Clinical sequencing of soft tissue and bone sarcomas delineates diverse genomic landscapes and potential therapeutic targets — **Source link** ⤢

Benjamin A. Nacev, Benjamin A. Nacev, Benjamin A. Nacev, Francisco Sanchez-Vega ...+63 more authors

Institutions: Memorial Sloan Kettering Cancer Center, Rockefeller University, Cornell University

Topics: Sarcoma

Related papers:

- Molecular testing of sarcomas

- Identification of prognostic gene signatures of glioblastoma: A study based on TCGA data analysis

- Molecular classification as prognostic factor and guide for treatment decision of pancreatic cancer.

- Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin

- Comprehensive genomic sequencing and the molecular profiles of clinically advanced breast cancer

Share this paper: 🅕 🐦 in ✉

View more about this paper here: https://typeset.io/papers/clinical-sequencing-of-soft-tissue-and-bone-sarcomas-1g7uxg60k4

Clinical sequencing of soft tissue and bone sarcomas delineates diverse genomic landscapes and potential therapeutic targets

Benjamin A. Nacev [1,2,3]*, Francisco Sanchez-Vega [4]*, Shaleigh A. Smith [4,5], Cristina R. Antonescu [6], Evan Rosenbaum [1,2], Hongyu Shi [7], Cerise Tang [6,8], Nicholas D. Socci [5,9], Satshil Rana [6], Rodrigo Gularte-Mérida [4], Ahmet Zehir [6], Mrinal M. Gounder [1,2], Timothy G. Bowler [1], Anisha Luthra [5,10], Bhumika Jadeja [4], Azusa Okada [4], Jonathan A. Strong [4], Jake Stoller [4], Jason E. Chan [1], Ping Chi [1,2,10], Sandra P. D'Angelo [1,2], Mark A. Dickson [1,2], Ciara M. Kelly [1,2], Mary Louise Keohan [1,2], Sujana Movva [1,2], Katherine Thornton [1,2], Paul A. Meyers [11], Leonard H. Wexler [11], Emily K. Slotkin [11], Julia L. Glade Bender [11], Neerav N. Shukla [11], Martee L. Hensley [1,2], John H. Healey [4], Michael P. La Quaglia [4,11,12], Kaled M. Alektiar [13], Aimee M. Crago [4,12], Sam S. Yoon [4,12], Brian R. Untch [4,12], Sarah Chiang [6], Narasimhan P. Agaram [6], Meera R. Hameed [6], Michael F. Berger [5,6,10], David B. Solit [1,2,5], Nikolaus Schultz [7,10], Marc Ladanyi [6,10], Samuel Singer [4,12‡], William D. Tap [1,2‡]

1. Department of Medicine, Memorial Sloan Kettering Cancer Center

2. Department of Medicine, Weill Cornell Medical College

3. The Laboratory of Chromatin Biology and Epigenetics, The Rockefeller University

4. Department of Surgery, Memorial Sloan Kettering Cancer Center

5. Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center

6. Department of Pathology, Memorial Sloan Kettering Cancer Center

7. Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center

8. Physiology, Biophysics and Systems Biology Graduate Program, Weill Cornell Medical College

9. Bioinformatics Core, Memorial Sloan Kettering Cancer Center

10. Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center

11. Department of Pediatrics, Memorial Sloan Kettering Cancer Center

12. Department of Surgery, Weill Cornell Medical College

13. Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center

* Equal contribution

‡ Co-corresponding authors

Running title: Genomic landscape of sarcomas

Keywords: mesenchymal neoplasm, targeted sequencing, sarcoma subtypes, genetic alteration, genetic variability

Corresponding authors:

William D. Tap, MD
1275 York Ave.
New York, NY 10065
646-888-4163
tapw@mskcc.org

Samuel Singer, MD
1275 York Ave.
New York, NY 10065
212-639-2164
singers@mskcc.org

Number of figures and tables: 7 main figures, 4 supplementary figures, 5 supplementary tables

1 **Abstract**

2 The genetic, biologic, and clinical heterogeneity of sarcomas poses a challenge for the

3 identification of therapeutic targets, clinical research, and advancing patient care. Because there

4 are > 100 sarcoma subtypes, in-depth genetic studies have focused on one or a few subtypes.

5 Herein, we report a comparative genetic analyses analysis of 2,138 sarcomas representing 45

6 pathological entities. This cohort was prospectively analyzed using targeted sequencing to

7 characterize subtype-specific somatic alterations in targetable pathways, rates of whole genome

8 doubling, mutational signatures, and subtype-agnostic genomic clusters. The most common

9 alterations were in cell cycle control and *TP53*, receptor tyrosine kinases/PI3K/RAS, and

10 epigenetic regulators. Subtype-specific associations included *TERT* amplification in intimal

11 sarcoma and SWI/SNF alterations in uterine adenosarcoma. Tumor mutational burden, while

12 low compared to other cancers, varied between and within subtypes. This resource will improve

13 sarcoma models, motivate studies of subtype-specific alterations, and inform investigations of

14 genetic factors and their correlations with treatment response.

Introduction

16 Sarcomas are mesenchymal malignancies of the bone or soft tissue that arise in diverse organ

17 sites and display a range of clinical behavior from indolent to aggressive. Sarcomas are also

18 rare tumors, representing < 1% of all malignancies in adults (1). Although the diagnosis and

19 management of sarcomas has slowly improved over the last decade, about 40% of patients with

20 newly diagnosed sarcoma eventually die of disease. One barrier to improving outcomes in

21 sarcoma patients is the cancer's genomic and biologic complexity, with more than 100 different

22 subtypes now recognized by the World Health Organization (2).

23 Advances in clinical tumor genomic analyses have improved tumor classification; sarcomas are

24 now classified into two broad genetic groups (3). Sarcomas often have either simple karyotypes,

25 harboring genetic translocations or activating mutations, or highly complex karyotypes, including

26 numerous genomic rearrangements and large chromosomal gains and losses, commonly

27 involving cell cycle genes such as *TP53*, *MDM2*, *RB1*, and *CDK4*. Toward identifying

28 therapeutic targets and designing precision oncology trials based on specific sarcomas' genetic

29 features, a comprehensive study of soft tissue sarcomas was performed by The Cancer

30 Genome Atlas network, which analyzed 206 samples within 7 common subtypes; rarer ones

31 were represented by as few as 5 cases (4). Analysis of a larger cohort could define the

32 frequency of potentially actionable alterations in rare sarcoma subtypes and broadly compare

33 the frequency of genetic alterations across subtypes to facilitate better diagnostic precision,

34 identify prognostic biomarkers, improve laboratory-based modeling of sarcomas, and generate

35 novel hypotheses on underlying disease mechanisms.

36 Here, we leveraged an institution-wide tumor genomic profiling initiative to prospectively analyze

37 2,138 sarcomas encompassing 45 subtypes to identify subtype-specific somatic mutations and

38 copy number alterations and characterize tumor mutation burden (TMB) and microsatellite

39   instability. Paired tumor and normal DNA samples were analyzed using the FDA-cleared

40   Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-

41   IMPACT) next generation sequencing platform (5).

42

43   **Results**

44   Study population characteristics

45   A total of 2,138 bone and soft tissue sarcoma samples were analyzed. Median patient age was

46   54 years (range < 1– > 90 years); 1,098 (51.4%) were female. Most were primary tumors; 790

47   samples (36.9%) were metastases (**Supplementary Table 1**). The analyzed dataset included

48   45 distinct pathologic entities as assessed by expert sarcoma pathologists. Twenty-two

49   subtypes were represented by ≥ 20 tumor samples and were therefore used as our core

50   subtype set for analyses (**Fig. 1A**). Data from less represented subtypes (**Fig. 1B**) are included

51   in this cohort as a resource. The most common subtypes were gastrointestinal stromal tumor

52   (GIST; n = 395, 18.5%), dedifferentiated liposarcoma (DDLS; n = 167, 5.4%), uterine

53   leiomyosarcoma (ULMS; n = 165, 5.3%), and undifferentiated pleomorphic sarcoma (UPS; n =

54   145, 4.6%) (**Fig. 1B**). Rare subtypes within the core set include angiosarcoma (ANGS; n = 101,

55   3.2%), desmoplastic small round cell tumor (DSRCT; n = 53, 1.7%), and perivascular epithelioid

56   cell neoplasms (PECOMA; n = 30, 0.96%) (**Fig. 1B**). As expected, the age distribution varied

57   among subtypes, as did tumor location (**Fig. 1A, Supplementary Table 1**). Among the more

58   common subtypes, myxofibrosarcoma (MFS) had the oldest median age (68 years), whereas

59   embryonal rhabdomyosarcoma (ERMS) had the youngest (8 years). Similarly, sex distribution

60   was not uniform among subtypes (**Fig. 1A**); PECOMA was more common in females (23/30;

61   76.6%) and DSRCT more common in males (48/53; 90.5%), as was DDLS (males 115/164;

62   68.8%) (**Fig. 1A**). Among the most common subtypes, survival rate differences were most

5

63   apparent starting at 3 years post-sequencing. Myxoid/round cell liposarcoma (MRLS) and GIST

64   patients had the highest 3-year survival rates (both > 75%), whereas ANGS and alveolar

65   rhabdomyosarcoma (ARMS) patients had the lowest (34% and 19%, respectively) (**Fig. 1C**).

66   Subtype-specific mutation analysis

67   Given the heterogeneity in sarcoma subtypes, their biologic behavior, and clinical presentation,

68   we sought to define the genetics of individual subtypes at both the gene (mutation, gene fusion,

69   and copy number alteration) and functional pathway levels. MSK-IMPACT identified at least one

70   driver mutation in the majority of subtypes (**Fig. 2A, Supplementary Table 2**). Overall, TMB

71   among sarcomas was low, whereas the fraction genome altered (FGA) in most cases was

72   relatively high compared with other cancers, consistent with prior reports (**Fig. 2A**) (4). Both

73   varied greatly among sarcoma subtypes, especially FGA. We performed MutSig and MuSiC

74   analyses to identify significantly recurrently mutated genes in each subtype (**Fig. 2B**). As

75   expected, *TP53* and *RB1* were significantly altered across multiple subtypes, but at markedly

76   different frequencies. Within GIST, we identified several frequently mutated genes in addition to

77   previously known drivers such as *KIT*, *SDHA*, and *PDGFRA* (6). These included the histone

78   methyltransferase *SETD2* (4%), the MYC binding partner and transcription factor *MAX* (4%),

79   and *MGA* (3%), which binds the MAX-MYC complex (7).

80   Additional subtypes with recurrently mutated genes of potential biologic or clinical relevance

81   included ANGS (n = 101), in which we identified recurrent mutations in receptor tyrosine kinases

82   involved in angiogenesis including *KDR* (VEGFR2; 19%) and *FLT4* (VEGFR3; 9%) as well as

83   another receptor tyrosine kinase, *EPHA5* (9% of cases). The mutations in *EPHA5* and *FLT4*

84   were all variants of unknown significance (VUS). The VUS in *FLT4* all affect the kinase domain

85   or the C-terminus, implying a possible functional consequence (**Fig. 2C**). In Ewing sarcoma (ES;

86   n = 99), 10% of samples carried mutations in the cohesion complex component *STAG2*,

87   confirming prior reports (8). In ULMS, *MED12*, a member of the transcription elongation

6

88   complex, was altered in 16% of cases, most frequently missense mutations at glycine 44, as

89   reported previously (**Fig. 2B**) (9).

90   In PECOMA, SFT, LMS, ULMS, and ES, driver mutations were represented in a cancer cell

91   fraction (CCF) of close to 1.0, suggesting that these represent a large clonal population (**Fig.**

92   **2D**). The CCF for VUS was overall similar to that of drivers within most subtypes, with the

93   exception of MFS, OS, and UPS, which suggests that in some cases these VUS could have an

94   unrecognized function, calling for further studies to determine their roles in oncogenesis and

95   progression.

96   <u>Copy Number Alterations by Subtype</u>

97   As many sarcomas are driven by copy number alterations, we analyzed these changes across

98   the whole cohort, including in subtypes not classically thought to be driven by them (**Fig. 3**). For

99   instance, in GIST patients (evaluable n = 371), there were frequent copy number loss events

100   involving chromosomes 1, 14, 15, and 22 **(Fig. 3A)**. Translocation-driven sarcomas, e.g. ES,

101   DSRCT, and SYNS, exhibited highly recurrent copy number changes, indicating that there may

102   be additional relevant genetic events beyond the driver translocations **(Fig. 3B)**. We identified a

103   diversity of chromosome arms (e.g. 5p, 8q, and 10p) that were recurrently affected by copy

104   number variation across multiple common subtypes (**Fig. 3B**). Of note, 12q amplifications in

105   DDLS and WDLS patients were not wide enough to be called in arm-level analysis (**Fig. 3B**,

106   left), though they were clearly observed as a strong focal event in copy number profiles (**Fig.**

107   **3A, B**, right). In most cases, these arm-level copy number events were not linked to a specific

108   gene. However, there were some exceptions including significant gains of *MYC* on chromosome

109   8q24 in OS, EPIS, ERMS, and ANGS, as well as significant gain of a negative regulator of NF-

110   kB signaling, *TNFAIP3*, in DDLS (**Fig. 3B**). As expected, we observed more widespread copy

7

111   number changes in classically copy number-driven subtypes such as LMS, ULMS, MFS, and

112   OS compared with the rest of the cohort.

113   Despite sharing *CDK4* and *MDM2* amplification events, DDLS is more aggressive than WDLS

114   and has increased risk for distant spread (10). Therefore, we compared rates of amplifications

115   between WDLS and DDLS, and found greater rates of amplification of the oncogenes *GLI1*

116   (8.5% vs. 25.3%), *TERT* (6.3% vs. 14.4%), and *JUN* (0% vs. 13.8%) in DDLS. The Jun

117   transcription factor positively regulates the expression of cyclin D1, a CDK4/6 cyclin partner,

118   and amplification of JUN in DDLS is associated with a more aggressive phenotype (4), calling

119   for investigation of whether *CDK4* and *JUN* co-amplification drives progression to DDLS or

120   modulates response to *CDK4* inhibition. Amplification of the GLI1 transcription factor,

121   downstream of Sonic Hedgehog (Shh) signaling, has previously been reported (11); this

122   confirmation furthers rationale for studying Shh pathway inhibition in DDLS. *GLI1* amplification

123   and *JUN* amplification were mutually exclusive.

