SUPPLEMENTARY INFORMATION

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

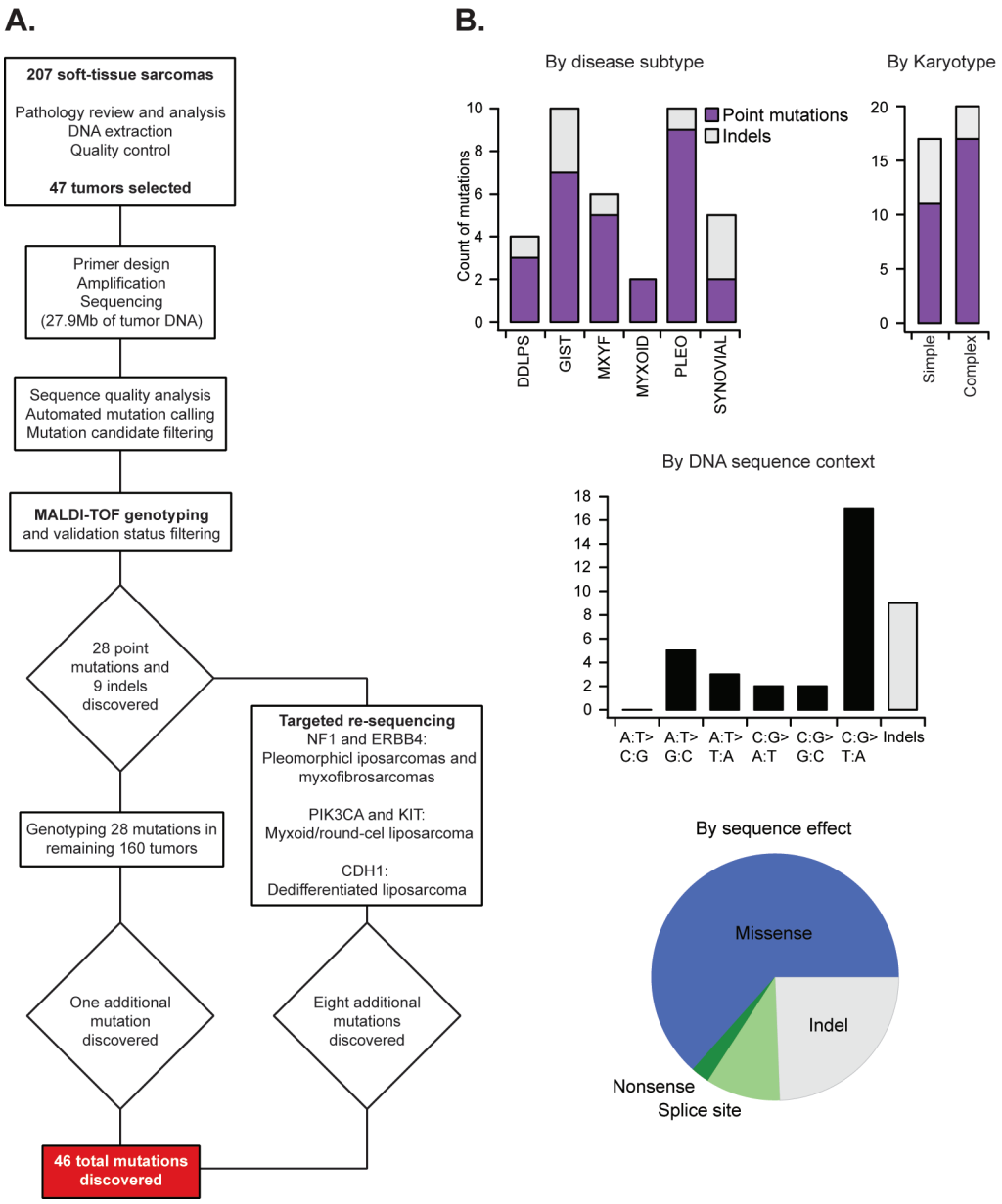
Jordi Barretina*, Barry S. Taylor*, Shantanu Banerji, Alexis H. Ramos, Mariana Lagos-Quintana, Penelope L. DeCarolis, Kinjal Shah, Nicholas D. Socci, Barbara A. Weir, Alan Ho, Derek Y. Chiang, Boris Reva, Craig Mermel, Gad Getz, Yevgenyi Antipin, Rameen Beroukhim, John E. Major, Charlie Hatton, Richard Nicoletti, Megan Hanna, Ted Sharpe, Tim Fennell, Kristian Cibulskis, Robert C. Onofrio, Tsuyoshi Saito, Neerav Shukla, Christopher Lau, Sven Nelander, Serena Silver, Carrie Sougnez, Agnes Viale, Wendy Winckler, Robert G. Maki, Levi A. Garraway, Alex Lash, Heidi Greulich, David Root, William R. Sellers, Gary K. Schwartz, Cristina R. Antonescu, Eric S. Lander, Harold E. Varmus, Marc Ladanyi, Chris Sander, Matthew Meyerson, Samuel Singer

