A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers

Ryan S. Lee,[1,2] Chip Stewart,[3] Scott L. Carter,[3] Lauren Ambrogio,[3] Kristian Cibulskis,[3] Carrie Sougnez,[3] Michael S. Lawrence,[3] Daniel Auclair,[3] Jaume Mora,[4] Todd R. Golub,[1,2,3,5] Jaclyn A. Biegel,[6,7] Gad Getz,[3] and Charles W.M. Roberts[1,2,8]

[1]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [2]Harvard Medical School, Boston, Massachusetts, USA. [3]Broad Institute, Cambridge, Massachusetts, USA. [4]Department of Pediatric Oncology, Hospital Sant Joan de Déu, Barcelona, Spain. [5]Howard Hughes Medical Institute, Chevy Chase, Maryland, USA. [6]Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA. [7]Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. [8]Division of Hematology/Oncology, Children's Hospital Boston, Boston, Massachusetts, USA.

Cancer is principally considered a genetic disease, and numerous mutations are thought essential to drive its growth. However, the existence of genomically stable cancers and the emergence of mutations in genes that encode chromatin remodelers raise the possibility that perturbation of chromatin structure and epigenetic regulation are capable of driving cancer formation. Here we sequenced the exomes of 35 rhabdoid tumors, highly aggressive cancers of early childhood characterized by biallelic loss of *SMARCB1*, a subunit of the SWI/SNF chromatin remodeling complex. We identified an extremely low rate of mutation, with loss of *SMARCB1* being essentially the sole recurrent event. Indeed, in 2 of the cancers there were no other identified mutations. Our results demonstrate that high mutation rates are dispensable for the genesis of cancers driven by mutation of a chromatin remodeling complex. Consequently, cancer can be a remarkably genetically simple disease.

Introduction

Cancer is generally considered to arise due to DNA mutations that alter the function of numerous genes (1). Indeed, most cancer genomes are aneuploid, contain amplifications and deletions, and typically have hundreds to thousands of DNA point mutations. However, not all cancers are necessarily so complex as some highly aggressive tumors are diploid. Large-scale sequencing projects have revealed that mutation rates display 10- to 100-fold differences among cancer types and even among different cancers of the same type (2, 3). Even in cancer types possessing the highest mutation rates, there are often individual cases that contain many fewer mutations. This raises a fundamental question about the nature of cancer: How simple can the genomes of highly aggressive cancers be?

Rhabdoid tumors (RTs) are a useful type of cancer with which to address these questions. These extremely aggressive pediatric cancers of the brain, kidney, and soft tissues are highly malignant, locally invasive, frequently metastatic, and particularly lethal (4), and yet, they are typically diploid and lack genomic aberrations detectable by SNP arrays (5). Early-onset cancers also offer the opportunity to examine cancer genomes prior to the acquisition of substantial numbers of age-related mutations, the vast majority of which are likely passenger mutations.

Near-uniform biallelic inactivating mutations in *SMARCB1* (also known as *SNF5*, *INI1*, and *BAF47*), a gene that encodes a core subunit of the SWI/SNF chromatin remodeling complex (6, 7), are a hallmark of RTs. Alterations in genes involved in chromatin remodeling, and particularly in genes encoding SWI/SNF subunits (8), are increasingly being identified in a wide variety of cancers, raising the possibility that epigenetic dysregulation may be a central mechanism of oncogenesis.

Results and Discussion

We obtained DNA from 32 diagnostic pretreatment *SMARCB1* mutant RT samples, of which 20 were from brain, 3 were from kidney, and 9 were from other soft tissues (Supplemental Table 1; supplemental material available online with this article; doi:10.1172/JCI64400DS1). The median age of the patients was 12 months. Matched non-tumor peripheral blood DNA was also obtained from each patient. Whole-exome sequencing and SNP array analysis was performed on all 32 sample pairs. Analysis of SNP arrays identified a single region with significant focal somatic copy number alterations (SCNAs): deletions at 22q11.23 that contained the *SMARCB1* gene were identified in 25 out of the 32 samples (GISTIC2.0, ref. 9; $q < 10^{-50}$), which comprised focal deletions in 16 cases, monosomy 22 in 15 cases, and both in 6 cases (Figure 1). One sample (08-262A) had a germ line focal deletion. Tumor purity ranged from 43% to 97%, so the lack of additional detected SCNAs was not likely due to stromal contamination (Supplemental Table 2).

