of diagnostic samples (Figure 1) and were observed in both diagnostic and relapse samples (Figures 5C and S8). By contrast, alterations in other driver genes were detected in either diagnostic (12/22) or relapse samples (10/22) alone (Figures 5C and S8), suggesting that less frequent driver events lead to substantial clonal diversification including spatial and temporal heterogeneity in MFS tissues.

In two cases multiregional sampling were performed. In each case, two tumor specimens with distinct histological components (myxoid and solid) were analyzed using WES (Figures 5B, C and S8). In each pair of samples, less than 50% of all SNVs and CNAs (SNVs: 20.8% and 45.5%, CNAs: 17.1% and 2.2%) were shared, while the majority of alterations were specific to each tumor site. In UPN021, both myxoid (site 1) and solid components (site 2) shared a biallelic deletion of CDKN2A/B, whereas the myxoid and the solid component harbored private lesions, such as a FAT1 amplification and amplifications of TRIO and COP53, respectively. In UPN025, both myxoid (site 1) and solid components (site 2) harbored an amplification of TEK. TEK encodes angiopoietin-1 receptor, also known as CD202B (cluster of differentiation 202B) or TIE2, which belongs to the Tie2 family of protein kinase. TEK amplification has been reported in breast carcinoma, lung adenocarcinoma, pancreatic adenocarcinoma and anaplastic astrocytoma. In our cohort, only four cases analyzed by WES (4.0%) were shown to harbor TEK amplifications. Although no plausible driver genetic alterations were found in myxoid component (site 1), only solid component (site 2) harbored biallelic deletions of CDKN2A/B. In both cases, the numbers of mutations detected in solid component (UPN021 site 2: 30 and UPN025 site 2: 42) were higher than those in myxoid component (UPN021 site 1: 14 and UPN025 site 1: 6) (Figure 5B). However, due to the small number of investigated cases, further studies are required to fully capture correlations between genetic alterations and the histological features.

Clonal evolution in MFSs was further analyzed by phylogenetic tree analysis (Figures 5C and S8). Among seven cases, five harbored somatic mutations or CNAs affecting TP53 or strong staining of p53 in IHC in the “trunk” of the tree, which means the alteration of the TP53 was the essential driver events in these tumors. Two cases without TP53 alterations harbored genetic alterations in other driver genes: one case with TEK amplification (Figure 5C, UPN025) and the other case with CDKN2A/B biallelic deletion (Figure 5C). Analysis by PyClone revealed that base substitutions in TP53 were clonal events in both cases, which support the notion that alterations in TP53 are fundamental events in the tumorigenesis of MFS. Supporting this notion, analysis of cell fraction with base substitution by PyClone or with CNA revealed that in primary MFS samples analyzed by WES (n = 102), almost all alterations in TP53 (58/60, 96.7%) were clonal events.

4 | DISCUSSION

Here we present a comprehensive genetic study, which to our knowledge included the largest number of MFS cases ever studied. We unravel genetic aberrations commonly observed in MFS and report a remarkable intra-/intertumor heterogeneity in MFS tumor tissue. Genetic aberrations affecting TP53 were the most frequent alterations and occurred in almost all cases as biallelic events, suggesting that biallelic TP53 alterations play a crucial role during MFS sarcomagenesis. While recurrent somatic SVs involving the first intron of TP53 were reported in osteosarcomas, our analysis did not identify similar breakpoints in MFS, although the intronic area of TP53 was captured by our custom baits. Frequent biallelic involvement of TP53 alterations resulted from different combinations of SVs, CNAs and SNVs. Based on the recent advances in the understanding of the molecular oncogenesis, many small molecules targeting specific oncogenic proteins, including mutant TP53 inhibitors, are under investigation in clinical trials. It will be of great interest to investigate whether small molecules that restore wild-type p53 functions in TP53-mutant cells such as Eprenetapopt (APR-246) and COTI-2 show a similar efficacy in MFS as reported for hematologic malignancies including acute myeloid leukemias.