* Contributed equally

CONTENTS

Nature Genetics: doi:10.1038/ng.619

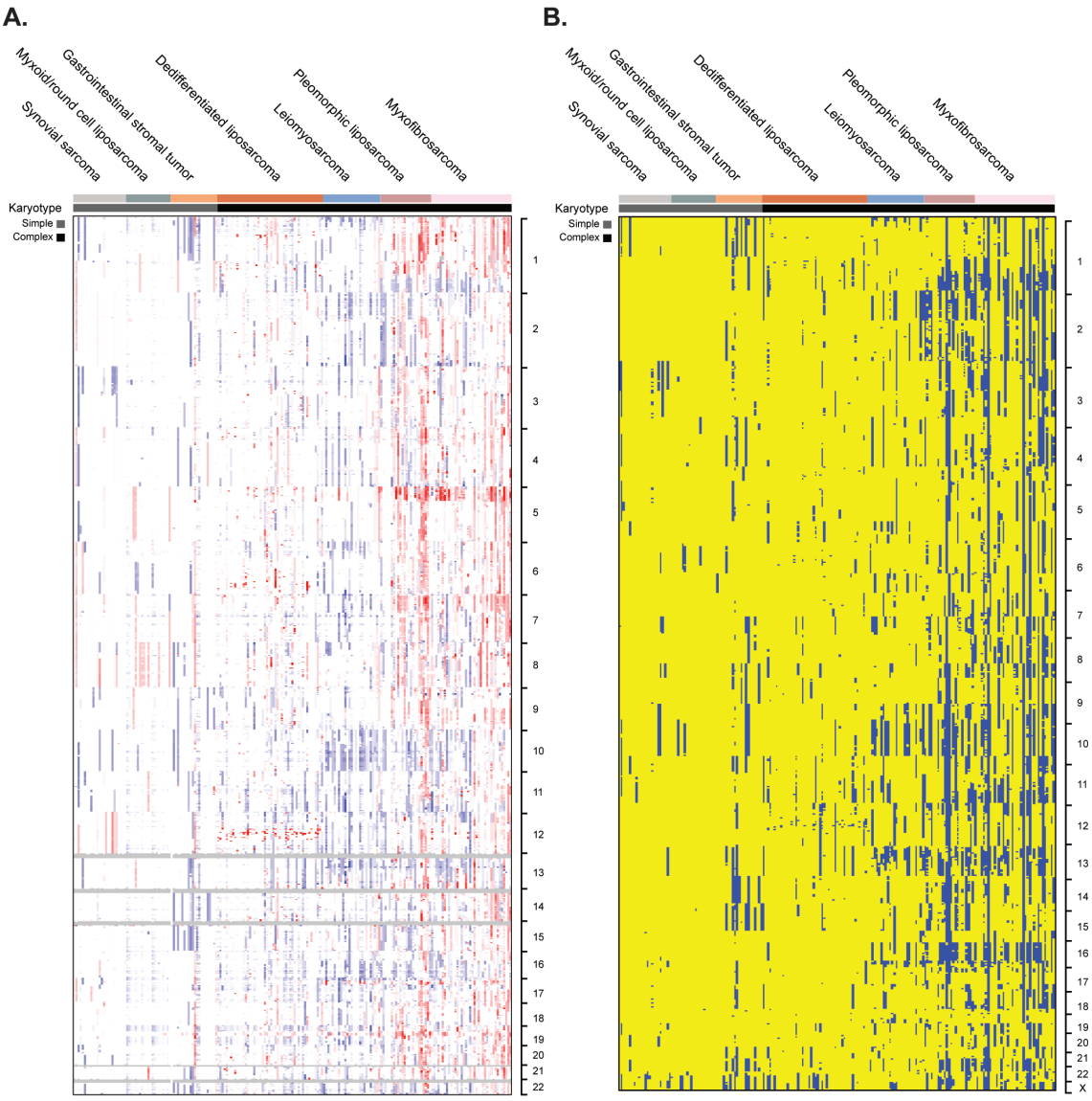
Supplementary Figure 1



Somatic mutation detection and results in soft tissue sarcoma. A.) This diagram schematizes the high-throughput re-sequencing workflow employed in this study. **B.)** The distribution of somatic mutations detected in soft tissue sarcomas is shown by subtype, karyotype, by DNA sequence change, and mutation type.

Nature Genetics: doi:10.1038/ng.619

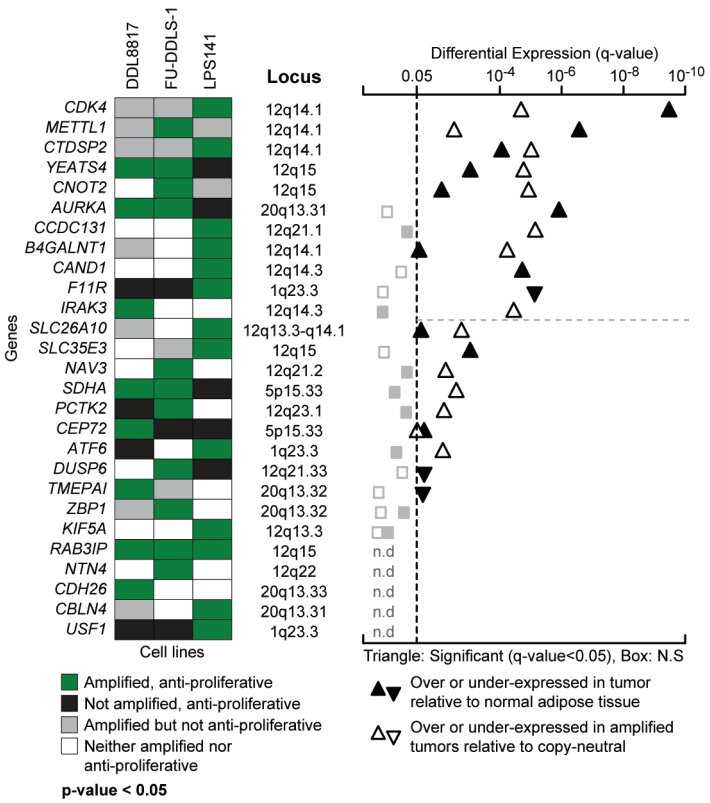
Supplementary Figure 2



Genome-wide copy-number alterations and loss-of-heterozygosity (LOH) in soft tissue sarcoma. A.) Plotted here is a heatmap representation of genome-wide copy-number changes in 208 soft tissue sarcomas (subtypes indicated at top, chromosomal position at right). Red indicates copy number gain/amplification, blue is loss/deletion, and white indicates copy-neutral regions. **B.)** Binary representation of genome-wide LOH with subtypes arranged as in panel A (chromosomal positions at left). Yellow indicates retention why blue indicates LOH.