We next performed exome sequencing of DNA to a mean coverage of 83-fold across 32.6 Mb of targeted coding regions for each sample (Supplemental Table 3). This level of coverage resulted in a "call-able" exome of 28.6 Mb. Detection of SCNAs by sequencing data was consistent with SNP array findings (Supplemental Figure 1). Analysis revealed a total of 172 somatic substitutions and insertions/deletions (indels) in the 32 tumors (Table 1 and Supplemental Table 12). Other than SMARCB1 loss, 2 tumors (08-114 and 09-223) had no detectable mutations, and 4 tumors (07-057, 07-221, 08-172, and 09-131) had only subclonal mutations (Figure 2A). The mean mutation rate was 0.19 mutations per Mb, with a minimum of 0 and a maximum of 0.45 mutations per Mb. This rate is, to our knowledge, the lowest of all cancers sequenced to date, particularly for such a high-grade and lethal type of cancer (Figure 2B). Consistent with our tumor selection process, all tumors had combinations of *SMARCB1* mutations and/or deletions predicted to cause homozygous loss of function (Supplemental Table 4 and Supplemental Figures 3 and 6). Over-
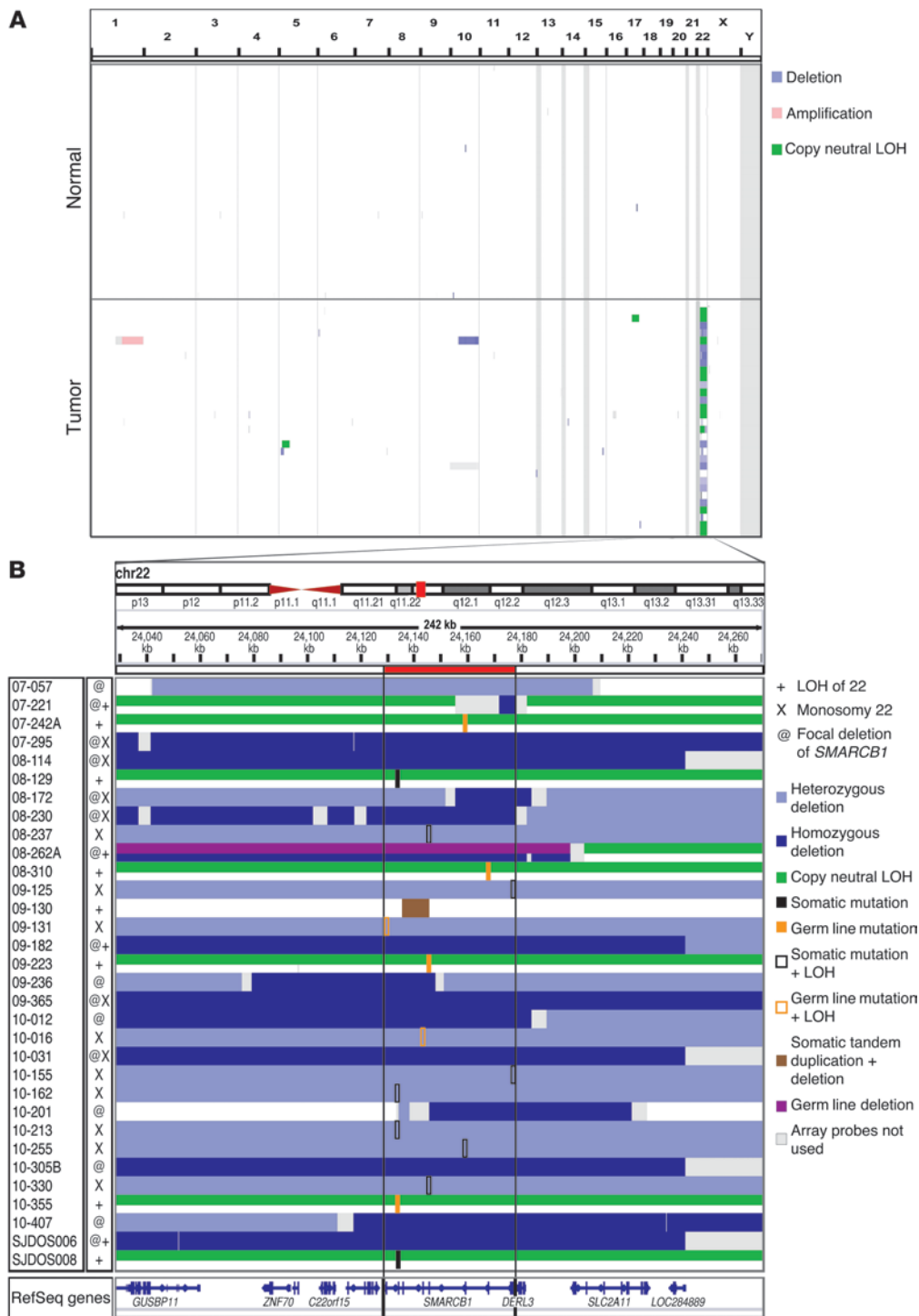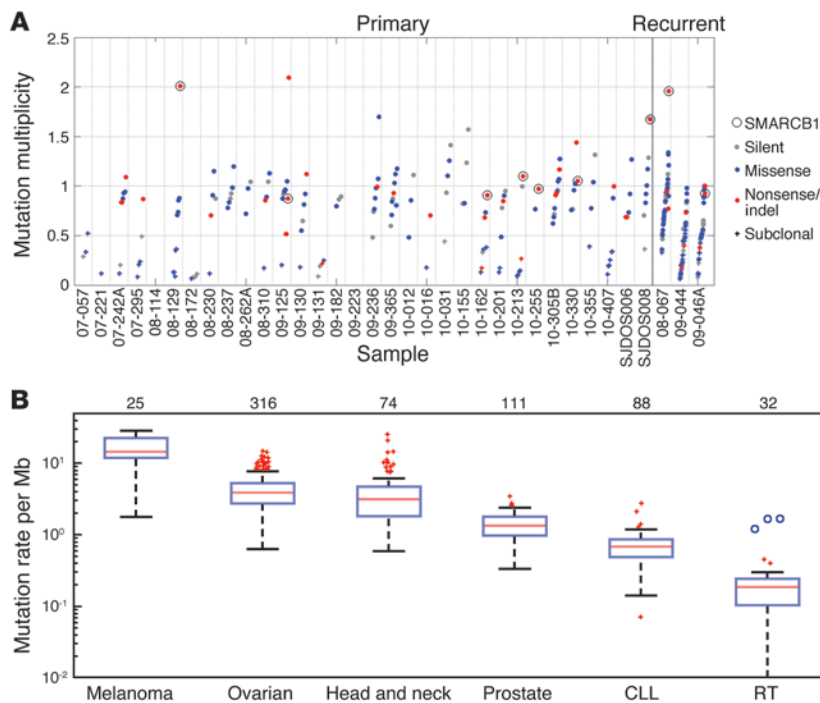
Figure 1
SNP arrays of primary RT samples and matched normal DNA. (**A**) Genome display of copy number changes. (**B**) Enlarged view of the *SMARCB1* locus. Mutations are overlaid on the SCNAs and loss-of-heterozygosity (LOH) tracks. Samples with focal deletions covering *SMARCB1* are marked with "@." Samples with monosomy 22 or loss of heterozygosity across 22 are marked with "x" or "+" next to the sample label. The red box represents the highlighted region of the chromosome, including the *SMARCB1* locus shown below. Red triangles represent the centromeric regions of the chromosome. Chr, chromosome.