124   OS, UPS, ERMS, and MPNST had high frequencies of WGD, all around 50%, ranking among

125   the highest even among a wide variety of cancers for which WGD was previously analyzed (**Fig.**

126   **3C; Supplementary Fig. S1A**) (12). In keeping with the notion that MFS is on a genetic

127   continuum with UPS (4), UPS and MFS had similar WGD frequencies. Despite being copy

128   number variation (CNV)-driven, WDLS and DDLS had lower rates of WGD frequency, as did

129   many translocation-driven subtypes including SYNS, ES, DSRCT, and MRLS. In sarcomas,

130   WGD was associated with worse overall survival among metastatic (p = 0.042) but not primary

131   cases (p = 0.391; **Supplementary Fig. S1B)**. Among specific subtypes, WGD was associated

132   with worse overall survival (from time of sequencing) in metastatic UPS (p = 0.022; **Fig. 3D**), but

133   not MFS (p = 0.78; **Supplementary Fig. S1C**).

134   Commonly Disrupted Pathways in Sarcoma

135   Pathway-specific analyses within each sarcoma subtype for which ≥ 20 samples were available

136   (**Fig. 4A**, genes in each pathway listed in **Supplementary Table 3)** revealed that a number of

137   pathways important in carcinomas were infrequently altered in sarcoma, including TGFβ, WNT,

138   Hippo, Notch, and NRF2 (**Fig. 4A**, right panel). By contrast, the cell cycle and *TP53* pathways

139   were altered in at least half of samples in 8 of the 22 most common subtypes. For instance,

140   DDLS and WDLS demonstrated cell cycle or *TP53* pathway alterations in 214/215 (99%) of

141   samples, most commonly through co-amplification of *CDK4* and the E3 ubiquitin ligase that

142   targets p53 for degradation, *MDM2* (**Fig. 4A**) (13,14). Many of the sarcomas with infrequent

143   alterations (< 10%) in the cell cycle and *TP53* pathways were driven by translocations (e.g.

144   MRLS, DSRCT, SYNS) or alteration in the SWI/SNF remodeling complex (epithelioid sarcoma

145   [EPIS]) (**Fig. 4A**), highlighting a distinct mechanism of pathogenesis. An exception was solitary

146   fibrous tumor (SFT), which is driven by the *NAB2-STAT6* fusion oncogene, and has oncogenic

147   *TP53* alteration in 28% of cases (**Fig. 4A**) (15).

148   The PI3K pathway was frequently altered in some of the most prevalent subtypes in our dataset

149   including MRLS (41%), PECOMA (40%), ULMS (30%), pleomorphic liposarcoma (PLLS; 22%),

150   UPS (20%), and soft tissue leiomyosarcoma (LMS; 20%) (**Fig. 4A**, right panel). Among these

151   subtypes, *PTEN* and *PIK3CA* were the most frequently affected genes except in PECOMA

152   where *TSC2* loss of function alterations were most common (30%) (**Fig. 4A, 4B**). *PTEN* loss of

153   function alterations predominate in LMS and ULMS (14% and 21%, respectively), whereas in

154   MRLS, *PIK3CA* mutations were most frequent, occurring in 25% of cases, consistent with our

155   prior findings (16). In MRLS, *PTEN* loss is observed in 21% of cases, some of which were

156   concurrent with *PIK3CA* mutations (4 *PIK3CA* mutations in 10 *PTEN* loss cases). In contrast, in

157   UPS, *PTEN* alterations were identified in 8% of samples and *PIK3CA* in 3%; only 1 of the 11

158   cases with a *PTEN* alteration had a concurrent *PIKC3A* mutation. Notably, *PTEN* loss of

9

159   function has also been proposed as a predictor of non-response to immune checkpoint inhibition

160   in ULMS (17).

161   Because a pan-cancer MSK-IMPACT analysis identified *TERT* promoter mutations in a subset

162   of sarcomas (18), we investigated their frequency as a function of sarcoma subtype (**Fig. 4C**).

163   We identified oncogenic *TERT* amplifications in 44% (8/18) of intimal sarcoma (INTS) and *TERT*

164   promoter mutations in 79% (38/48) of MRLS, 46% (24/52) of SFT, and 35% (5/14) of

165   dedifferentiated chondrosarcoma (DDCHS). In DDLS, oncogenic *TERT* promoter alterations

166   were present in 16% of samples (27/167) and were almost entirely amplifications (n = 24).

167   *TERT* copy number alterations have not yet been described in INTS, perhaps due to the low

168   incidence of this rare subtype. The *TERT* locus is distinct from that of the *MDM2* and *CDK4*

169   amplifications (19) that are hallmarks of INTS, implicating *TERT* amplification as a potential

170   independent contributor to pathogenesis.

171   Alterations in DNA damage repair (DDR) pathway genes have been associated with

172   development of sarcomas (20), and are of particular clinical interest as PARP inhibition has

173   activity in select carcinomas with homologous recombination deficiency and immune checkpoint

174   blockade has activity in certain tumors with microsatellite instability (21,22). Our analysis of

175   DDR pathway alterations found that 9.6% of all samples harbored an oncogenic somatic

176   alteration in a DDR pathway. Among subtypes with more than 20 samples, the frequency of

177   DDR gene alterations was highest in ULMS (24%), MPNST (16%), PLLS (13%), PECOMA

178   (13%), ANGS (13%), LMS (10%), and OS (10%) (**Fig. 4A**, right panel). The most frequently

179   altered genes across subtypes were *BRCA2* (1.4% of all samples)*, RAD51B* (1.1%)*, CHEK2*

180   (1.0%) *ATM* (0.9%)*, FANCA* (0.6%), and *RAD51* (0.6%)*. Consistent with a previous report in

181   uterine sarcomas (23), nearly half of *BRCA2* (41%) and *RAD51B* (47%) alterations occurred in

182   sarcomas of uterine origin, with *RAD51B* or *BRCA2* each mutated in 7% of ULMS cases.

183   Similarly, 35% of the 14 uterine adenosarcomas also had an altered DDR gene, all deep

184   deletions. Five percent of ANGS had oncogenic mutations and another 5% had a VUS in *ATM*.

185   Given the association of ANGS with prior ionizing radiation, *ATM* mutations may represent a

186   convergent pathogenic mechanism for accumulation of DNA damage. Of the 15 sarcomas

187   (0.7%) with an altered mismatch repair (MMR) gene (*MLH1, MSH2, MSH6,* or *PMS2*), one

188   (LMS) was microsatellite instability (MSI)-high by MSISensor and had a high TMB.

189   Epigenetic dysregulation contributes to the pathogenesis of several sarcoma subtypes (24). In

190   SYNS, EPIS, malignant rhabdoid tumors, and MPNST, this occurs through alterations of

191   chromatin-remodeling and -modifying complexes; in chondroblastoma, CHS, UPS, giant cell

192   tumors of bone, and osteosarcoma through oncogenic histone mutations. In light of emerging

193   pharmacologic strategies to study and therapeutically target epigenetic regulatory proteins, we

194   identified sarcoma subtypes characterized by epigenetic pathway alterations (**Fig. 4A**;

195   **Supplementary Table 3**) (25). As expected, 75% of EPIS had loss-of-function deletions,

196   truncating mutations, or intragenic fusions in *SMARCB1*. In addition, the epigenetic pathway

197   was one of the most altered pathways among the highly prevalent subtypes in our dataset.

198   Pathogenic alterations in epigenetic pathway genes (**Supplementary Table 3**) were observed

199   in 64% of MPNST, 49% of ULMS, 45% of PLLS, 43% of CHS, 42% of UPS, 36% of MFS, and

200   32% of OS (**Fig. 4A**). By contrast, these alterations were infrequently observed (< 10%) in

201   WDLS, ARMS, and MRLS, suggesting subtype specificity.

202   We determined the association with specific subtypes of epigenetic pathway genes contributing

203   to a specific biochemical function (e.g. DNA methylation, chromatin remodeling) and complex

204   (e.g. PRC1, PRC2, MLL3/4) (**Fig. 4D**; **Supplementary Fig. 2; Supplementary Table 4**). Genes

205   involved in histone modification were altered in 48% of MPNST, 42% of sclerosing epithelioid

206   fibrosarcoma (SEF), 36% of uterine adenosarcoma (UAS), and 36% of high-grade endometrial

207   stromal sarcoma (HGESS). ERMS had frequent alterations in in the transcriptional co-repressor

208   and non-canonical PRC1 complex member *BCOR* (19% total, 16% oncogenic), which were

11

209   mutually exclusive with *DICER1* alterations (12% oncogenic). Both alterations were more

210   prevalent in our population than in prior studies (26,27).

211   Genes involved in chromatin remodeling were altered at similarly high frequencies: 76% of

212   epithelial sarcoma (EPIS), 39% of ULMS, 26% of UPS, 24% of MFS, and 18% of MPNST. In a

213   significant portion of these cases, the histone chaperone *ATRX* drove these high rates

214   (**Supplementary Fig. 2**). We also note the unexpected finding that UAS (n = 14), a rare

215   sarcoma subtype, had oncogenic alterations in genes encoding subunits of the SWI/SNF

216   chromatin remodeling complex in 43% of patients, with *ARID1A* and *PBRM1* most frequently

217   affected (**Fig. 4D**, **Supplementary Fig. 2**). Interestingly, UAS also had alterations in histone-

218   modifying genes in 36% of cases.

219   As epigenetic alterations are more frequent in DDLS (25%) than WDLS (8%) and we have

220   previously found epigenetic dysregulation to contribute to DDLS (28), we further examined

221   differences between DDLS and WDLS. Histone-modifying and histone chaperone/chromatin-

222   remodeling alterations occur in 15% and 13% of DDLS cases, respectively, compared with 4%

223   each of WDLS (**Fig. 4D**). This suggests that loss of epigenetic regulation could be an important

224   contributor to dedifferentiation.

225   We also examined epigenetic pathways without filtering for alterations already established as

226   oncogenic, which is a strategy we recently employed to generate hypotheses in an analysis of

227   genetic alterations in OS (**Supplementary Fig. 3**) (29). This analysis identified the histone

228   methyltransferase *KMT2D/MLL4* as more frequently altered in MFS (16%) compared with other

229   subtypes including the closely related UPS (6%) (4). Histone-modifying enzymes were altered in

230   20% of SYNS, among which *KMT2B* and *SETD2* were altered in 6% and 7% of samples,

231   respectively. We also found that the transcriptional corepressor *NCOR1*, which complexes with

232   *HDAC3* and other deacetylases to regulate the activity of transcription factors such as the

233   retinoic acid receptor and thyroid hormone receptor (30), was altered in 10% of ULMS, 19% of

234   LMS, and 21% of OS, mostly through amplification. *NCOR1* has also been shown to regulate

235   transcription factors important in mesenchymal lineages including the MEF2 family and PPARγ,

236   which regulate myo- and adipogenesis, respectively (31). Several other genes within the same

237   cytoband as *NCOR1*, 17p12-p11.2, were co-amplified, including *FLCN*, *MAP2K4*, *AURKB*, and

238   *ALOX12B* (**Supplementary Fig. 4A**)*.* Amplifications of *MYOCD*, whose genomic location is

239   within a region previously found to be amplified in LMS (32), were not detected because this

240   gene is not represented on the MSK-IMPACT panel. Except for *ALOX12B*, the 17p copy

241   number gains of MSK-IMPACT-assessed genes were associated with increased gene

242   expression in the sarcoma TCGA analysis (**Supplementary Fig. 4B**). Thus, one or more of

243   these genes could play a pathogenic role.

244   <u>Mutual Exclusivity and Co-occurrence</u>

245   To better understand how gene- and pathway-level alterations interact, we analyzed their co-

246   occurrence and mutual exclusivity (**Fig. 5A**). As expected, *KIT* and *PDGFRA* alterations were

247   mutually exclusive in GIST and *CDK4* and *MDM2* co-occurred in DDLS. In OS, *KDR* alterations

248   co-occurred with *KIT* and *PDGFRA,* as did the latter two with each other, suggesting

249   dysregulation of signaling through these 3 RTK genes located at the 4q12 locus (29). *TP53*

250   alterations were mutually exclusive with *CDKN2A/B* in GIST and ULMS, but not in UPS. In SFT

251   and ES, *TP53* alterations co-occurred with *STAG2* and *TERT* alterations, respectively,

252   suggesting context dependence for alterations co-occurring with *TP53*. In UPS, *ATRX* and *NF1*

253   alterations, which are mostly loss-of-function events, were mutually exclusive, suggesting

254   biologically different subgroups.

255   At the pathway level, cell cycle and DDR pathway alterations significantly co-occurred (false

256   discovery rate [FDR] < 0.05) with those in other pathways. For instance, cell cycle alterations

13

257   co-occurred with MYC pathway alterations in GIST, with PI3K pathway alterations in GIST and

258   ULMS, and with RTK/RAS alterations in GIST, OS, and SYNS (**Fig. 5A**). DDR pathway

259   alterations co-occurred with MYC pathway alterations in ULMS and MRLS, epigenetic pathway

260   alterations in GIST and DDLS, Hippo pathway alterations in DDLS, and cell cycle alterations in

261   ES. There were no examples of significant mutual exclusivity at the pathway level.

262   *ATRX* Alterations across Subtypes

263   *ATRX* stood out across subtypes as frequently affected by loss-of-function events (**Fig. 5B**); this

264   gene was altered in ≥ 10% of cases in 7 subtypes: ULMS, PLLS, UPS, MFS, PECOMA, LMS,

265   and ANGS. In ULMS, which had the highest rate of *ATRX* alterations, the frequency was

266   roughly 1 in 3 cases. That *ATRX* loss-of-function events occur in both copy number- and

267   translocation-driven subtypes, although at lower frequency in the latter, raises the possibility that

268   they may serve a fundamental role in the biology of a molecular subset of these subtypes.