Nature Genetics: doi:10.1038/ng.619

Supplementary Figure 3



RNAi identifies cancer proliferation genes in dedifferentiated liposarcoma. On the left, 27 genes whose knockdown impairs proliferation in at least one cell line in which they are also amplified (green; remaining statuses indicated at bottom). Genes are sorted by descending significance of differential expression in human tumors (right). Expression was compared in normal adipose tissue versus tumors (filled symbols), and in tumors in which the gene was amplified versus tumors where it was copy-neutral (open symbols) using matched copy number and expression data available for 46 of 60 samples in our study. The direction of differential expression is reflected in the direction of the triangle (up/down respectively). The vertical dashed line indicates the threshold of statistical significance; horizontal dashed line separates the two classes of genes as described in the Supplementary Note; gray boxes reflect sub-significant expression changes in either comparison; N.D. not determined.

Nature Genetics: doi:10.1038/ng.619

Supplementary Table 1: *See supplied file.*

Supplementary Table 2: *See supplied file.*

Supplementary Table 3: *See supplied file.*

Nature Genetics: doi:10.1038/ng.619

Supplementary Table 4

Chr	Start	End	Q-value*	Number of genes	Genetic element of interest
Dedifferentiated liposarcoma					
1	157245008	158910992	0.0114	42	*SDHC*
1	225664000	232960992	0.0034	46	
2	174611008	178686992	0.2431	29	
6	110391	1694929	0.1261	9	*IRF4*
6	129814000	131723000	0.1261	8	
9	68978504	69946096	0.2431	5	
11	106774000	111406000	0.0012	34	*ATM*
12	20662800	21191300	0.0114	5	*PDE3A*
12	2461291	4422489	0.0034	16	*CCND2*
13	18097312	51194100	0.1261	165	*FOXO1A,RB1,BRCA2,ELF1*
14	25382600	29912400	0.2431	3	
14	93906200	94941696	0.2431	12	*DICER1*
21	39953500	41000600	0.1261	4	
22	24821200	35620800	0.2431	103	*EWSR1,NF2*
Leiomyosarcoma					
4	7277771	24856400	0.2294	52	
8	189547	22618700	0.2294	119	
13	39605000	47593700	3.12E-10	42	*FOXO1A,ELF1,LCP1,CPB2*
19	61069600	63811652	0.0295	79	
22	19785600	25899400	0.2294	67	*BCR,SMARCB1*
Myxofibrosarcoma					
1	210503008	216664992	8.29E-06	17	*ESRRG,PTPN14,TGFB2*
2	231507008	234006000	0.0020	33	
3	38001700	40760300	0.1297	31	
4	136587008	143304992	0.2314	19	*ELF2*
7	146951008	151078992	0.2314	52	*EZH2,CDK5*
8	87515904	92016896	0.1297	14	
9	4049111	13055599	0.0238	29	*JAK2,PTPRD*
11	120307000	124503000	1.61E-04	51	
13	29923600	41013500	1.61E-04	48	*BRCA2,FOXO1A,LHFP,ELF1*
16	87804904	88827256	5.83E-04	22	*FANCA,CDK10*
17	4878631	6860979	1.61E-04	27	*USP6*
18	65638600	70586400	7.05E-03	14	*DOK6*
22	32604800	33596100	0.0619	1	
Pleomorphic liposarcoma					
1	212594000	218830992	3.37E-04	19	*TGFB2,ESRRG*
2	7972791	18514100	7.54E-04	40	*E2F6*
2	169535008	175860992	0.0088	45	
2	241792000	243018224	0.0088	15	
10	72210904	74517600	7.54E-04	19	
11	124609000	131730000	0.0088	40	
13	44402400	47545800	6.75E-04	19	*LCP1*
15	68571504	73858896	0.0088	63	*CSK,PML*
16	50273000	53651500	7.54E-04	8	

* GISTIC q-value<0.25 (False Discovery Rate < 25%)

Nature Genetics: doi:10.1038/ng.619

Supplementary Table 5: *See supplied file.*

Supplementary Table 6: *See supplied file.*

Nature Genetics: doi:10.1038/ng.619

Supplementary Note

Somatic mutations in soft-tissue sarcomas

Here, we describe in detail the results from our tumor re-sequencing for somatic mutations in coding exons and adjacent splice sites of 226 genes and DNA encoding 496 microRNAs (Supplementary Table 2) in six subtypes. Altogether, we found a total of 46 mutations involving 21 genes in 35 of the sequenced samples (Table 2). While we discovered mutations in subtypes with both simple and complex karyotypes, we observed no relationship between the frequency of mutation and the karyotypic complexity of the subtypes in which they occurred (Supplementary Figure 1B). The most prevalent transition was C:G>T:A, consistent with prior reports in other cancer types[1]. Most individual tumors had 2 or fewer mutations with one pleomorphic liposarcoma and a GIST having 3 and 4 mutations respectively. Mutation prevalence in soft tissue sarcomas (1.004 per Mb for point mutations) is similar to mutation rates in ovarian, colorectal, and renal carcinomas[1]. In total, twenty-four mutations affected protein and lipid kinase genes: 6 in *KIT*, 6 and 2 in the lipid kinases *PIK3CA* and *PI4K* respectively, 2 in the *EGFR* homolog *ERBB4* and single mutations in several members of the Ephrin receptor tyrosine kinase family (*EPHA1*, *EPHA5*, and *EPHA7*) and the tyrosine kinases *LTK*, *MST1R*, *NTRK1*, *PTK2B,* and *SYK* (Table 2). Some of these, while perhaps rare, may serve as therapeutically important targets in specific sarcoma subtypes.