all, 71.5% of the mutations were classified as clonal. All 7 of the point mutations in *SMARCB1* were classified as clonal (Figure 2A and Supplemental Table 5).

We looked for recurrent mutations that may cooperate with *SMARCB1* loss to drive RTs. The only other recurrently mutated gene was *GABRB2*, a subunit of the GABA A receptor, which was found to be clonally mutated in 2 out of the 32 samples (10-330 and SJDOS006; Supplemental Figure 4). However, the Catalog

of Somatic Mutations in Cancer (COSMIC v51) database contains only 2 other instances of this gene being mutated, neither of which matched the RT mutations (10, 11). Only 2 mutations found among the 32 sequenced RTs, aside from *SMARCB1* and *GABRB2*, were present in the COSMIC database: 1 in *NF2* (10-213) and the other in *TP53* (10-330). The *NF2* nonsense mutation (Y144*) may have been contributory; although subclonal (26% of cells), the mutation was present on a background of hemizygous deletion, implying

Figure 2
Somatic mutations in RTs. (**A**) Mutation multiplicity for each sample. Multiplicity is a measure of the average number of alternate alleles per tumor cell for each mutation. Heterozygous clonal mutations have a multiplicity near 1, while events below 1 are subclonal. Multiplicities close to 2 tend to be the result of mutations in loss-of-heterozygosity regions. Circles indicate the 9 *SMARCB1* mutations. (**B**) Logarithmic plot of mutation rates in 5 other types of cancer compared with those in RTs. Blue circles represent recurrent the RT samples. For box-and-whisker plots, red horizontal bars indicate medians, boxes indicate 25th and 75th percentiles, lower whiskers indicate lowest datum within 1.5 times the interquartile range (1.5xIQR) of the lower quartile, upper whiskers indicate highest datum within 1.5xIQR of the upper quartile, and red dots represent outliers. CLL, chronic lymphocytic leukemia.

homozygous loss in the mutant subclone. The relevance of the *TP53* mutation (D49N) was unclear; although clonal, the variant did not occur in one of the canonical mutation domains and was predicted not to be detrimental.

In addition to the 32 primary samples, we analyzed 3 independent recurrent tumor/normal pairs after chemotherapy (Supplemental Table 6). Sample 09-044 was found to be aneuploid (Figure 3A). The range of purity and coverage was comparable to that in the primary samples (Supplemental Tables 7 and 8). While the largest number of mutations per sample found in the 32 primary tumors was 13, the recurrent cancers contained 38, 47, and 47 mutations, resulting in a rate of 1.53 mutations per Mb (Supplemental Table 9), 8 times higher than that in the primary tumors (Figure 3B; $P < 0.005$). Other than *SMARCB1*, which was homozygously lost in all 3 recurrent samples (Supplemental Table 10), none of the mutations matched any in the COSMIC database and no gene contained recurrent, clonal mutations (Supplemental Figure 5). Notably, the treated samples had significantly more subclonal mutations than the primary tumors ($P = 7 \times 10^{-4}$; Figure 2A and Supplemental Table 11). Despite their absence from the COSMIC database, we could not exclude the possibility that these mutations could be conferring a growth advantage for subclones that could ultimately contribute to recurrent or refractory disease. Regardless, these mutations are unlikely to constitute effective therapeutic targets up front, given their absence from the predominant cancer population.

The mutational profile was also distinct in the recurrent samples, as they contained a significantly reduced proportion of C→T transitions ($P < 10^{-5}$, 2-proportion z test) and increased proportions of A→T ($P < 0.05$) and C→A ($P < 0.005$) transversions (Figure 3C). Overall, the recurrent tumors had a greater percentage of transversions ($P < 10^{-4}$; Figure 3D and Supplemental Figure 2).

In part based upon the large number of mutations commonly present in tumors, genetic alterations that affect numerous pro-

tein coding genes have typically been considered a fundamental requirement for cancer development. The finding that *SMARCB1* is the sole gene recurrently mutated at high frequency in extremely aggressive and lethal RTs, and in some cases may be the only mutated gene, prompts essential questions: What accounts for the extreme paucity of mutations, and how can these data be reconciled with current models of cancer that estimate that 5 to 15 driving mutations are required for oncogenesis (12)?

We considered 4 possible explanations. First, as we have only sequenced exome DNA, we cannot exclude the existence of mutations in noncoding portions of the genome, such as in noncoding RNAs or regulatory elements or in mutations in low coverage areas. Further, we cannot exclude balanced translocations or inversions, although these are not characteristic of RTs by karyotype (13). Nonetheless, occult events could cooperate with *SMARCB1* loss. Second, as mutations were largely identified based upon differences between tumor and normal DNA, contributions from germ line events are difficult to exclude. However, since genetically engineered models have demonstrated that inactivation of *Smarcb1* drives extremely rapid formation of cancer in all mice and since this occurs on several genetic backgrounds (14, 15), it seems unlikely that germ line alterations are essential for cancer formation driven by *SMARCB1* loss. Third, the developmental stage/epigenetic state may serve a contributory role. During development, there is relative enrichment of minimally differentiated cell populations that have a high proliferative capacity. Consequently, it is possible that developmentally restricted or lineage-specific populations of cells characterized by a certain epigenetic state are particularly susceptible, such that mutation of a single chromatin remodeler can drive transformation. This is consistent with our mouse model in which *Smarcb1* deletion in the T cell lineage leads to transformation of a highly specific cell type, $CD8^+CD44^{hi}CD122^{lo}$ memory T cells, a population that has a high intrinsic capacity for proliferation. Notably, this transformation arises due to aberrant responses to lineage-specific T cell receptor