269   *ATRX* loss-of-function mutations were more frequent than deletion events, independent of

270   subtype, despite the overall low mutation rate. Our analysis also captured intra- and intergenic

271   *ATRX* fusion events.

272   Unsupervised Clustering of Subtypes

273   To assess genetic similarities among subtypes, we grouped samples on the basis of genetic

274   alterations by unsupervised clustering (**Fig. 5C**), which generated 15 distinct clusters

275   subsequently named according to their prevailing subtype and/or genetic feature. Some

276   subtypes and clusters were closely associated (**Fig. 5D**). These included EPIS and the

277   *SMARCB1* cluster*,* DSRCT and *WT1,* and WDLS and DDLS with *MDM2-CDK4.* These

278   groupings largely reflect known or presumed drivers in these subtypes and reinforces their

279   central roles therein.

280   Other clusters were heterogeneous and comprised of many subtypes. One such cluster was

281   associated with frequent alterations in *TERT*, which dominated by MRLS and SFT, but also

282   included small populations of other sarcoma subtypes. Another cluster lacked any

283   predominantly altered gene. This 'other' cluster included the majority of samples in some

284   histotypes (e.g. ANGS and CHS) but also included samples from multiple subtypes that are

285   represented more commonly in other clusters (e.g. LMS and UPS), suggesting that they may be

286   genetic outliers among those subtypes.

287   Notably, alterations in *TP53*, thought of as a canonical driver in many sarcomas, were

288   associated with 4 closely related clusters, each with a distinct association with co-occurring

289   *ATRX* and/or *RB1* alterations. In contrast, the *RB1*-altered/*TP53*-WT ("*RB1*") group clustered

290   distantly from the *TP53*-altered groups and was more closely related with *CDKN2A*/*B* and 'other'

291   groups. ULMS, PLLS, and UPS were represented in both the *TP53*-altered groups and the *RB1*-

292   altered/*TP53*-WT group.

293   For each subtype we also assigned an entropy score with respect to the clustering assignments

294   (**Fig. 5D**). WDLS, DDLS, and DSRCT had the lowest entropy, suggesting relatively uniform

295   genomic profiles within each subtype, whereas ULMS, UPS, and OS had high entropy,

296   suggesting that these pathologically defined entities harbor multiple distinct genetic variants.

297   Tumor Mutational Burden, Microsatellite Instability, and Mutational Signatures

298   Two recent immune checkpoint blockade trials in sarcoma demonstrated low overall response

299   rates, though rates varied among subtypes (33,34). Thus, predictive biomarkers for response to

300   checkpoint blockade are needed to deconvolute this heterogeneity and aid in the design of

301   future clinical trials. As microsatellite instability predicts response to pembrolizumab (21), and

302   tumors with high TMB are more likely to respond to immune checkpoint blockade (35), we

303   determined MSI status and TMB for each subtype (**Fig. 6A and 6B, Supplementary Table 1**).

15

304   While the median TMB for sarcomas was low compared to many carcinomas (4), there is

305   considerable heterogeneity within and between the more common subtypes in our cohort (inter-

306   subtype median range 0.9–3.0) (**Fig. 6A**). The median TMB was greatest in ANGS (3.0), UPS

307   (2.6), and ULMS (2.6) and lowest in WDLS (0.9), EPIS (0.9), and RCS (other) (0.9). However, in

308   certain subtypes, the distribution of TMB had a long upper tail and was skewed towards higher

309   TMB (**Fig. 6A**). TMB was ≥ 5 mut/Mb in 25% of ANGS, 15% of ULMS and UPS, and 13% of

310   ERMS. Only two subtypes had ≥ 5% of samples with a TMB of ≥ 10 mut/Mb: ANGS (7.6%) and

311   UPS (6.7%).

312   Only 5 of 1893 samples evaluable for MSI status were MSI-high (by MSIsensor score ≥ 10),

313   including one UPS, one LMS, and 3 ULMS (**Fig. 6B**). Of these, 4 were confirmed to be MSI-high

314   by a conventional PCR-based MSI assay. MSIsensor scores varied widely between subtypes

315   (**Supplementary Table 1**). Overall, while microsatellite instability corresponded with high TMB,

316   the inverse was not true.

317   To understand mechanisms contributing to extensive genetic alterations, we examined

318   mutational signatures in samples with ≥ 15 single nucleotide variants (SNVs) (**Fig. 6C**). A UV

319   mutational signature was observed in a subset of ANGS and was most strongly observed in

320   samples at the highest end of the TMB spectrum within that group. Of these 16 samples, 11 had

321   a head and neck primary site. Interestingly, a subset of UPS also harbored a UV signature and

322   a higher TMB, while another subset of highly mutated UPS was dominated by an aging

323   signature, suggesting alternative mechanisms for high TMB within UPS.

324   <u>Subtype-specific actionable alterations</u>

325   Toward improved detection of targetable alterations for each subtype and patient, we analyzed

326   genetic alterations by actionability according to OncoKB (**Fig. 7**) (36). As expected, level 1

327   alterations, defined as FDA-recognized biomarkers for response to an FDA-approved drug,

16

328   were most frequent in GIST due to *KIT* and *PDGFRA* mutations. Similarly, level 1 *SMARCB1*

329   deletion was noted in 66% of EPIS. Level 2 alterations, defined as guideline-supported

330   standard-of-care biomarkers for an FDA-approved drug, were seen in > 90% of WDLS and

331   DDLS related to *CDK4* amplification. In the same subtypes, *MDM2* amplifications in > 90% of

332   cases were deemed Level 3A, for which compelling evidence supports use as a predictive

333   biomarker. Many other observed alterations were classified as Level 3B,

334   defined as standard-of-care or investigational biomarkers that predict response to an FDA-

335   approved or investigational drug in another cancer. Notable examples included a combined 37%

336   prevalence of actionable *TSC1/2* deletions in PECOMA, *IDH1/2* alterations in 27% of CHS, and

337   targetable PI3K pathway (*PIK3CA*, *ATK1*, *MTOR*, or *TSC1*) alterations in a collective 31% of

338   MRLS. Notably, 21% of MRLS cases had Level 4 *PTEN* deletions, for which compelling

339   biological evidence supports their use as a predictive biomarker. Other intriguing Level 4

340   alterations included somatic *NF1* deletions in MPNST (32%), UPS (14%), ERMS (14%), and

341   PLLS (14%), and *CDKN2A* deletions in many subtypes at a rate of up to 48% as seen in

342   MPNST.

343

Discussion

345   To better understand genetic heterogeneity in sarcomas, we analyzed prospectively generated

346   tumor next generation sequencing data from a cohort of 2,138 sarcoma samples representing

347   45 histological subtypes. Across all subtypes, the most common alterations we identified were in

348   cell cycle control and *TP53*, receptor tyrosine kinases/PI3K/RAS, and epigenetic regulators.

349   Previously unreported subtype-specific associations included *TERT* amplification in intimal

350   sarcoma and SWI/SNF complex alterations in uterine adenosarcoma. Tumor mutation burden

351   varied widely between and within subtypes.

17

352   Epigenetic pathway mutations frequently occurred in many subtypes in our cohort, in keeping

353   with an emerging recognition of epigenetic dysregulation as an important factor in the

354   pathogenesis of sarcomas (24). A common epigenetic pathway alteration was amplification of

355   *NCOR1*, particularly in ULMS, LMS, and OS. *NCOR1* is a transcriptional corepressor that

356   regulates transcription factors specific to mesenchymal lineages and can suppress

357   differentiation when overexpressed (31,37). If amplification of *NCOR1* correlates with increased

358   protein levels in these sarcomas, which RNA sequencing data suggests it may, this could lead

359   to altered differentiation and transcriptional programs. Moreover, since the activity of *NCOR1* is

360   modulated by PI3K/Akt-mediated control of nuclear localization, both inhibition of that pathway

361   and of *HDAC3* warrant further exploration as potential therapeutic strategies in *NCOR1*-

362   amplified ULMS, LMS, or OS (38).

363   In uterine adenosarcoma, we identified genetic alterations in the SWI/SNF chromatin

364   remodeling complex in 43% of cases, mostly loss-of-function alterations in *ARID1A* or *PBRM1*.

365   Uterine adenosarcoma is a rare subtype composed of both sarcomatous stroma and benign

366   epithelium, which can behave aggressively, especially in the setting of sarcomatous overgrowth

367   (39). Given the role of epigenetic regulation in determining differentiation, impaired SWI/SNF

368   function could contribute to this phenotype. Histone mutations have been observed in ovarian

369   carcinosarcoma, suggesting that epigenetic dysregulation may be a common mechanism for

370   impaired lineage commitment in Müllerian tumors (40). Given synthetic lethality between the

371   loss of the SWI/SNF component *SMARCB1* in epithelioid sarcoma and EZH2 inhibition with the

372   now FDA-approved drug tazemetostat, EZH2 inhibition may represent a future therapeutic

373   strategy in uterine adenosarcoma (41).

374   We also analyzed genes involved in maintenance of telomeres, whose tumor-suppressive

375   function is dependent on the silencing of *TERT*, a reverse transcriptase and core component of

376   telomerase. Mutations in the *TERT* promoter, first identified in melanoma, lead to increased

18

377   transcription of the TERT gene (42,43). Within our cohort, *TERT* amplification occurs in 44% of

378   intimal sarcomas, which to our knowledge has not been previously reported. Whether this

379   amplification leads to increased expression of the TERT gene product should be investigated,

380   as *TERT* overexpression is known to be oncogenic in certain contexts (44). In addition, our data

381   validate prior findings of *TERT* mutations in MLPS and SFT. However, the rate in SFT was

382   greater than observed in prior studies (45), which may be explained by differences in disease

383   aggressiveness, as *TERT* mutations associate with worse prognosis (46).

384   While we included *ATRX* in the epigenetic pathway gene list owing to its fundamental role,

385   along with *DAXX*, as a histone variant H3.3 chaperone, *ATRX* also participates in other

386   pathways including telomerase-independent alternative lengthening of telomeres (ALT), which

387   has been observed in a number of soft tissue sarcomas including UPS and liposarcoma (47).

388   Because UPS and liposarcoma also harbor *TERT* alterations in a largely non-overlapping

389   pattern, these sarcomas may acquire the ability to aberrantly maintain telomeres through

390   multiple independent mechanisms. In addition to epigenetic and ALT functions, *ATRX* helps

391   maintain genomic integrity (48). Because of the diversity of the physiologic functions of *ATRX,*

392   the role(s) of *ATRX* alterations in sarcomagenesis are difficult to predict *a priori*. Thus,

393   developing tools such as patient-derived cell lines and xenografts to study the impact of these

394   alterations on *ATRX*-dependent functions will be informative. Given the relative frequency of

395   *ATRX* alterations and the inclusion of *ATRX* on MSK-IMPACT and other tumor sequencing

396   platforms, such investigations are eminently feasible.

397   Toward identifying predictive biomarkers for response to immune checkpoint blockade in

398   sarcoma, we analyzed the distribution of MSI-H and high TMB, which are associated with

399   response to these agents in other solid tumors. Almost none of the samples had microsatellite

400   instability and there were relatively few samples with high TMB. However, the upper tail of TMB

401   was relatively long in certain subtypes such as UPS, ANGS, and ULMS. Moreover, we do not

402   yet know whether the TMB cutoff of 10 mutations per megabase, which defines high TMB for

403   carcinomas and predicts response to immune checkpoint blockade, is the appropriate threshold

404   for TMB as a predictive biomarker in mesenchymal neoplasms, let alone specific sarcoma

405   subtypes. Indeed, recent work suggests that the highest quintile of TMB within a specific cancer

406   type is associated with improved outcomes following checkpoint inhibitor therapy and, following

407   from that observation, that the TMB threshold for benefit is not absolute (49). Because both MSI

408   status (via MSIsensor) and TMB can be readily determined from targeted sequencing,

409   correlative analysis of both MSI status and TMB in sarcoma immunotherapy trials on a subtype-

410   specific basis is needed to inform our understanding.

411   Determining the clinical relevance of the landscape of genetic alterations in sarcomas described

412   herein requires a further phase of investigation. Toward improved designs of clinical trials in

413   sarcoma, which have often grouped multiple subtypes together despite significant inter- and

414   intra-subtype genetic variability, future studies should investigate which genetic alterations result

415   in functional effects. This may be particularly important in the subtypes we identified as having

416   high entropy in their genomic clustering. That knowledge will enable the establishment of

417   subtype *and* genotype-based trials to study the effect of novel agents in better defined biologic

418   groups of tumors. The data we present herein and via an accompanying interactive database

419   (cBioPortal link to be provided) will serve as a resource for the field to explore and compare

420   subtype-specific alterations to facilitate this transition in approach.

421

422   **Methods**

423   Patient Cohort

424   This study was approved by the Institutional Review Board at Memorial Sloan Kettering Cancer

425   Center (MSK). We identified patients with a diagnosis of soft tissue or bone sarcoma who had

20

426   tumor and matched normal (white blood cell) tissue sequenced using the MSK-IMPACT assay

427   through December 19, 2019 (5,18). Tumors were sequenced using one of 3 versions of MSK-

428   IMPACT, including 341, 410, or 468 genes, with results reported in the medical record. In

429   patients with multiple samples, only one sample was included in the cohort; those collected

430   earliest and of highest purity and highest average coverage were selected in that order of

431   priority. Clinical characteristics such as patient age, sex, race, and metastatic versus primary

432   site, were annotated per the standard MSK-IMPACT workflow (18).