Although our power to detect recurrently mutated genes was limited by the modest number of genes and samples sequenced, we nonetheless detected both mutations previously described in sarcoma and other cancers as well as 30 (65.2%) mutations not previously reported in cancer (see Catalog of Somatic Mutations in Cancer[2], release 37). Among known mutations, we confirmed *KIT* mutations (including imatinib-resistance mutation L576P) in GIST. Similar activating mutations occur in ~70% of GISTs and most can be successfully treated with imatinib[3]. The majority of *KIT* mutations (5/7) in our GIST samples were identified during exon sequencing, while only two were identified by subsequent genotyping (Supplementary Figure 1A). This is likely due to the criteria used to select samples for sequencing that required LOH on at least one chromosome arm to qualify a sample (see Methods). Because such LOH is a feature of advanced GISTs driven by *KIT* mutations, these 6 samples were biased towards those known to harbor mutations in *KIT*. Beyond *KIT* mutations found in GIST, we were surprised to identify K778N KIT in a myxoid/round-cell liposarcoma. This represents the first report of a *KIT* mutation in this subtype. The K778N mutation alters a conserved residue of the kinase domain, has not been previously reported, and is predicted to be highly functionally significant (Supplementary Table 3), possibly affecting interactions and resulting in increased KIT activity. In GIST, mutations in the kinase domain of *KIT* are less common than those in extracellular domains, and they tend to emerge in tumors resistant to imatinib[4]. Functional experiments are necessary to determine the effect of K778N on KIT activity in myxoid/round-cell liposarcoma.

Among other previously identified mutations in sarcoma, we also detected a D32V mutation in *CTNNB1* in synovial sarcomas. Other mutations found represent the first report of somatic mutations in the sarcoma subtypes studied here, including *CDH1* and *NF1* mutations. While we discuss *NF1* mutations in the primary text, the *CDH1* mutations were also intriguing. Loss of function of *CDH1*, encoding E-cadherin, is

Nature Genetics: doi:10.1038/ng.619

thought to contribute to proliferation, invasion, and metastasis[5]. We identified mutations N238D and A617T in a dedifferentiated liposarcoma and a GIST sample, respectively. The A617T mutation was previously detected in endometrial cancer[2]. Both mutations fall in regions encoding the extracellular cadherin domain, which mediates cell-cell adhesion. Furthermore, integrating mutations with copy number data (see description of copy number changes in dedifferentiated liposarcoma below) revealed that heterozygous deletions of the locus encoding *CDH1* (16q22.1) was common in multiple subtypes: dedifferentiated (24%), pleomorphic (37.5%), and myxoid/round-cell liposarcomas (14.3%), leiomyosarcomas (44%) and myxofibrosarcomas (23.7%) (Figure 1 and Supplementary Table 5). Indeed, the mutant dedifferentiated liposarcoma harbored heterozygous loss of the remaining allele, indicating biallelic inactivation. *CDH1* expression levels were low across all dedifferentiated liposarcoma samples in our dataset (data not shown), which may be a result of mutation, deletion, or even epigenetic inactivation that remains to be explored further. Although genomic loss/mutation was not associated with metastasis (for dedifferentiated liposarcoma p-value=0.06, one-sided Fisher's exact test), our analysis, despite being underpowered, suggests a role for CDH1 loss-of-function in invasive sarcomas.

Among the newly reported amino acid changes caused by mutations detected here, 18 mutations (excluding indels) were predicted to have a functional effect by Mutation Assessor analysis (see Methods and Supplementary Table 3)**.** Beyond the *PIK3CA* and *NF1* mutations described in the main text, and those described above, we also predicted functionally significant activation for a Y796H mutation in the tyrosine kinase domain of *EPHA5* and for two mutations in protein receptor ligand-binding domains: A212T in *EPHA1* and C520S in *ERBB4* (Supplementary Table 3). The two latter mutations affect evolutionarily conserved residues in extracellular ligand-binding domains. The likely consequence is the activation of receptor signaling through their destabilizing effect on dimerization or ligand interaction, something recently observed in glioblastoma[6]. In addition to the previously noted D32V *CTNNB1* mutation in synovial sarcoma, we noted *CTNNB1* was also mutated in a dedifferentiated liposarcoma sample. This suggests activation of the Wnt/ß-catenin pathway, previously described as anti-adipogenic, and consequent suppression of *PPARγ* and *C/EBPα*[7]. Other predicted functionally significant mutations include S300P in *MOS* and P410L in *MSTR1*, both of which may destabilize or otherwise affect protein interactions. To facilitate end-user's comprehensive analysis of the mutations described in this study in the context of their protein-family sequence alignments and residue placement in known or homology-deduced three-dimensional protein and complex structures, we refer readers to:
http://awabi.cbio.mskcc.org/oma/Report.jspx?mutanId=159

Integrated subtype-specific CNAs and gene expression analysis

In addition to mutation sequencing, we also characterized patterns of genomic aberrations in soft tissue sarcomas with both copy number and expression analysis. Tumor and matched normal DNA samples were analyzed with 250K SNP arrays for copy number alterations (CNAs: n=207 tumor and 205 matched normal) and LOH (n=200), as well as gene expression arrays (n=149 tumor samples) (see Methods, Supplementary Figure 2, Supplementary Table 5).