Table 1
Somatic mutation types in 32 primary RT samples

Sample	Silent	Missense	Nonsense/indel/splice site	Total	Rate per Mb
07-057	1	2	0	3	0.10
07-221	0	1	0	1	0.03
07-242A	1	4	3	8	0.28
07-295	1	3	1	5	0.17
08-114	0	0	0	0	0
08-129	0	9	1	10	0.34
08-172	3	1	0	4	0.14
08-230	1	3	1	5	0.17
08-237	2	4	0	6	0.21
08-262A	1	2	0	3	0.10
08-310	1	3	1	5	0.17
09-125	1	5	3	9	0.31
09-130	1	4	1	6	0.21
09-131	2	1	1	4	0.14
09-182	2	1	0	3	0.10
09-223	0	0	0	0	0
09-236	2	5	1	8	0.28
09-365	2	6	1	9	0.31
10-012	1	2	0	3	0.10
10-016	0	1	1	2	0.07
10-031	3	2	0	5	0.17
10-155	3	1	0	4	0.14
10-162	1	4	3	8	0.28
10-201	2	4	1	7	0.24
10-213	1	4	2	7	0.24
10-255	0	1	1	2	0.07
10-305B	2	8	3	13	0.45
10-330	1	3	2	6	0.21
10-355	1	3	0	4	0.14
10-407	1	6	1	8	0.28
SJDOS006	0	3	2	5	0.17
SJDOS008	4	4	1	9	0.31
Total	41	100	31	172	
Mean	1.28	3.13	0.97	5.38	0.19
Median	1	3	1	5	0.17

Numbers in the "total" column represent the total number of somatic substitutions and indels.

signaling caused by *SMARCB1* loss (16). Such specificity in lineage effects could explain why mutation of *SMARCB1* is largely restricted to RTs and a few other cancers. It should be noted that the SWI/SNF complex contributes to differentiation control in many tissues and that several other members of the SWI/SNF complex are mutated across a variety of adult cancers, including subsets of ovarian and renal cancer, among other cancers, with each mutated subunit having a distinct profile of associated cancers (8, 17–19). We speculate that this specificity may be related to epigenetic state and distinct roles for the subunits in modulating interactions with particular transcription factors. Fourth, it is conceivable that 5 to 15 is an overestimation of the number of mutations required for oncogenic transformation. While adult cancers that arise in tissues exposed to mutagens, such as those of skin, lung, and the gastrointestinal tract, generally contain an extremely high number of mutations, other adult cancers, such as acute myeloid leukemia (AML) and breast cancer, are typically much simpler. Similarly, pediatric cancers, such as osteosarcomas, have extremely complex genomes, while

retinoblastomas, poorly differentiated but generally good prognosis cancers, have extremely low mutation rates (20, 21). Regardless of the explanation, our data from RTs demonstrate that collaboration between multiple coding mutations is not essential for the genesis of extremely aggressive and highly lethal cancers.

It is interesting that a type of cancer that has extremely few gene mutations at the time of initial diagnosis is characterized by a much higher number of mutations in recurrent samples. The reason for this is unclear. While we do not have specific treatment data, these patients were treated with chemotherapy and potentially radiation therapy. It is possible that genotoxic chemotherapy directly causes such damage. This possibility is supported by the presence in recurrent samples of a high percentage of transversions, a mutation type known to be associated with chemotherapy and similarly seen in recurrent AML (22). This raises the possibility that chemotherapy can cause the conversion of a remarkably simple cancer genome into one with 8-fold more mutations, a possibility with substantial clinical implications, as such mutations could potentially contribute to resistance. Alternatively, the selective pressure of chemotherapy may result in the outgrowth of subclones with high rates of mutation, or, conceivably, the fundamental nature of the recurrent disease has changed such that these cancers have acquired genetic instability and a high mutation rate.

Finally, while there is increasing evidence that epigenetic regulators are mutated in a large variety of cancers, it has been unclear whether these alterations are selected for because they act to facilitate genomic instability, because they potentiate the effects of other mutations, or because they directly drive oncogenic transformation. Particularly in cancers in which mutations in chromatin regulators exist in a genetically complex background, it has been extremely difficult to determine the relative contribution of epigenetic alterations. Our findings from RTs demonstrate that mutations in the SWI/SNF chromatin remodeling complex can act as potent drivers of cancer. Understanding the contributions of mutations in these remodelers to oncogenesis has the potential to facilitate development of targeted therapies for the wide variety of SWI/SNF mutant cancers.