433   <u>Histologic Analysis</u>

434   Histologic diagnosis was annotated according to the standard MSK-IMPACT workflow. In the

435   case of sarcomas characterized by canonical fusion events, the medical record was queried to

436   ensure that the appropriate fusion event was detected and if not, the sample was reviewed with

437   the assistance of an expert sarcoma pathologist. Similarly, samples harboring a canonical

438   fusion but with a discordant pathologic diagnosis were further reviewed to assign the most

439   appropriate diagnosis. Fusions other than those identified by MSK-IMPACT were annotated at

440   the patient (not sample) level. Samples originally annotated as sarcoma or round cell sarcoma

441   not otherwise specified, rhabdomyosarcoma (without further classification), spindle cell

442   rhabdomyosarcoma, and fibrosarcoma underwent additional medical record review and, in

443   some cases, pathology review to render the most accurate diagnosis possible. In some

444   additional cases with ambiguity in subtype assignment, the diagnosis was updated upon further

445   review by an expert pathologist. We further standardized diagnoses by mapping each tumor to a

446   unique code from the OncoTree ontology (50) except for round cell sarcoma other (RCS (other))

447   and extraskeletal osteosarcoma, which were categories created for this study. Samples that

448   could not be assigned to one of the Oncotree codes (n = 243) were excluded from our analysis

449   cohort.

450   Computational Genomic Analysis

451   Genomic alterations were annotated using the OncoKB precision oncology knowledge base,

452   which identifies functionally relevant cancer variants and their potential clinical actionability (36).

453   Except where otherwise specified in the text, variants of unknown significance, i.e. not labeled

454   as *oncogenic*, *likely oncogenic*, or *predicted oncogenic* in OncoKB were excluded from the

455   analysis. Therapeutically targetable somatic alterations were labeled using levels of clinical

456   actionability defined in OncoKB, which range from level 1, FDA-recognized biomarkers of

457   response to FDA-approved drugs, to level 4, biomarkers of hypothetical relevance based on

458   compelling preclinical biological evidence. Analyses of alterations in oncogenic signaling

459   pathways were performed using the set of pathway definitions previously curated by our group,

460   which we expanded to include the DDR and epigenetic modifier pathways using additional

461   templates curated from literature subtypes (24,51-53).

462   Tumor mutation burden (TMB) was computed as the total number of nonsynonymous mutations

463   divided by the total number of base pairs sequenced per sample. The fraction of the genome

464   altered (FGA) was defined as the fraction of genome with $\log_2$ copy number gain > 0.2 or loss <

465   −0.2 relative to the size of the genome for which copy number was profiled. We computed

466   MSIsensor scores for all samples in the cohort and used a threshold of MSIsensor score ≥ 10 to

467   identify tumors with microsatellite instability (MSI-high) (54). MSI-high was confirmed by a PCR-

468   based assay (Idylla). MSIsensor ≥ 3 and < 10 were labeled indeterminate and samples that did

469   not meet quality control for assigning MSI status were labeled do not report (DNR).

470   Allele-specific copy number estimates at both the gene and chromosome arm levels were

471   computed using the FACETS (Fraction and Allele-Specific Copy Number Estimates from Tumor

472   Sequencing) algorithm, which also provided purity-corrected segmentation files and allowed

473   identification of whole-genome duplication events (55). FACETS output was also used to infer

22

474   the cancer cell fraction associated with individual mutations for clonality analyses. Significantly

475   recurrently mutated genes were identified using the MuSic and MutSigCV 1.4 algorithms, with a

476   threshold q-value of 0.1 (56,57). Dimensionality reduction was performed using Uniform

477   Manifold Approximation and Projection (UMAP) (http://arxiv.org/abs/1802.03426) and clusters

478   were identified using Hierarchical Density-Based Spatial Clustering of Applications with Noise

479   (HDBSCAN) (58). Shannon entropy was calculated from observed cluster assignment by

480   subtype and reported in natural units.

481   Mutational signatures for samples with ≥ 15 synonymous and nonsynonymous single nucleotide

482   variants (SNVs) were extracted using the COSMIC v3 catalog of exome reference signatures

483   and default parameters (59) (https://github.com/mskcc/tempoSig). For mutational signatures to

484   be considered detectable, we required a p-value < 0.05 and a minimum of 1 observed mutation

485   attributed to the signature, where the number of observed mutations was defined as the

486   observed mutational signature fraction multiplied by the number of SNVs per sample.

487   Data Availability

488   All clinical and genomic data described in this manuscript will be accessible online and publicly

489   available for bulk download through the cBioPortal for Cancer Genomics (60).

490

491   **Author contributions**

492   Conception and design: Nacev, Sanchez-Vega, Gounder, Bowler, Singer, Tap

493   Development of methodology: Nacev, Sanchez-Vega, Smith, Shi, Tang, Socci

494   Data collection: Nacev, Antonescu, Zehir, Gounder, Chi, D'Angelo, Dickson, Kelly, Keohan,

495   Movva, Thornton, Meyers, Wexler, Slotkin, Glade Bender, Shukla, Hensley, Healey, La

496   Quaglia, Crago, Yoon, Untch, Chiang, Agaram, Hameed, Berger, Solit, Ladanyi, Singer, Tap

23

497   Data curation: Nacev, Sanchez-Vega, Smith, Antonescu, Bowler

498   Resources: Sanchez-Vega, Smith, Shi, Tang, Zehir, Berger, Schultz

499   Analysis and interpretation of data: Nacev, Sanchez-Vega, Smith, Antonescu, Rosenbaum, Shi,

500   Tang, Socci, Rana, Gularte-Mérida, Zehir, Luthra, Jadeja, Okada, Strong, Stoller, Berger,

501   Solit, Schultz, Singer, Tap

502   Drafting the manuscript – Nacev, Sanchez-Vega, Rosenbaum, Singer, Tap

503   Review and revision of the manuscript: All authors

504   Visualization of data: Sanchez-Vega, Smith, Shi, Tang

505   Study supervision and funding acquisition: Singer, Tap

506

507   **Competing interests**

508   MMG has served on advisory boards for Athenex, Ayala, Bayer, Boehringer Ingelheim, Daiichi

509   Sankyo, Epizyme, Karyopharm, Rain, SpringWorks Therapeutics, Tracon, and TYME

510   Technologies; provides consulting services through Guidepoint, GLG Pharma, Third Bridge, and

511   Flatiron Health; has received speaking honoraria from Medscape, More Health, Physicians

512   Education Resource and touchIME; receives publishing royalties from Wolters Kluwer; holds a

513   patent for a patient-reported outcome tool licensed through the institution; and has performed

514   research without compensation in collaboration with Foundation Medicine. TGB is currently

515   employed by Pfizer. PC has served on advisory boards or consulted for Deciphera, Exelixis,

516   NingboNewBay Medical Technology, Novartis, and Zai Lab, and has received institutional

517   research funding from Deciphera, Ningbo NewBay Medical Technology, Novartis, and

518   Pfizer/Array. SPD has received institutional research funding from Amgen, Bristol Meyers

519   Squibb, Deciphera, EMD Serono, Incyte, Merck, and Nektar Therapeutics, has served as a

520   consultant or on advisory boards for Adaptimmune, Amgen, EMD Serono, GlaxoSmithKline,

521   Immune Design, Immunocore, Incyte, Merck, and Nektar Therapeutics, and has served on data

24

25

548   Pfizer, Scorpion Therapeutics, and Vividion Therapeutics. All author authors have no financial

549   relationships to disclose.

References

1.   Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin **2018**;68(1):7-30 doi 10.3322/caac.21442.

2.   WHO. Soft tissue and bone tumours. Lyon (France): International Agency for Research on Cancer; 2020.

3.   Schaefer IM, Cote GM, Hornick JL. Contemporary Sarcoma Diagnosis, Genetics, and Genomics. J Clin Oncol **2018**;36(2):101-10 doi 10.1200/JCO.2017.74.9374.

4.   Cancer Genome Atlas Research N. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. Cell **2017**;171(4):950-65 e28 doi 10.1016/j.cell.2017.10.014.

5.   Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A*, et al.* Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. J Mol Diagn **2015**;17(3):251-64 doi 10.1016/j.jmoldx.2014.12.006.

6.   von Mehren M, Joensuu H. Gastrointestinal Stromal Tumors. J Clin Oncol **2018**;36(2):136-43 doi 10.1200/JCO.2017.74.9705.

7.   Schaefer IM, Wang Y, Liang CW, Bahri N, Quattrone A, Doyle L*, et al.* MAX inactivation is an early event in GIST development that regulates p16 and cell proliferation. Nat Commun **2017**;8:14674 doi 10.1038/ncomms14674.

8.   Agelopoulos K, Richter GH, Schmidt E, Dirksen U, von Heyking K, Moser B*, et al.* Deep Sequencing in Conjunction with Expression and Functional Analyses Reveals Activation of FGFR1 in Ewing Sarcoma. Clin Cancer Res **2015**;21(21):4935-46 doi 10.1158/1078-0432.CCR-14-2744.

574   9.   Perot G, Croce S, Ribeiro A, Lagarde P, Velasco V, Neuville A*, et al.* MED12 alterations

575   in both human benign and malignant uterine soft tissue tumors. PLoS One

576   **2012**;7(6):e40015 doi 10.1371/journal.pone.0040015.

577   10.   Tan MC, Brennan MF, Kuk D, Agaram NP, Antonescu CR, Qin LX*, et al.* Histology-

578   based Classification Predicts Pattern of Recurrence and Improves Risk Stratification in

579   Primary Retroperitoneal Sarcoma. Ann Surg **2016**;263(3):593-600 doi

580   10.1097/SLA.0000000000001149.

581   11.   Agaram NP, Zhang L, Sung YS, Singer S, Stevens T, Prieto-Granada CN*, et al.* GLI1-

582   amplifications expand the spectrum of soft tissue neoplasms defined by GLI1 gene

583   fusions. Mod Pathol **2019**;32(11):1617-26 doi 10.1038/s41379-019-0293-x.

584   12.   Bielski CM, Zehir A, Penson AV, Donoghue MTA, Chatila W, Armenia J*, et al.* Genome

585   doubling shapes the evolution and prognosis of advanced cancers. Nat Genet

586   **2018**;50(8):1189-95 doi 10.1038/s41588-018-0165-1.

587   13.   Oliner JD, Kinzler KW, Meltzer PS, George DL, Vogelstein B. Amplification of a gene

588   encoding a p53-associated protein in human sarcomas. Nature **1992**;358(6381):80-3 doi

589   10.1038/358080a0.

590   14.   Singer S, Socci ND, Ambrosini G, Sambol E, Decarolis P, Wu Y*, et al.* Gene expression

591   profiling of liposarcoma identifies distinct biological types/subtypes and potential

592   therapeutic targets in well-differentiated and dedifferentiated liposarcoma. Cancer Res

593   **2007**;67(14):6626-36 doi 10.1158/0008-5472.CAN-07-0584.

594   15.   Chmielecki J, Crago AM, Rosenberg M, O'Connor R, Walker SR, Ambrogio L*, et al.*

595   Whole-exome sequencing identifies a recurrent NAB2-STAT6 fusion in solitary fibrous

596   tumors. Nat Genet **2013**;45(2):131-2 doi 10.1038/ng.2522.

597   16.   Barretina J, Taylor BS, Banerji S, Ramos AH, Lagos-Quintana M, Decarolis PL*, et al.*

598   Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy.

599   Nat Genet **2010**;42(8):715-21 doi 10.1038/ng.619.

600   17.   George S, Miao D, Demetri GD, Adeegbe D, Rodig SJ, Shukla S, *et al.* Loss of PTEN Is

601   Associated with Resistance to Anti-PD-1 Checkpoint Blockade Therapy in Metastatic

602   Uterine Leiomyosarcoma. Immunity **2017**;46(2):197-204 doi

603   10.1016/j.immuni.2017.02.001.

604   18.   Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, *et al.* Mutational landscape

605   of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.

606   Nat Med **2017**;23(6):703-13 doi 10.1038/nm.4333.

607   19.   Bode-Lesniewska B, Zhao J, Speel EJ, Biraima AM, Turina M, Komminoth P, *et al.*

608   Gains of 12q13-14 and overexpression of mdm2 are frequent findings in intimal

609   sarcomas of the pulmonary artery. Virchows Arch **2001**;438(1):57-65 doi

610   10.1007/s004280000313.

611   20.   Ballinger ML, Goode DL, Ray-Coquard I, James PA, Mitchell G, Niedermayr E, *et al.*

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

613   Lancet Oncol **2016**;17(9):1261-71 doi 10.1016/S1470-2045(16)30147-4.

614   21.   Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, *et al.* Mismatch repair

615   deficiency predicts response of solid tumors to PD-1 blockade. Science

616   **2017**;357(6349):409-13 doi 10.1126/science.aan6733.

617   22.   Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science

618   **2017**;355(6330):1152-8 doi 10.1126/science.aam7344.

619   23.   Hensley ML, Chavan SS, Solit DB, Murali R, Soslow R, Chiang S, *et al.* Genomic

620   Landscape of Uterine Sarcomas Defined through Prospective Clinical Sequencing. Clin

621   Cancer Res **2020** doi 10.1158/1078-0432.CCR-19-3959.

622   24.   Nacev BA, Jones KB, Intlekofer AM, Yu JSE, Allis CD, Tap WD, *et al.* The epigenomics

623   of sarcoma. Nat Rev Cancer **2020** doi 10.1038/s41568-020-0288-4.

624 25.   Cermakova K, Hodges HC. Next-Generation Drugs and Probes for Chromatin Biology:

625   From Targeted Protein Degradation to Phase Separation. Molecules **2018**;23(8) doi

626   10.3390/molecules23081958.

627 26.   Shern JF, Chen L, Chmielecki J, Wei JS, Patidar R, Rosenberg M*, et al.* Comprehensive

628   genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a

629   common genetic axis in fusion-positive and fusion-negative tumors. Cancer Discov

630   **2014**;4(2):216-31 doi 10.1158/2159-8290.CD-13-0639.

631 27.   Doros L, Yang J, Dehner L, Rossi CT, Skiver K, Jarzembowski JA*, et al.* DICER1

632   mutations in embryonal rhabdomyosarcomas from children with and without familial

633   PPB-tumor predisposition syndrome. Pediatr Blood Cancer **2012**;59(3):558-60 doi

634   10.1002/pbc.24020.

635 28.   Taylor BS, DeCarolis PL, Angeles CV, Brenet F, Schultz N, Antonescu CR*, et al.*

636   Frequent alterations and epigenetic silencing of differentiation pathway genes in

637   structurally rearranged liposarcomas. Cancer Discov **2011**;1(7):587-97 doi

638   10.1158/2159-8290.CD-11-0181.