Nature Genetics: doi:10.1038/ng.619

As stated in the main text, we identified several recurrent CNAs, both novel and previously described, in soft tissue sarcomas and many genes in regions of altered copy number showed correspondingly altered mRNA transcript levels (Supplementary Table 5). Several of these CNAs are large and encompass many genes, including known oncogenes or tumor suppressor genes. A more detailed description of both common and histology-specific genomic alterations, and the potential target genes encoded by those regions, is presented below.

<u>*Dedifferentiated liposarcoma*</u>. In addition to the canonical amplifications of *MDM2, HMGA2*, and *CDK4*, usually in double minutes, ring chromosomes, and large marker chromosomes[8,9], the 12q13.2-q23.1 locus harbors complex rearrangements (Figure 4). On the basis of the rearrangements and correlated over-expression results, 12q may contain additional driver genes besides *MDM2, HMGA2*, and *CDK4*. These potentially include genes known to be involved in liposarcomagenesis, as well as genes not previously involved in this disease or in cancer. Examples are *NAV3*, *WIF1*, *MDM1*, *DYRK2*, *ELK3*, *DUSP6*, *YEATS4*, *TBK1*, and *FRS2*, which were amplified at frequencies ranging from ~14% to 80% of tumors (Supplementary Table 5).

The amplicon including *CDK4* (56.098-58.273mb) likely had its peak obscured due to reduced probe coverage and a high density of copy number/rearrangement breakpoints at this locus (nearest 3' region is amplified in 88% of the tumors). The regional boundaries of this amplicon contain 26 additional genes, 18 of which showed correlated changes in expression. In addition to *CDK4*, these include *METTL1, INHBE, GLI1, MARS, DDIT3, DCTN2, PIP5K2C, DTX3, SLC26A10, OS9, TSPAN31, CYP27B1, FAM119B, TSFM, AVIL, B4GALNT1,* and *CTDSP2*. Our copy number analysis of the *CDK4* region actually located the peak of amplification directly 3' of this commonly accepted target, alternatively indicating *CYP27B1*. Both *CYP27B1* and *CYP2J2* [amp(1p32.2)] encode members of the cytochrome P450 super-family of enzymes, whose over-expression may be related to the chemoresistance of dedifferentiated liposarcomas. *CYP27B1*, also amplified in osteosarcoma[10], glioma[11], lung cancer[12], and a small number of melanomas[13], is also involved in the synthesis of cholesterol, steroids, and other lipids, and its deregulation may be related to the profound dedifferentiation phenotype in this subtype.

The *HMGA2* locus (12q14.3, amplified in ~70% of dedifferentiated liposarcomas) is another complex amplicon. The gene is translocated in several cancers including benign lipomas and well-differentiated liposarcomas, both speculated to be precursors of dedifferentiated liposarcomas[14-16]. Notably, only the first three exons of *HMGA2* are significantly recurrently amplified (15 of 50 tumors bearing 5' amplification), and the resulting truncated protein is oncogenic[17]. This is one of several internal gene events in our data, another example of which was interstitial amplifications of *NAV3* at 12q21.2. We note with interest that chromosome 12 is frequently amplified in several cancers including malignant melanoma[13,18], lung adenocarcinoma[19], and glioma[20], suggesting this genomic region is not only prone to rearrangements that are likely under selection, but also that these regions encode genes selected for by a broad range of epithelial and mesenchymal tumor types.

Aside from 12q aberrations, we detected significant gains/amplifications of 1p, 1q, 5p, 6q, and 20q (Supplementary Figure 2A and Figure 1). The peak of amplification on chromosome 1p was centered on *DAB1* in RAE and *OMA1* in GISTIC, but the two

Nature Genetics: doi:10.1038/ng.619

analyses yielded very similar regional boundaries spanning seven additional genes, four of which had correlated expression changes. In fact, as seen in previous studies[21], we detected both *JUN* amplification (1p32) and up-regulated expression of its cognate transcript in this region (~24% in this series; correlation of copy number to transcript expression, p-value=$1.3x10^{-6}$ and R=0.646, Pearson's product moment correlation). While Mariani et al. demonstrated a role for oncogenic *JUN* in controlling adipocyte dedifferentiation through repression of C/EBPβ, the 24% frequency of amplification in our series couldn't explain the adipogenesis block in all tumors. Other genes, including *DDIT3*, exert a similar negative regulatory effect on C/EBPβ[22]. Critically, *DDIT3* is on 12p13.3 near *CDK4* and was also frequently amplified and over-expressed in dedifferentiated liposarcoma (p-value=$2.3x10^{-15}$ and R=0.874, Pearson's product moment correlation). *DDIT3* alterations were mutually exclusive with those affecting *JUN* (data not shown). It is therefore tempting to speculate that *DDIT3* serves as an alternative effector of deregulated adipogenesis in tumors bearing wild-type *JUN*.

On chromosome 5p, *TERT* amplification (5p13.33; ~16% of tumors) may be responsible for the high level of telomerase gene expression associated with alternative lengthening of telomeres (ALT), frequently observed in liposarcomas[23,24]. Apart from *TERT*, other genes amplified on 5p include *NDUFS6* and *CCT5*. *NDUFS6* is involved in mitochondrial oxidative phosphorylation and over-expressed in cervical cancer[25], while increased expression of *CCT5* is related to docetaxel resistance in breast cancer. The 5p amplification might represent another critical event perhaps partially responsible for the low chemotherapy responses of dedifferentiated liposarcomas.