Another potential target to restore normal p53 function might be an inhibition of overexpressed MDM2 by gene amplification. MDM2 is known to promote degradation of p53 through ubiquitination. MDM2 amplifications are common in adult soft tissue sarcoma in a mutually exclusive manner with TP53 mutations, which is a finding that we were able to confirm for MFS. Although amplification of MDM2 is, as previously described by Ogura et al, a typical marker of the dedifferentiated liposarcoma, it is not a diagnostic finding. In our cohort, all the cases with MDM2 amplifications arose at the peripheral sites, not the truncal area, and careful pathological review proved that all the cases in our cohort did not harbor liposarcoma-like components. Considering these findings, in concordance with the previous report, we diagnosed these cases as MFS with MDM2 amplification. Many small molecules targeting MDM2 are currently under evaluation in clinical trials. In addition to mutual exclusiveness with TP53 aberrations, we identified a significant enrichment of MDM2 amplifications in samples without alterations in genes involved in DSB repair. Through synthetic lethality, poly (adenosine diphosphate [ADP]-ribose) polymerase (PARP) has antitumor activity in cancer associated with BRCA associated gene alterations. Although identification of SBS-3 is known to be difficult on the analysis of mutational signatures, and SBS-3 was identified without concomitant Indel6/8 signatures in our cohort, validation and analysis in even larger MFS cohorts will help to assess the frequency of SBS3-mutations and its usefulness to predict sensitivity to drugs that target defective DSB
repair pathway. In line with this, defective DSB repair pathway as a potentially actionable feature was recently described in leiomyosarcoma (LMS), a sarcoma entity with many genetic commonalities to MFS. LMS and MFS harbor similar genetic alterations such as mutations in genes including TP53, RB1 and ATRX, and complex CNAs. However, it should be noted that methylation class prediction revealed substantial epigenetic differences between LMS and MFS, resulting in different class assignments with MFS cases clustering in the undifferentiated sarcoma class. Although mutational landscape and analysis on the tumor evolution revealed that biallelic inactivation of TP53/RB1 are frequent and fundamental event in both MFS (Figures 4 and 5) and LMS, in comparison with MFS genetics, LMS less frequently harbors alterations in CDKN2A (8% vs 30% in MFS) and NF1 (4% vs 22% in MFS), but more commonly affected DSB repair pathway (73.5% vs 32% in MFS). In MFS, various alterations in receptor tyrosine kinase (RTK)-RAS-PI3K cancer pathway (eg, KRAS, BRAF, PIK3CA and NF1) were more frequently detected (36.2% vs not described in LMS), which may result in inter- and intratumor heterogeneity of MFS but offer therapeutic windows.

Another disease entity which is known to harbor similar genetic alterations is UPS. Recent profiling of genetic alterations in UPS revealed frequent alterations in TP53, RB1, ATRX and CDKN2A in UPS (65%, 47%, 28% and 25%, respectively), which is similar to the frequency in MFS in our current study. Methylation class prediction cannot differentiate UPS and MFS, both of which was classified as undifferentiated sarcoma class. Considering these findings, until more detailed profiling of genetic/epigenetic/transcriptomic shows clear difference between MFS and UPS, classical histological survey for distinguishing MFS from UPS is mandatory.

In addition to known genetic alterations, we uncovered several recurrent driver genes, which have not previously been identified as driver genes in MFS. These not only included genes that have been found in other cancer and sarcoma entities such as HDLBP, but also other genes whose role in carcinogenesis is poorly understood (eg, ZNF780A). HDLBP encodes vigilin, an RNA-binding protein that has been implicated in the induction of heterochromatin, and was reported as candidate target for chromosomal 2q37.3 deletions in Sleeping Beauty transposons screens of sarcoma cell lines. So far, although overexpression of vigilin is reported to be associated with tumor growth of hepatocellular carcinoma, association between sarcomagenesis and vigilin is not yet reported. Detailed analysis of the impact of loss of vigilin and sarcomagenesis is required.