639 29.   Suehara Y, Alex D, Bowman A, Middha S, Zehir A, Chakravarty D*, et al.* Clinical

640   Genomic Sequencing of Pediatric and Adult Osteosarcoma Reveals Distinct Molecular

641   Subsets with Potentially Targetable Alterations. Clin Cancer Res **2019**;25(21):6346-56

642   doi 10.1158/1078-0432.CCR-18-4032.

643 30.   Mottis A, Mouchiroud L, Auwerx J. Emerging roles of the corepressors NCoR1 and

644   SMRT in homeostasis. Genes Dev **2013**;27(8):819-35 doi 10.1101/gad.214023.113.

645 31.   Yamamoto H, Williams EG, Mouchiroud L, Canto C, Fan W, Downes M*, et al.* NCoR1 is

646   a conserved physiological modulator of muscle mass and oxidative function. Cell

647   **2011**;147(4):827-39 doi 10.1016/j.cell.2011.10.017.

648   32.   Agaram NP, Zhang L, LeLoarer F, Silk T, Sung YS, Scott SN*, et al.* Targeted exome

649   sequencing profiles genetic alterations in leiomyosarcoma. Genes Chromosomes

650   Cancer **2016**;55(2):124-30 doi 10.1002/gcc.22318.

651   33.   D'Angelo SP, Mahoney MR, Van Tine BA, Atkins J, Milhem MM, Jahagirdar BN*, et al.*

652   Nivolumab with or without ipilimumab treatment for metastatic sarcoma (Alliance

653   A091401): two open-label, non-comparative, randomised, phase 2 trials. Lancet Oncol

654   **2018**;19(3):416-26 doi 10.1016/s1470-2045(18)30006-8.

655   34.   Tawbi HA, Burgess M, Bolejack V, Van Tine BA, Schuetze SM, Hu J*, et al.*

656   Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a

657   multicentre, two-cohort, single-arm, open-label, phase 2 trial. Lancet Oncol

658   **2017**;18(11):1493-501 doi 10.1016/s1470-2045(17)30624-1.

659   35.   Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ*, et al.* Cancer

660   immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small

661   cell lung cancer. Science **2015**;348(6230):124-8 doi 10.1126/science.aaa1348.

662   36.   Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J*, et al.* OncoKB: A

663   Precision Oncology Knowledge Base. JCO Precis Oncol **2017**;2017 doi

664   10.1200/PO.17.00011.

665   37.   Park DM, Li J, Okamoto H, Akeju O, Kim SH, Lubensky I*, et al.* N-CoR pathway

666   targeting induces glioblastoma derived cancer stem cell differentiation. Cell Cycle

667   **2007**;6(4):467-70 doi 10.4161/cc.6.4.3856.

668   38.   Hermanson O, Jepsen K, Rosenfeld MG. N-CoR controls differentiation of neural stem

669   cells into astrocytes. Nature **2002**;419(6910):934-9 doi 10.1038/nature01156.

670   39.   Carroll A, Ramirez PT, Westin SN, Soliman PT, Munsell MF, Nick AM*, et al.* Uterine

671   adenosarcoma: an analysis on management, outcomes, and risk factors for recurrence.

672   Gynecol Oncol **2014**;135(3):455-61 doi 10.1016/j.ygyno.2014.10.022.

31

673   40.   Zhao S, Bellone S, Lopez S, Thakral D, Schwab C, English DP, *et al.* Mutational

674   landscape of uterine and ovarian carcinosarcomas implicates histone genes in epithelial-

675   mesenchymal transition. Proc Natl Acad Sci U S A **2016**;113(43):12238-43 doi

676   10.1073/pnas.1614120113.

677   41.   Gounder M, Schöffski P, Jones RL, Agulnik M, Cote GM, Villalobos VM*, et al.*

678   Tazemetostat in advanced epithelioid sarcoma with loss of INI1/SMARCB1: an

679   international, open-label, phase 2 basket study. Lancet Oncol **2020**;21(11):1423-1432

680   doi 10.1016/S1470-2045(20)30451-4.

681   42.   Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A*, et al.* TERT promoter

682   mutations in familial and sporadic melanoma. Science **2013**;339(6122):959-61 doi

683   10.1126/science.1230062.

684   43.   Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. Highly recurrent TERT

685   promoter mutations in human melanoma. Science **2013**;339(6122):957-9 doi

686   10.1126/science.1229259.

687   44.   Peifer M, Hertwig F, Roels F, Dreidax D, Gartlgruber M, Menon R*, et al.* Telomerase

688   activation by genomic rearrangements in high-risk neuroblastoma. Nature

689   **2015**;526(7575):700-4 doi 10.1038/nature14980.

690   45.   Campanella NC, Penna V, Abrahao-Machado LF, Cruvinel-Carloni A, Ribeiro G, Soares

691   P*, et al.* TERT promoter mutations in soft tissue sarcomas. Int J Biol Markers

692   **2016**;31(1):e62-7 doi 10.5301/jbm.5000168.

693   46.   Bahrami A, Lee S, Schaefer IM, Boland JM, Patton KT, Pounds S*, et al.* TERT promoter

694   mutations and prognosis in solitary fibrous tumor. Mod Pathol **2016**;29(12):1511-22 doi

695   10.1038/modpathol.2016.126.

696   47.   Henson JD, Hannay JA, McCarthy SW, Royds JA, Yeager TR, Robinson RA*, et al.* A

697   robust assay for alternative lengthening of telomeres in tumors shows the significance of

32

698   alternative lengthening of telomeres in sarcomas and astrocytomas. Clin Cancer Res

699   **2005**;11(1):217-25.

700   48.   Huh MS, Ivanochko D, Hashem LE, Curtin M, Delorme M, Goodall E*, et al.* Stalled

701   replication forks within heterochromatin require ATRX for protection. Cell Death Dis

702   **2016**;7:e2220 doi 10.1038/cddis.2016.121.

703   49.   Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY*, et al.*

704   Tumor mutational load predicts survival after immunotherapy across multiple cancer

705   types. Nat Genet **2019**;51(2):202-6 doi 10.1038/s41588-018-0312-8.

706   50.   Kundra R, Zhang H, Sheridan R, Sirintrapun SJ, Wang A, Ochoa A*, et al.* OncoTree: A

707   Cancer Classification System for Precision Oncology. JCO Clin Cancer Inform

708   **2021**;5:221-30 doi 10.1200/CCI.20.00108.

709   51.   Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC*, et al.* Oncogenic

710   Signaling Pathways in The Cancer Genome Atlas. Cell **2018**;173(2):321-37 e10 doi

711   10.1016/j.cell.2018.03.035.

712   52.   Nacev BA, Feng L, Bagert JD, Lemiesz AE, Gao J, Soshnev AA*, et al.* The expanding

713   landscape of 'oncohistone' mutations in human cancers. Nature **2019**;567(7749):473-8

714   doi 10.1038/s41586-019-1038-1.

715   53.   Rosenbaum E, Jonsson P, Seier K, Qin LX, Chi P, Dickson M*, et al.* Clinical Outcome of

716   Leiomyosarcomas With Somatic Alteration in Homologous Recombination Pathway

717   Genes. JCO Precis Oncol **2020**;4 doi 10.1200/PO.20.00122.

718   54.   Middha S, Zhang L, Nafa K, Jayakumaran G, Wong D, Kim HR*, et al.* Reliable Pan-

719   Cancer Microsatellite Instability Assessment by Using Targeted Next-Generation

720   Sequencing Data. JCO Precis Oncol **2017**;2017 doi 10.1200/PO.17.00084.

721   55.   Shen R, Seshan VE. FACETS: allele-specific copy number and clonal heterogeneity

722   analysis tool for high-throughput DNA sequencing. Nucleic Acids Res **2016**;44(16):e131

723   doi 10.1093/nar/gkw520.

724   56.   Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, Koboldt DC*, et al.* MuSiC:

725   identifying mutational significance in cancer genomes. Genome Res **2012**;22(8):1589-98

726   doi 10.1101/gr.134635.111.

727   57.   Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A*, et al.*

728   Mutational heterogeneity in cancer and the search for new cancer-associated genes.

729   Nature **2013**;499(7457):214-8 doi 10.1038/nature12213.

730   58.   Campello RJGB, Moulavi D, Sander J. Density-Based Clustering Based on Hierarchical

731   Density Estimates. Advances in Knowledge Discovery and Data Mining; 2013; Berlin,

732   Heidelberg. Springer Berlin Heidelberg. p 160-72. (Advances in Knowledge Discovery

733   and Data Mining).

734   59.   Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y*, et al.* The

735   repertoire of mutational signatures in human cancer. Nature **2020**;578(7793):94-101 doi

736   10.1038/s41586-020-1943-3.

737   60.   Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA*, et al.* The cBio cancer

738   genomics portal: an open platform for exploring multidimensional cancer genomics data.

739   Cancer Discov **2012**;2(5):401-4 doi 10.1158/2159-8290.CD-12-0095.

740

741

742   **Supplementary Tables**

743   Supplementary Table 1. Master sample dataset. MSI, microsatellite instability; FGA, fraction of

744   the genome altered; OS, overall survival; FACETS, Fraction and Allele-Specific Copy Number

745   Estimates from Tumor Sequencing algorithm; QC, quality control; CNA, copy number

746   alterations; WGD, whole genome doubling.

747   Supplementary Table 2. Patient characteristics by subtype.

748   Supplementary Table 3. Number of oncogenic fusion events in individual genes in each

749   subtype.

750   Supplementary Table 4. Pathway gene composition. CC, cell cycle; EPI, epigenetic.

751   Supplementary Table 5. Biochemical function/complex assignments for epigenetic pathway
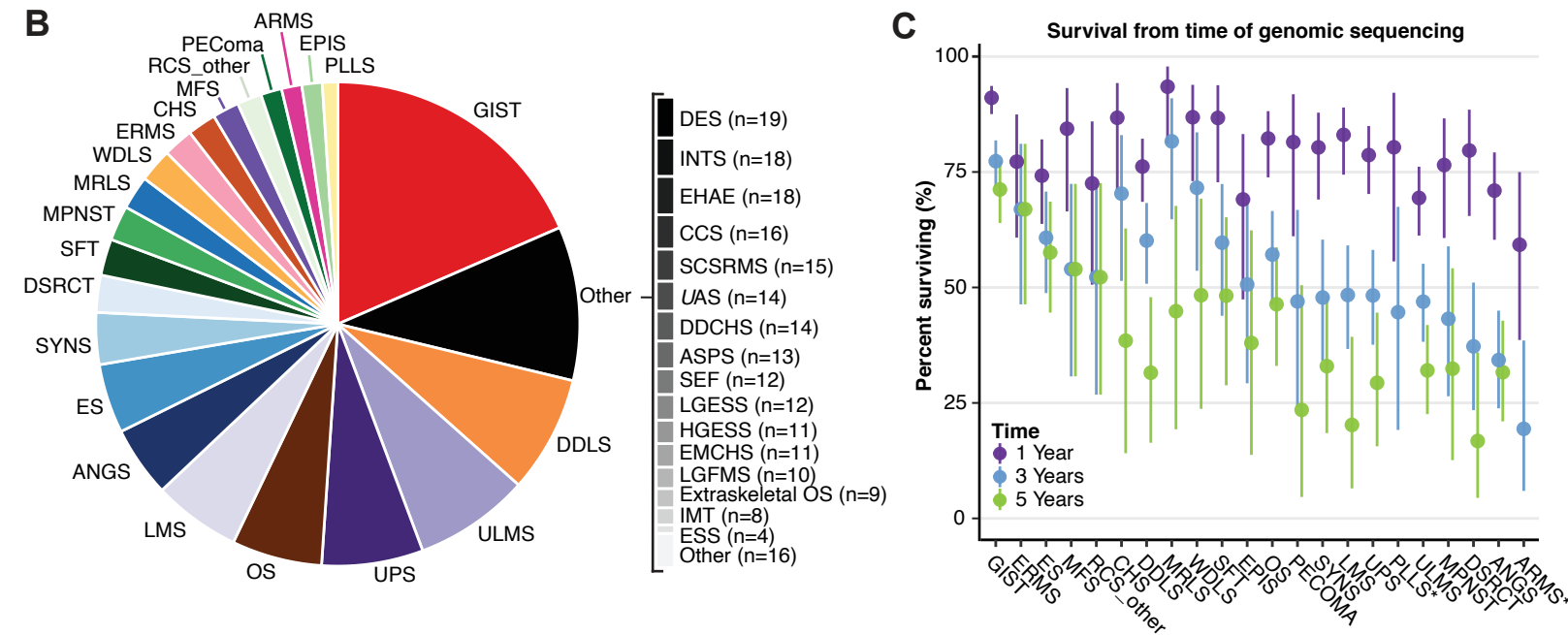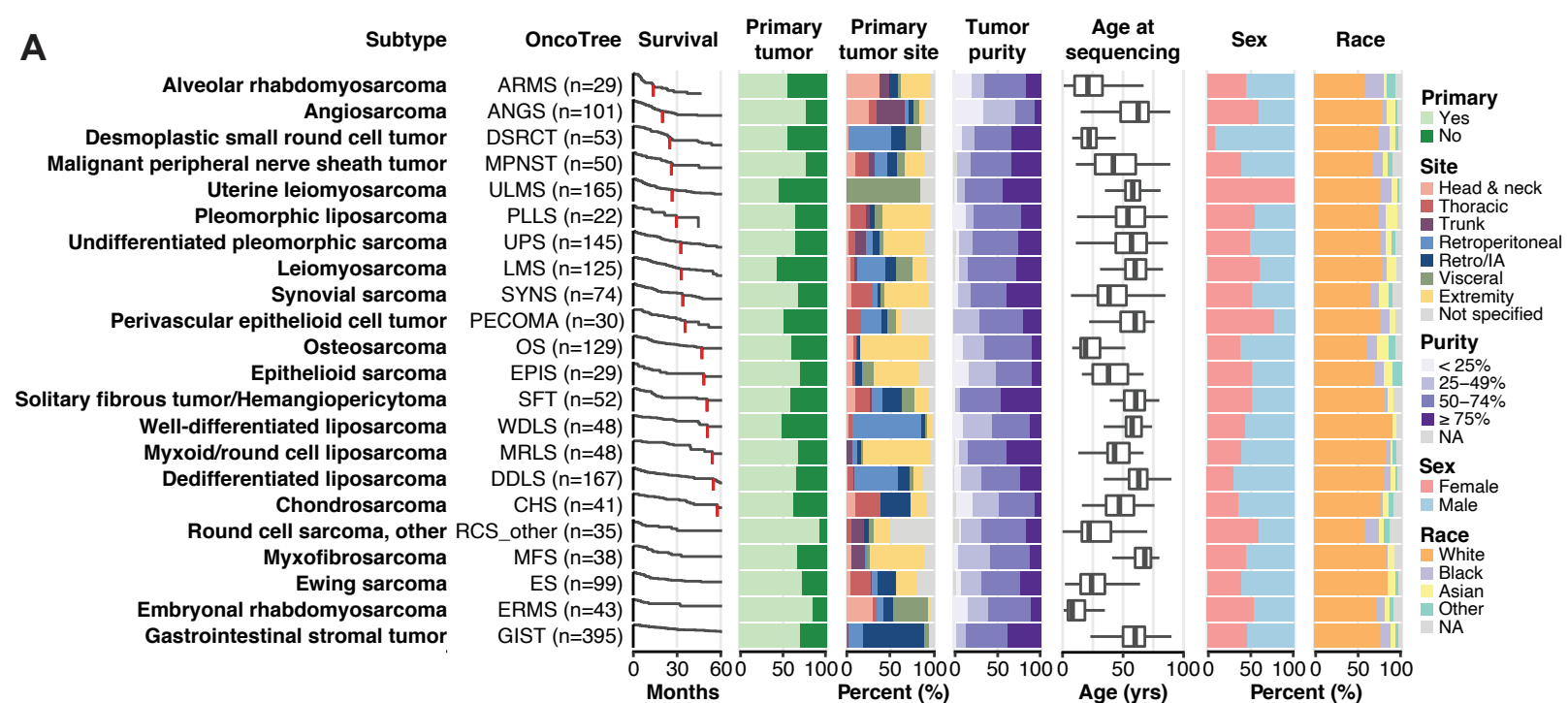
752   genes.