Dedifferentiated liposarcoma had fewer robustly recurrent deletions than amplifications (Figure 1). Loss of 19p (~34%) was the most common arm-length event (Supplementary Table 5), while other genomic losses encompassed well-known tumor suppressors not yet implicated in liposarcoma including *MEN1, ATM, CDKN1B*, *WNT1*, *PTPN11*, *CDH1, TSC2*, *FANCA*, *TP53*, *SUZ12*, *NF1*, *BRCA1* (Supplementary Table 5). In addition, *RB1* and *BRCA2*, among other genes, were also affected by copy-neutral LOH (Supplementary Table 4). Finally, we also noted frequent deletion of genes encoding BCL-2 family pro-apoptotic proteins in dedifferentiated liposarcoma, as well as other subtypes, a finding also evident in a larger tumor set including multi-lineage carcinomas and hematologic malignancies[26].

GIST. The most common copy-number changes in GIST were deletions of 1p, 14q, and 15q, all appearing in ~32% of cases (Figure 1 and Supplementary Figure 2A, Supplementary Table 5). Potentially important aberrations included homozygous deletion of the 9p12.3 locus encoding *CDKN2A/CDKN2B,* and monoallelic loss of *LRP1B* (2q22.1-2), encoding the low-density lipoprotein-related protein 1B and affected by both whole-gene and intragenic deletions. Deletions and somatic mutations of *LRP1B* have also been identified in lung cancer[27].

Leiomyosarcoma. In this karyotypically complex subtype, genomic deletions were more common than amplifications, and the most prominent changes were chromosome 10 deletions (~50-70% of cases) (Figure 1, Supplementary Table 5). Indeed, a recently developed murine model recapitulates leiomyosarcoma by genetically inactivating Pten (human 10q23.21) in smooth muscle cells[28], suggesting 10q loss occurs early in leiomyosarcomagenesis. In addition to PTEN inactivation, we identified homozygous deletions in *FRAP1,* encoding mTOR. Because PTEN is a repressor of Akt, both these

Nature Genetics: doi:10.1038/ng.619

events suggest a role for aberrant Akt-mTOR signaling in leiomyosarcoma. As with other subtypes, the deletions included well-characterized tumor suppressor genes like *TP53*, *BRCA2*, *RB1*, and *FANCA*.

Myxofibrosarcoma. The most common copy-number gain/amplification in this subtype (~55% of tumors) affected chromosome 5p (Figure 1, Supplementary Table 5). This region contains *RICTOR* (the rapamycin-insensitive binding partner of mTOR)[29], *CDH9*, and *LIFR*. Other amplified regions included several discontinuous loci on 1p and 1q spanning *PIK4CB*, *ETV3*, and *MCL1* among others. *MCL1,* an anti-apoptotic gene[30], was also concomitantly over-expressed in these tumors. Myxofibrosarcomas also harbored deletions of classic tumor suppressors including *CDKN2A*/*CDKN2B*, *RB1*, and *TP53*. These events, in combination with the inactivating mutations we detected in *PTEN* and *NF1* in this subtype, demonstrated extensive loss of function in several known tumor suppressors. Other tumor suppressors were implicated by significant copy-neutral LOH; these include *JAK2*, *PTPRD*, and *DOK6* (Supplementary Table 4).

Myxoid/round-cell liposarcoma. In this translocation-driven liposarcoma, the most common copy-number aberrations were gains of 8p and q (~33% of tumors), and losses of 19p and 19q (38% and 14% respectively) (Figure 1, Supplementary Table 5). The chromosome 8 gains included *MYC* and other well-known oncogenes, but also over-expressed genes that have not been previously involved in cancer, such as *JRK*. We note that a subset of samples were hypersegmented, i.e. the profile contained many short adjacent segments that differed in copy number, and the count of copy number segments was an outlier among all samples segmented. The cause of this phenomenon has been described previously[31]. This, in combination with the low count of available tumors (n=21) due to their rarity, affected our power to detect significant alterations.

Synovial sarcoma. The most common copy-number alterations in this subtype were 3p deletion and 12q trisomy (24% and 20% of cases, respectively) (Figure 1 and Supplementary Table 5). Large-scale low-amplitude gains of 8q and 12p (spanning ~14 and 8mb of average sequence respectively) were also statistically significant, as well as deletions of both *PDE4B* and *PTPRD*, the latter frequently lost in lung cancer[19] and glioblastoma[32]. As in myxoid/round-cell liposarcoma, altered regions on the q-arm of chromosome 8 included *MYC*, *JRK*, and the Ser/Thr protein kinase *SGK3*, encoding a PDK1 target, recently involved in AKT-independent signaling downstream of oncogenic *PIK3CA* mutations[33]. A similar pattern of large-scale gains on 12p affected *KRAS*, *CCND2,* and *ETV6*, but also *GAPDH*, which has demonstrated a pro-survival effect during apoptosis[34].

Pleomorphic liposarcoma. The most significant genomic alterations in this karyotypically complex subtype have been extensively described[31]. Integration with the gene expression results of this study confirmed many of the alterations described, and refined others. For gains/amplifications, multiple genes of interest demonstrated correlated expression including those on 19q13.12 (*ETV2, MLL4, PSENEN*), 5p13.3-p12 (*AMACR, RAD1, SKP2*), 5p15.1-p13.3 (*RNASEN* encoding Drosha), and 5p15.2 (*DAP*). A similar pattern arose for genomic deletions, including those on 1q41-q42.12 (*MARK1, TP53BP2*), 12p13.33 (*RAD52, WNT5B, JARID1A*), 13q14.2 (*RB1*), 17p31.1 (*TP53*), 17q11.2 (*NF1*), and 22q13.2 (*ST13*), among others.