With respect to prognostic impact of genetic alterations in MFS, we were able to confirm an inferior overall survival for patients harboring RB1 alterations. However, in contrast to previous reports, we did not observe an association between survival and alterations in TP53 nor CDKN2A/B. It should be noted that substantial heterogeneity with respect to presurgical application of chemotherapy with or without hyperthermia and radiotherapy may account for these differences. Even larger cohorts, preferentially enrolled on prospective trials, are warranted to further establish prognostic significance of genetic lesions for individual risk-guidance in MFS.

Taken together, our data provide a comprehensive genetic atlas of MFS sarcomagenesis and suggest at least three avenues for precision medicine guided treatment approaches to further improve patient outcome.

AUTHOR CONTRIBUTIONS
Yasuhide Takeuchi, Kenichi Yoshida, Frederik Damm and Seishi Ogawa designed the study; Annegret Kunitz, Adriane Halik, Berit Maria Pfützner, Lars Bullinger, Takeshi Okamoto, Hironori Haga and Frederik Damm provided specimens and clinical information; Yasuhide Takeuchi and Hironori Haga performed histological analysis; Yasuhide Takeuchi and Annegret Kunitz performed sample preparation; Yasuhide Takeuchi, Kenichi Yoshida, Hiromichi Suzuki, Nobuyuki Kakiuchi, Yusuke Shiozawa, Akira Yokoyama, Yoshikage Inoue, Tomonori Hirano, Tetsuichi Yoshizato, Kosuke Aoki, Yoichi Fujii, Yasuhiro Nannya and Hideki Makishima performed sequencing, mutation calling, validated the results and analyzed CNAs and mutational signatures; Masahiro Hirata and Keita Jinnouchi performed immuno-histochemical staining; Yuichi Shiraishi, Kenichi Chiba, Hiroko Tanaka and Satoru Miyano performed bioinformatics analysis; Yasuhide Takeuchi, Kenichi Yoshida, Frederik Damm and Seishi Ogawa prepared the article. The work reported in the article has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS
We are grateful to the TCGA Research Network (cancergenome.nih.gov/) for providing the data analyzed in this article. We also thank Maki Nakamura, and Itsuko Koyanagi, for their technical assistance. The authors acknowledge the assistance of the Central Biobank Charité (ZeBanC) and thank the Center for Anatomical, Pathological and Forensic Medical Research, Kyoto University Graduate School of Medicine for preparing microscope slides. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST
Lars Bullinger reports honoraria from and/or advisory roles for Abbvie, Amgen, Astellas, Bristol-Myers Squibb, Celgene, Daichi Sankyo, Gilead, Hexal, Janssen, Jazz Pharmaceuticals, Menarini, Novartis, Pfizer, Sanofi, Seattle Genetics. Lars Bullinger received research funding from Bayer, Jazz Pharmaceuticals. All other authors declare no competing financial interests.

DATA AVAILABILITY STATEMENT
The data sets have been deposited in the European Genome-Phenome Archive (EGA) under the accession number WES: EGAS00001005442, WGS: EGAS00001005443, TS: EGAS00001005444. For legal reasons, access to the datasets is restricted. Note that the EGA provides secure access to restricted data for authorized researchers and clinicians. Further information is available from the corresponding author upon request.

ETHICS STATEMENT
All samples were collected with informed consent following approval of the institutional review boards of the respective institutions.

ORCID
Frederik Damm https://orcid.org/0000-0001-5553-1173
REFERENCES

1. WHO. Classification of Tumours Editorial B: Soft Tissue and Bone Tumours. Lyon, France: International Agency for Research on Cancer (IARC); 2020.

SUPPORTING INFORMATION
Additional supporting information may be found in the online version of the article at the publisher’s website.