753

35

Figure 1. Sample and patient characteristics. This analysis includes 2,138 bone and soft tissue sarcoma samples, each from distinct patients. Subtypes with ≥ 20 samples in the dataset are displayed. **A**, Distribution of number of samples, survival from time of sequencing, sample type (primary or metastatic site), tumor site, sample purity, age, sex, and race in each subtype. **B**, Overall distribution of sample number for the entire cohort. **C**, 1,3, and 5-year survival from the time of sequencing. *, 5-year survival = 0. DES, desmoid tumor; ESS, endometrial stromal sarcoma; INT, intimal sarcoma; LGFMS, low-grade fibromyxoid sarcoma; EMCHS, extraskeletal myxoid chondrosarcoma; HGESS, high-grade endometrial stromal sarcoma; LGESS, low-grade endometrial stromal sarcoma; SEF, sclerosing epithelioid fibrosarcoma; ASPS, alveolar soft part sarcoma; DDCHS, dedifferentiated chondrosarcoma; UAS, uterine adenosarcoma; SCSRMS, spindle cell/sclerosing rhabdomyosarcoma; CCS, clear cell sarcoma; EHAE, epithelioid hemangioendothelioma.
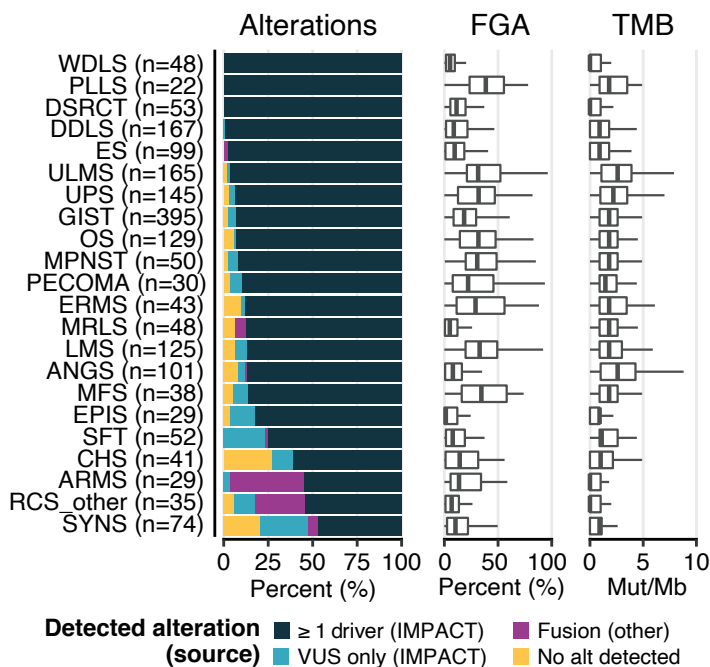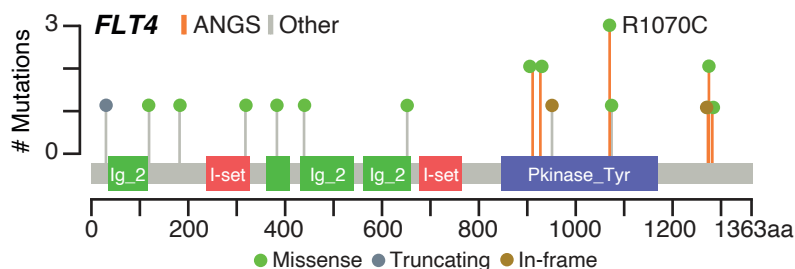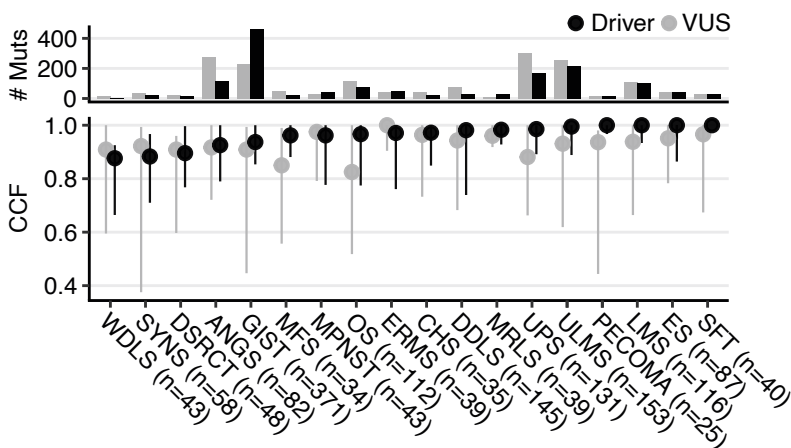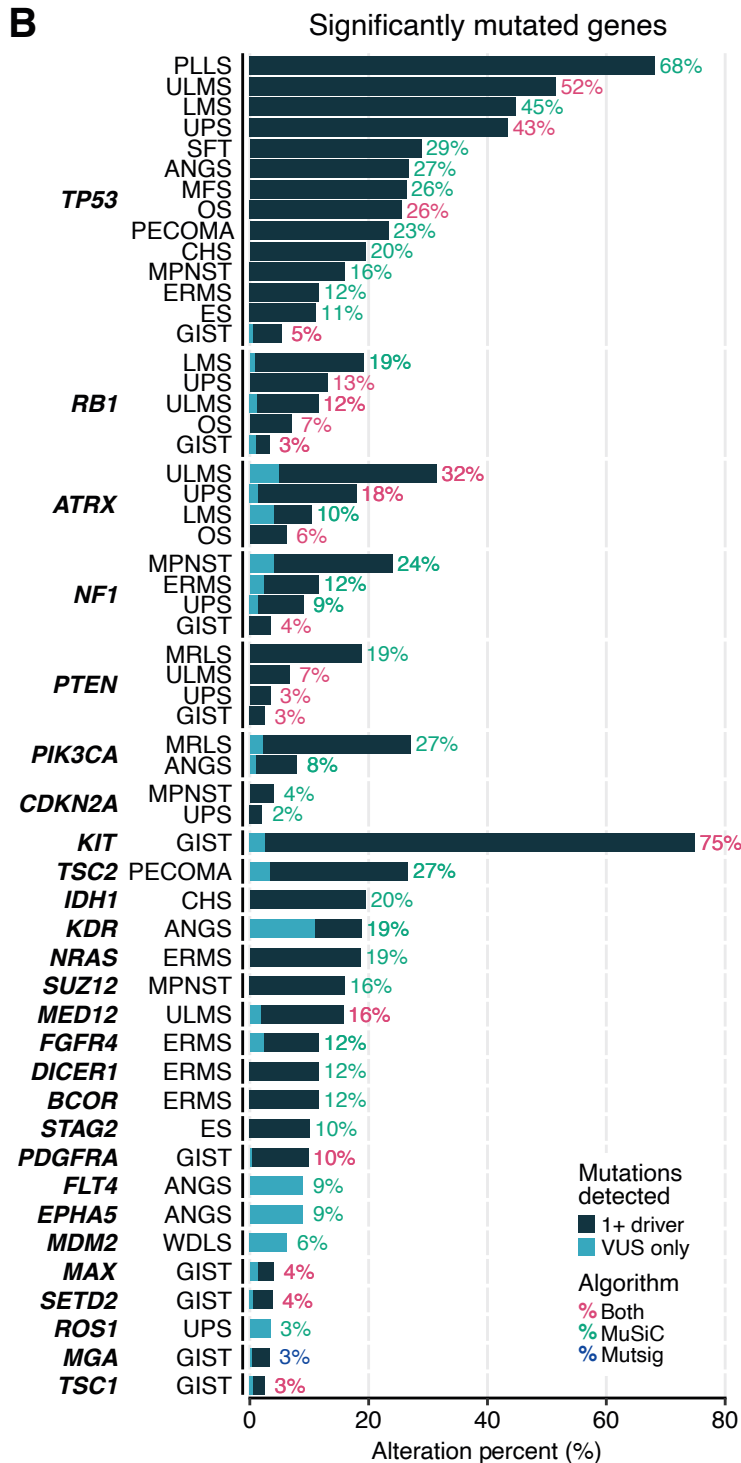
Figure 2. Mutation analysis by subtype. A, Alteration type and frequency, fraction of the genome altered (FGA) and tumor mutation burden (TMB) by subtype. Oncogenic fusions detected by MSK-IMPACT are classified as drivers. **B**, Significant mutations were identified in all subtypes with n ≥ 20 in our dataset using both MutSig and MuSiC analysis. Percentages indicate the percentage of samples with an oncogenic mutation in the corresponding gene. **C**, *FLT4* mutation type, frequency, and location in ANGS vs. other subtypes. **D**, Cancer cell fraction (CCF) and number of mutations for driver mutations and VUS by subtype.
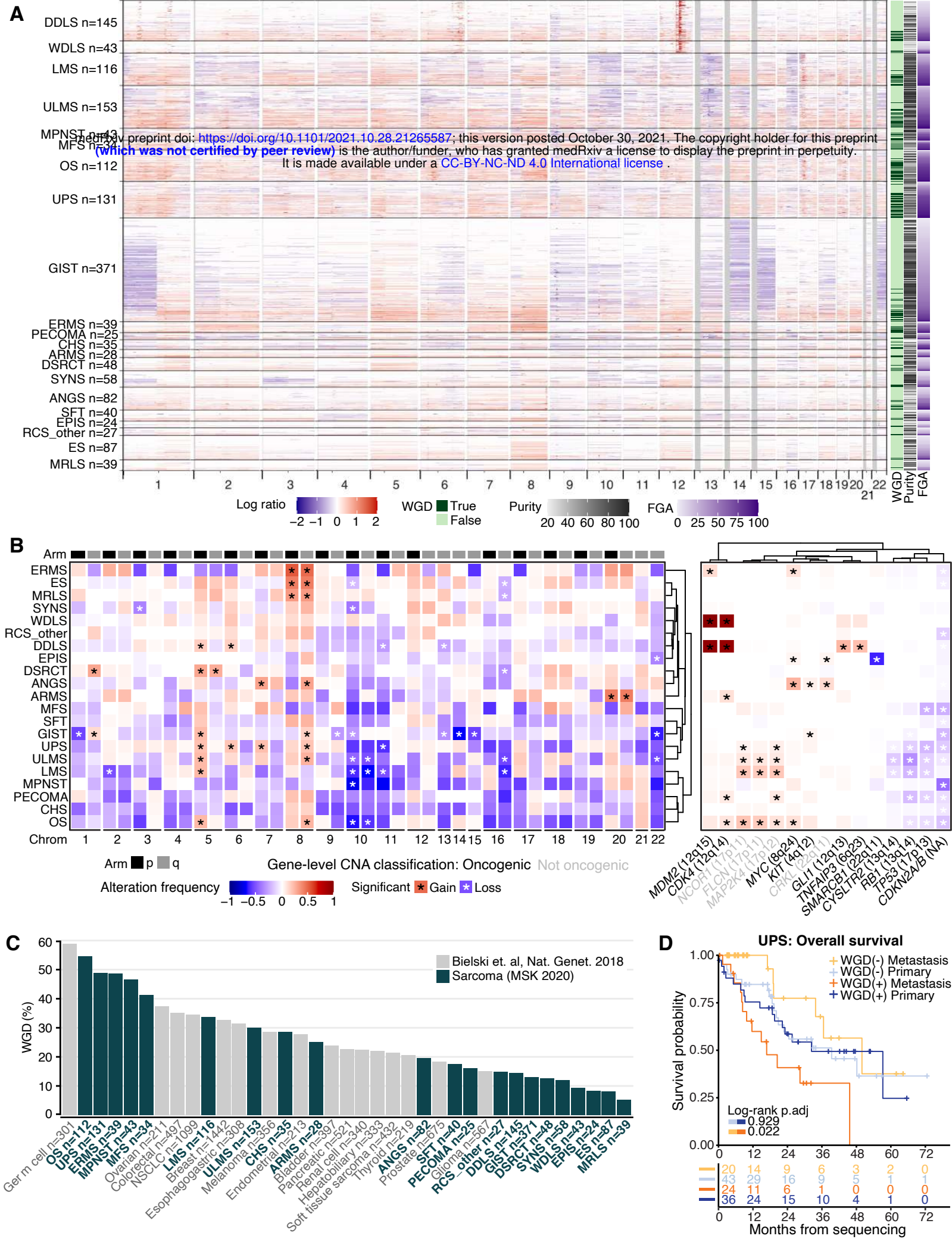
Figure 3. Copy number changes by subtype. A, Individual sample copy number variation (CNV) across the genome for each subtype. Whole genome doubling (WGD), fraction genome altered (FGA), and purity are shown at right. **B**, Aggregate arm-level (left) and gene-level events (right) grouped by subtype. *, significant change based on Bonferroni corrected p-values. Oncogenic vs. non-oncogenic CNV classifications according to OncoKB. **C**, Frequency of WGD by subtype (green) compared to other cancers with ≥ 200 samples available for comparison (gray) (12). **D**, Overall survival based on WGD status within primary and metastatic samples.
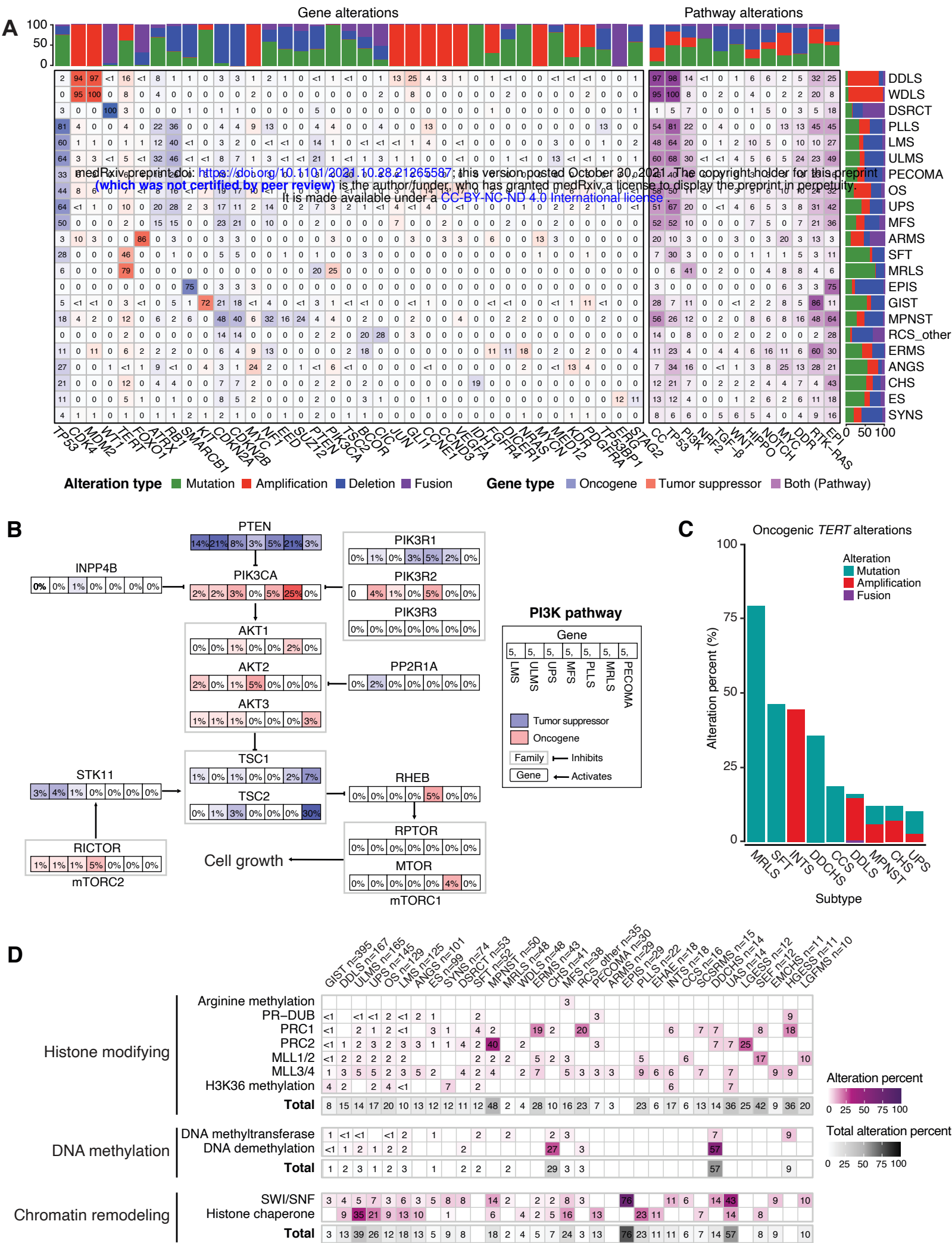
Figure 4. Integrated pathway analysis. A, Oncogenic alterations within each of 12 pathways with relevance to cancer biology in each subtype. Numbers in each cell indicate percentage of samples harboring alterations. Stacked bar graphs indicate the distribution of the type of oncogenic alteration per gene or pathway (top) or subtype (right). CC, cell cycle; EPI, epigenetic. **B,** PI3K pathway alterations in specific subtypes. The percentage of samples with an alteration in a specific gene in each subtype is indicated in each box. **C,** Oncogenic *TERT* promoter alterations in each of the 9 most altered subtypes. **D,** Oncogenic epigenetic pathway alterations by subtype, grouped by complex and/or biochemical function of the encoded protein. Totals include all alterations in genes that belong to a parent category, not only those affecting specific complexes listed.