Nature Genetics: doi:10.1038/ng.619

Functional genetic screen of amplified genes in dedifferentiated liposarcoma

Here, we describe in more detail the results from our loss-of-function RNAi screen against significantly amplified genes in dedifferentiated liposarcoma. To identify those genes whose knockdown caused significant reduction in cell proliferation in each of the three cell lines screened here, we exploited the redundancy of our shRNA library (median of 5 hairpins per gene; range 1-19). Specifically, we used the probabilistic methodology RSA (redundant siRNA activity) to aggregate the evidence of individual hairpin/well activities targeting that gene (described in main text)[31]. RSA was run once per cell line. Parameter values were -1 and -0.5 for the lower and upper bounds respectively, defining a Z-score below which a hairpin is considered to suppress proliferation and at which no anti-proliferative affect is observed. An anti-proliferative affect was defined as genes with a nominal p-value < 0.05 in any cell line.

Our analysis revealed 99 genes (25.7% of those screened, Supplementary Table 6) whose RNA knockdown significantly inhibited proliferation in at least one representative cell line. When combined with cell line copy number data, two groups of genes emerged: (i) genotype-dependent essential genes: those amplified in at least one cell line in which they were essential (n=27) (Supplementary Figure 3), and (ii) genes amplified in some human tumors, but diploid in cell lines in which they were essential (n=72). The latter group may represent genes essential for the survival of all cells independent of amplification status, or suggest alternative mechanisms of oncogene activation beyond gene amplification.

Ranking the genotype-dependent genes by differential mRNA expression (both normal adipose tissue versus tumor samples and amplified versus non-amplified tumor samples) again produced two classes (Supplementary Figure 3, right panel): genes whose amplification correlated with over-expression, as is believed to be typical of oncogenes[35], and genes that lacked substantial differential expression. The latter were nevertheless required for cancer cell proliferation, therefore, non-overexpressed genes within amplicons may expand our notion of the mechanisms through which amplified genes may exert functionally important effects[36]. Of the 27 genotype-essential genes, we noted enrichment for genes on 1q, 5p, 12q (q14 and q15) and 20q (Figure 4A). Among them, the kinases *CDK4*, *IRAK3*, *PCTK2*, *AURKA,* together with *NAV3*, have all been found mutated in cancer[2]. Additionally, *CDK4*, *PCTK2*, and *DUSP6* are all regulators of the cell cycle. *AURKA* at 20q13.31 has been shown to participate in the regulation of the mitotic spindle during mitosis and is over-expressed in a number of cancers[37]. Two of the three cell lines used in the arrayed screen showed genotype-specific essentiality for *AURKA* while the third line was still sensitive despite no amplification being present. Recently developed *AURKA*-specific inhibitors may therefore have an important role in the treatment of dedifferentiated liposarcoma, and this deserves further investigation[38]. Of the genes whose amplification correlated with over-expression, 4 of 11 were found in the 12q14.1 amplicon (*CDK4*, *METTL1*, *CTDSP2*, and *B4GALNT1*). As described in the main text, prolonged CDK4 RNA knockdown led to decreased proliferation in cell lines with amplification of this gene. *METTL1* and *CTDSP2* have also been found to be co-amplified with *CDK4* in lung cancer and glioma respectively[12,20]. The results of our screen might point to a yet unknown cooperation between multiple genes within this amplicon. Taken together, our data suggest that several genes in addition to *CDK4* may contribute to the disease phenotype when amplified and may be candidate therapeutic targets.

Nature Genetics: doi:10.1038/ng.619

The methodological approach used here, while an efficient avenue for cancer gene discovery when informed by genomic data, is limited in several important ways. Like similar screens with technical and biological noise, it may produce false negatives; genes whose knockdown may otherwise be cell-lethal but were not observed as such in our screen. For instance, as noted in the main text, the limited duration of the proliferation analysis in the high-throughput screen may have caused us to miss certain essential genes. Also, and of particular interest in dedifferentiated liposarcoma, a subset of hairpins in the library used here target the wild-type region of genes with intragenic structural mutations that amplify only a fraction of the coding locus. These hairpins will not produce a significant genotype-dependent effect. Another possible source of false negatives is high-level amplification, a common feature of 12q genes, which may raise gene expression to a level that overwhelms the effects of the hairpins. Finally, additional genes are certainly over-expressed through reversible mechanisms, such as epigenetic, and therefore not included in our screening set.

Another issue, mentioned above, is the subset of genes amplified but apparently not essential for proliferation. We note that our loss-of-function screen tested only a single phenotype of proliferation and therefore it is likely that a subset of these amplified genes are functionally significant, but contribute to other malignant phenotypes that require alternative assays. Finally, genes neither amplified nor lethal in our screen indicate the importance of expanding the set of liposarcoma cell lines to more fully recapitulate the fine-scale structure of the tumor genotype.

Nature Genetics: doi:10.1038/ng.619

Supplementary References

1. Greenman, C. et al. Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153-8 (2007).
2. Bamford, S. et al. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* **91**, 355-8 (2004).
3. Heinrich, M.C. et al. Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. *J Clin Oncol* **21**, 4342-9 (2003).
4. Fletcher, J.A. & Rubin, B.P. KIT mutations in GIST. *Curr Opin Genet Dev* **17**, 3-7 (2007).
5. Strathdee, G. Epigenetic versus genetic alterations in the inactivation of E-cadherin. *Semin Cancer Biol* **12**, 373-9 (2002).
6. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-8 (2008).
7. Rosen, E.D. & MacDougald, O.A. Adipocyte differentiation from the inside out. *Nat Rev Mol Cell Biol* **7**, 885-96 (2006).
8. Heidenblad, M. et al. Genomic profiling of bone and soft tissue tumors with supernumerary ring chromosomes using tiling resolution bacterial artificial chromosome microarrays. *Oncogene* **25**, 7106-16 (2006).
9. Pedeutour, F. et al. Structure of the supernumerary ring and giant rod chromosomes in adipose tissue tumors. *Genes Chromosomes Cancer* **24**, 30-41 (1999).
10. Atiye, J. et al. Gene amplifications in osteosarcoma-CGH microarray analysis. *Genes Chromosomes Cancer* **42**, 158-63 (2005).
11. Maas, R.M. et al. Amplification and expression of splice variants of the gene encoding the P450 cytochrome 25-hydroxyvitamin D(3) 1,alpha-hydroxylase (CYP 27B1) in human malignant glioma. *Clin Cancer Res* **7**, 868-75 (2001).
12. Wikman, H. et al. CDK4 is a probable target gene in a novel amplicon at 12q13.3-q14.1 in lung cancer. *Genes Chromosomes Cancer* **42**, 193-9 (2005).
13. Muthusamy, V. et al. Amplification of CDK4 and MDM2 in malignant melanoma. *Genes Chromosomes Cancer* **45**, 447-54 (2006).
14. Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer. *Nat Rev Cancer* **7**, 899-910 (2007).
15. Italiano, A. et al. Gains and complex rearrangements of the 12q13-15 chromosomal region in ordinary lipomas: the "missing link" between lipomas and liposarcomas? *Int J Cancer* **121**, 308-15 (2007).
16. Narita, M. et al. A novel role for high-mobility group a proteins in cellular senescence and heterochromatin formation. *Cell* **126**, 503-14 (2006).
17. Fedele, M. et al. Truncated and chimeric HMGI-C genes induce neoplastic transformation of NIH3T3 murine fibroblasts. *Oncogene* **17**, 413-8 (1998).
18. Lin, W.M. et al. Modeling genomic diversity and tumor dependency in malignant melanoma. *Cancer Res* **68**, 664-73 (2008).
19. Weir, B.A. et al. Characterizing the cancer genome in lung adenocarcinoma. *Nature* **450**, 893-8 (2007).
20. Fischer, U. et al. A different view on DNA amplifications indicates frequent, highly complex, and stable amplicons on 12q13-21 in glioma. *Mol Cancer Res* **6**, 576-84 (2008).

Nature Genetics: doi:10.1038/ng.619

21. Mariani, O. et al. JUN oncogene amplification and overexpression block adipocytic differentiation in highly aggressive sarcomas. *Cancer Cell* **11**, 361-74 (2007).
22. Fawcett, T.W., Eastman, H.B., Martindale, J.L. & Holbrook, N.J. Physical and functional association between GADD153 and CCAAT/enhancer-binding protein beta during cellular stress. *J Biol Chem* **271**, 14285-9 (1996).
23. Cairney, C.J., Hoare, S.F., Daidone, M.G., Zaffaroni, N. & Keith, W.N. High level of telomerase RNA gene expression is associated with chromatin modification, the ALT phenotype and poor prognosis in liposarcoma. *Br J Cancer* **98**, 1467-74 (2008).
24. Johnson, J.E. et al. Multiple mechanisms of telomere maintenance exist in liposarcomas. *Clin Cancer Res* **11**, 5347-55 (2005).
25. Scotto, L. et al. Integrative genomics analysis of chromosome 5p gain in cervical cancer reveals target over-expressed genes, including Drosha. *Mol Cancer* **7**, 58 (2008).
26. Beroukhim, R. et al. The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899-905 (2010).
27. Ding, L. et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069-75 (2008).
28. Hernando, E. et al. The AKT-mTOR pathway plays a critical role in the development of leiomyosarcomas. *Nat Med* **13**, 748-53 (2007).
29. Sarbassov, D.D., Guertin, D.A., Ali, S.M. & Sabatini, D.M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098-101 (2005).
30. Warr, M.R. & Shore, G.C. Unique biology of Mcl-1: therapeutic opportunities in cancer. *Curr Mol Med* **8**, 138-47 (2008).
31. Taylor, B.S. et al. Functional copy-number alterations in cancer. *PLoS ONE* **3**, e3179 (2008).
32. Veeriah, S. et al. The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated and mutated in glioblastoma and other human cancers. *Proc Natl Acad Sci U S A* **106**, 9435-40 (2009).
33. Vasudevan, K.M. et al. AKT-independent signaling downstream of oncogenic PIK3CA mutations in human cancer. *Cancer Cell* **16**, 21-32 (2009).
34. Colell, A. et al. GAPDH and autophagy preserve survival after apoptotic cytochrome c release in the absence of caspase activation. *Cell* **129**, 983-97 (2007).
35. Albertson, D.G. Gene amplification in cancer. *Trends Genet* **22**, 447-55 (2006).
36. Solimini, N.L., Luo, J. & Elledge, S.J. Non-oncogene addiction and the stress phenotype of cancer cells. *Cell* **130**, 986-8 (2007).
37. Marumoto, T., Zhang, D. & Saya, H. Aurora-A - a guardian of poles. *Nat Rev Cancer* **5**, 42-50 (2005).
38. Manfredi, M.G. et al. Antitumor activity of MLN8054, an orally active small-molecule inhibitor of Aurora A kinase. *Proc Natl Acad Sci U S A* **104**, 4106-11 (2007).

Nature Genetics: doi:10.1038/ng.619