Figure 5. Mutual exclusivity, co-occurrence, *ATRX* alterations, and unsupervised clustering based on genetic signatures.
A, Co-occurrence and mutual exclusivity of gene- (top) and pathway-level (bottom) alterations in each subtype. **B**, Frequency and types of oncogenic *ATRX* alterations in each of the 14 most altered subtypes. **C**, Unsupervised clustering of samples based on oncogenic alteration patterns. D, Subtype-specific cluster associations and entropy scores.
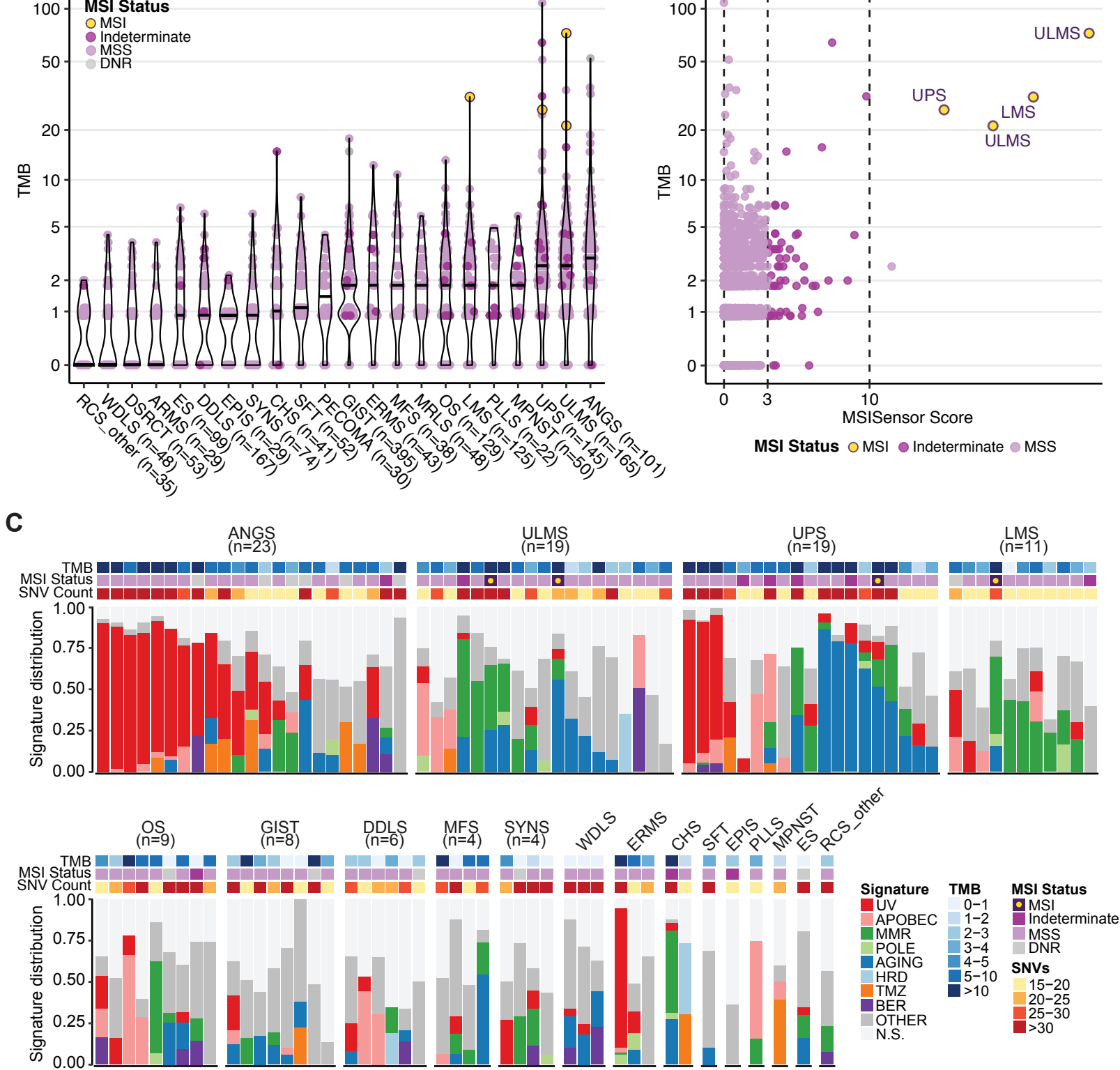
Figure 6. Tumor mutation burden (TMB), microsatellite instability (MSI) status, and mutational signatures. A, Distribution of TMB (mutations/Mb) by subtype with sample-level MSI status. **B**, MSISensor score versus TMB for the cohort with sample-level MSI status. **C**, Mutational signature profile of each evaluable sample. Each bar represents the signature distribution of an individual sample, grouped by subtype. TMB, MSI status, and single nucleotide polymorphism (SNV) count are shown above.
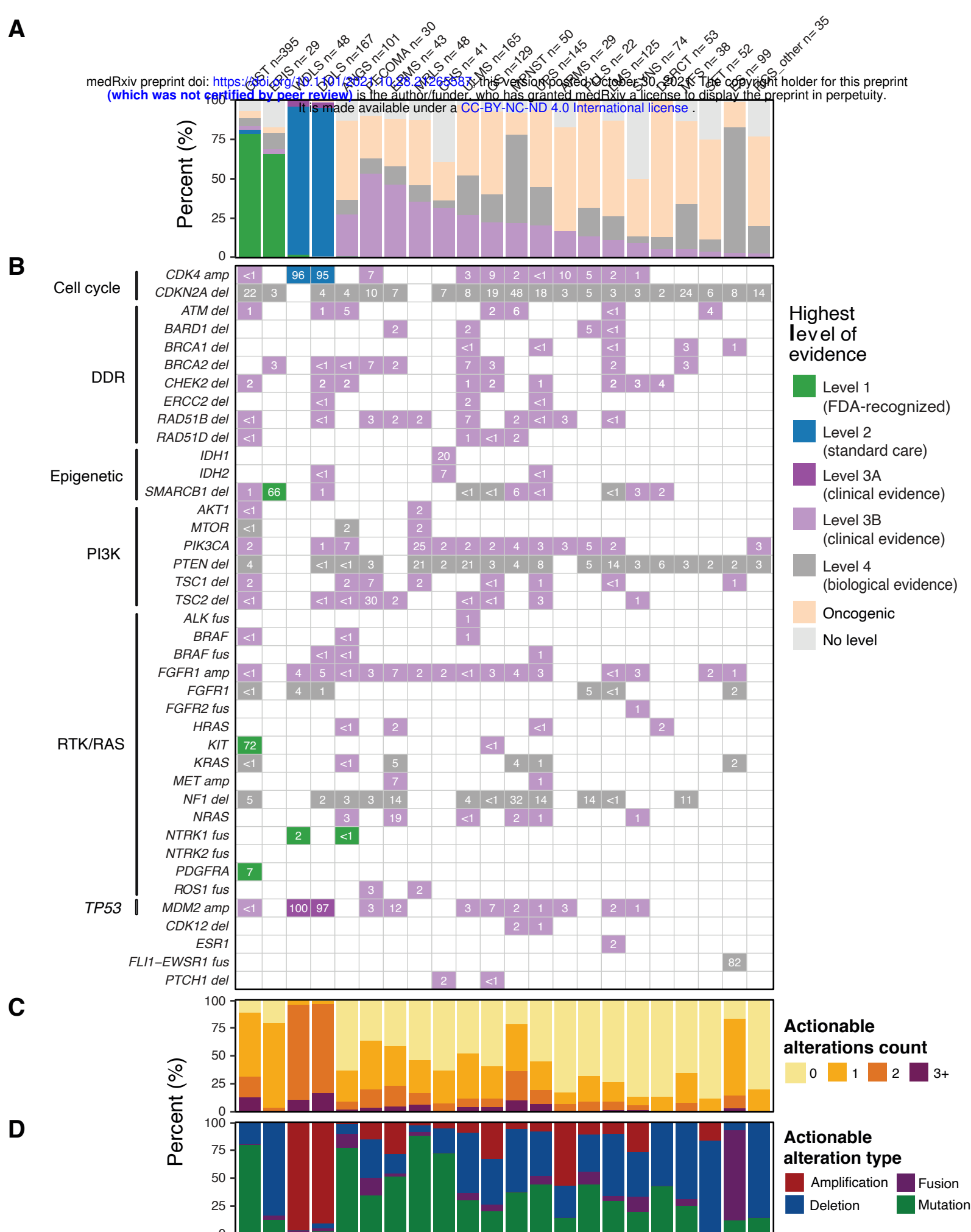
Figure 7. Actionability of mutations by subtype and gene. For each of the 22 most common subtypes: **A**, Frequency of actionable alterations by level of evidence. **B**, Actionable alterations in individual genes, grouped by pathway. Numbers in each cell represent the percentage of samples with actionable alterations in that gene. **C**, Number of actionable alterations per sample. **D**, Frequency of actionable alterations classified by alteration type.
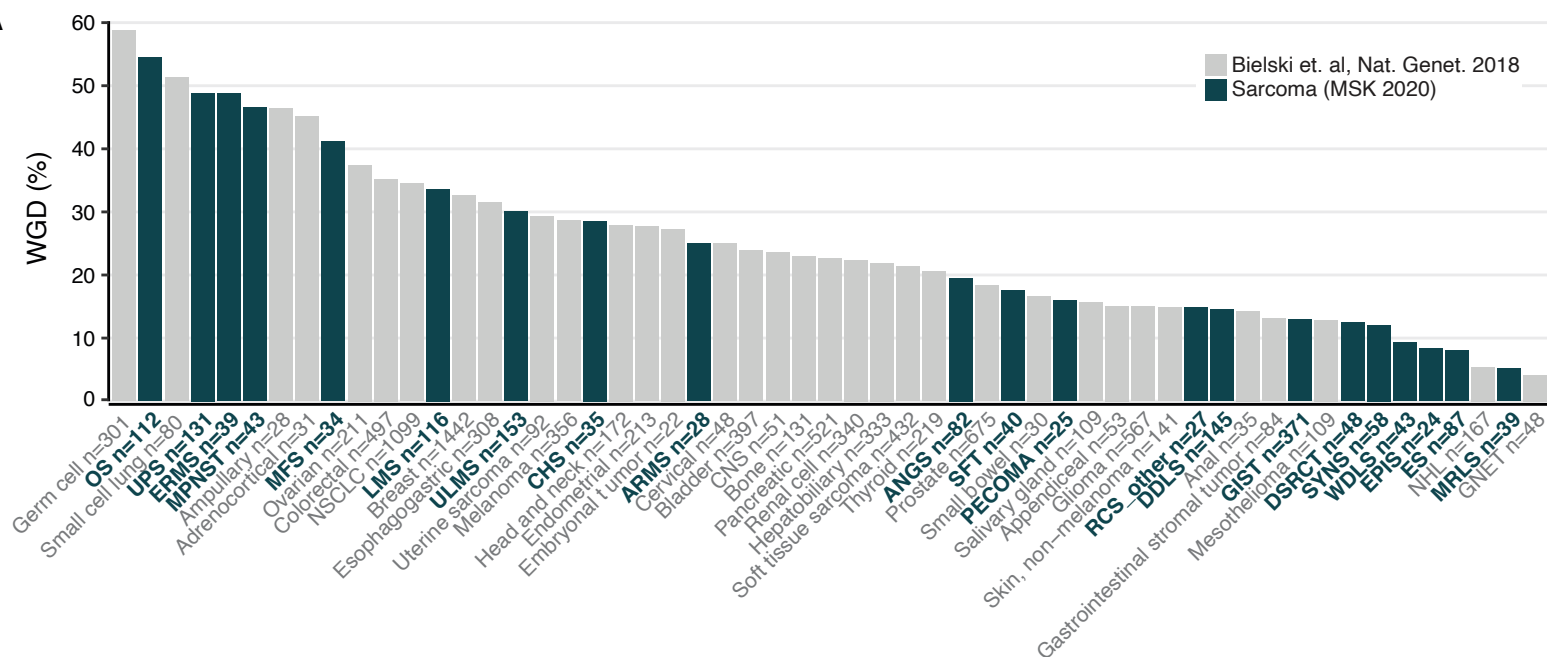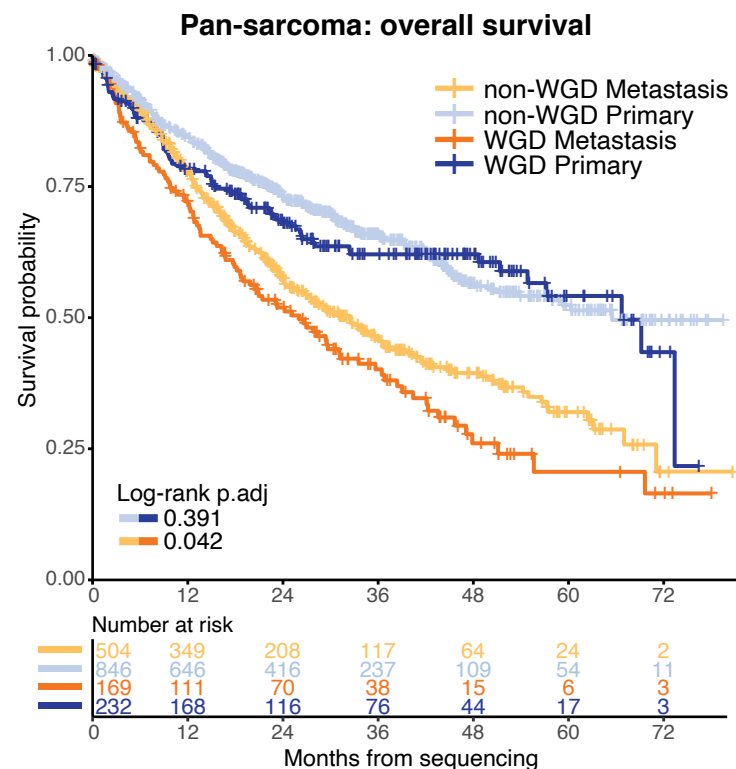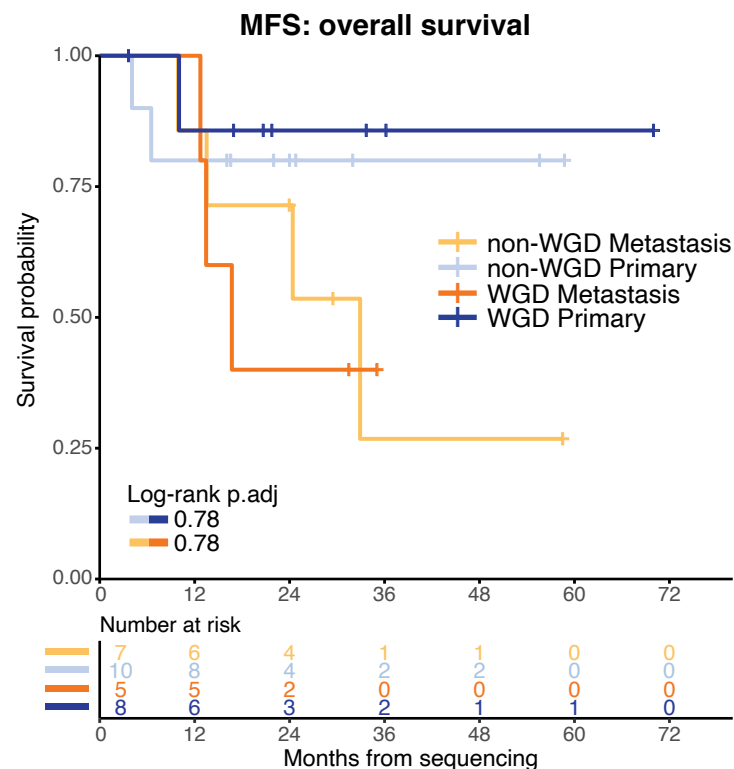
Supplementary Figure 1. Whole genome doubling (WGD) and survival probability. A, Frequency of WGD by subtype (green) compared to other cancers (all available samples) (gray). **B-C**, Overall survival based on WGD status within metastatic and primary tumor cohorts in **B**, all sarcoma subtypes; **C**, MFS.

Supplementary Figure 2. Oncogenic epigenetic pathway alterations. Frequency of oncogenic alterations in specific epigenetic pathway genes in each subtype with ≥ 10 samples. Top box, aggregate number of alterations in each gene family/biochemical process.
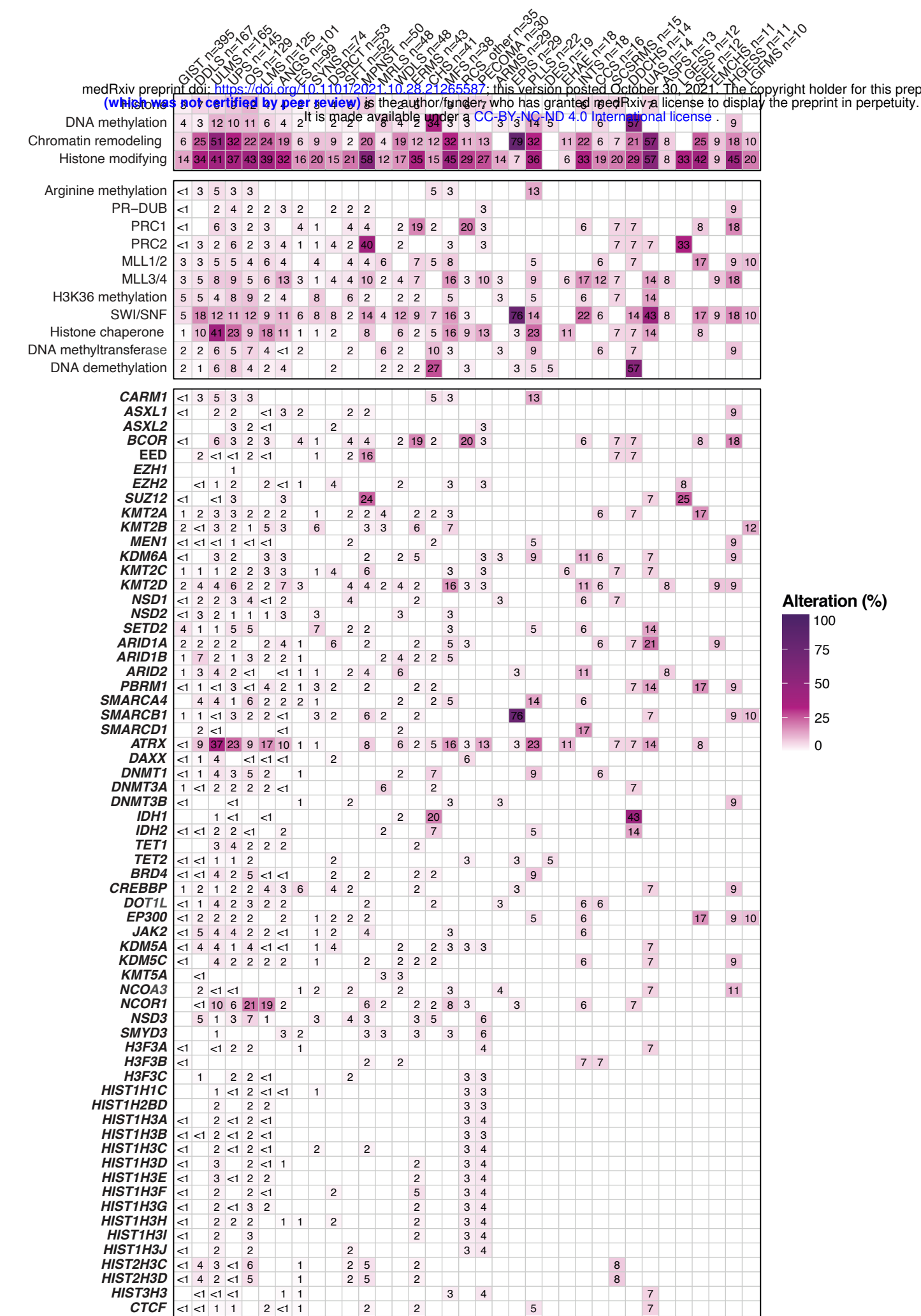
Supplementary Figure 3. All epigenetic pathway alterations, including variants of unknown significance (VUS).
Frequency of somatic alterations in epigenetic pathway genes in each subtype with ≥ 10 samples. Top boxes, aggregate number of alterations in each gene family and biochemical process.

Column subtypes (with sample numbers): GIST n=395, DDLS n=167, ULMS n=165, UPS n=145, OS n=129, LMS n=125, ANGS n=101, ES n=99, SYNS n=74, DSRCT n=53, SFT n=52, MPNST n=50, MRLS n=48, WDLS n=48, LERMS n=43, CHS n=41, MFS n=38, RCS other n=35, PECOMA n=30, ARMS n=29, EPIS n=29, PLLS n=22, DES n=19, EHAE n=19, INTS n=18, CCS n=18, SCSRMS n=16, DDCHS n=15, UAS n=14, ASPS n=14, LGESS n=13, SEF n=12, EMCHS n=11, HGESS n=11, LGFMS n=10

Biochemical process (aggregate) rows: DNA methylation, Chromatin remodeling, Histone modifying.

Gene family rows: Arginine methylation, PR–DUB, PRC1, PRC2, MLL1/2, MLL3/4, H3K36 methylation, SWI/SNF, Histone chaperone, DNA methyltransferase, DNA demethylation.

Gene rows: CARM1, ASXL1, ASXL2, BCOR, EED, EZH1, EZH2, SUZ12, KMT2A, KMT2B, MEN1, KDM6A, KMT2C, KMT2D, NSD1, NSD2, SETD2, ARID1A, ARID1B, ARID2, PBRM1, SMARCA4, SMARCB1, SMARCD1, ATRX, DAXX, DNMT1, DNMT3A, DNMT3B, IDH1, IDH2, TET1, TET2, BRD4, CREBBP, DOT1L, EP300, JAK2, KDM5A, KDM5C, KMT5A, NCOA3, NCOR1, NSD3, SMYD3, H3F3A, H3F3B, H3F3C, HIST1H1C, HIST1H2BD, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I, HIST1H3J, HIST2H3C, HIST2H3D, HIST3H3, CTCF.

Legend — Alteration (%): 0, 25, 50, 75, 100 (white to dark purple color scale).