Methods

Samples. Tumor tissue and matched blood from 32 newly diagnosed patients with cancer and from 3 recurrent tumors were collected. Mutation and deletion analysis of the *SMARCB1* gene was performed as previously described (6). Tumors were reviewed to confirm the diagnosis and to estimate tumor content (The Children's Hospital of Philadelphia, Hospital Sant Joan de Déu). DNA was extracted using standard techniques.

SNP arrays. All of these samples were processed and hybridized to Affymetrix SNP 6.0 arrays for genotyping and copy number analysis (2). SNP array data were further analyzed using the ABSOLUTE tool (23) to infer the tumor purity and ploidy (11, 24, 25).

Whole-exome sequence data. Library construction followed the procedure of previous publications (11, 24, 25). Descriptions of sequencing and analysis methods are in the Supplemental Methods. Data were deposited in dbGaP (accession no. phs000508). The complete list of all detected mutations can be found in Supplemental Table 12.

Statistics. Comparison of mutation rates was performed using a 2-tailed Welch's t test for samples with unequal variance. Two-proportion z test and Pearson's χ^2 test with a Yates's correction for continuity were used to analyze the different proportions of mutation type. Subclonal mutation frequency was analyzed using Fisher's exact test. Data in figures are shown as mean ± SEM.
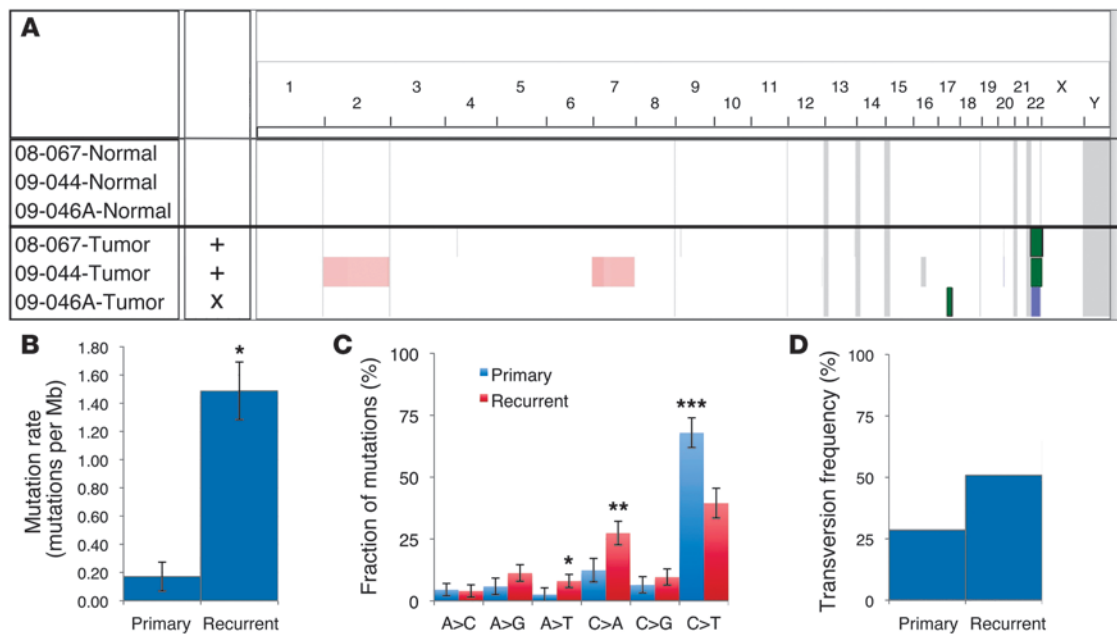
Figure 3
Recurrent RTs have more mutations than primary tumors. (**A**) SNP array of 2 matched tumor/normal pairs from recurrent tumors reveals an aneuploid tumor sample (09-044). Blue represents deletion; red represents amplification; and green represents copy neutral LOH. (**B**) The mutation rate in recurrent RTs is significantly higher (*$P < 0.005$) than that in primary RT samples. (**C**) While primary samples had a greater proportion of C→T transitions, recurrent samples had a greater proportion of C→A and A→T transversions. Significant differences between primary and recurrent samples are indicated. *$P < 0.05$, **$P < 0.005$, ***$P < 10^{-5}$. (**D**) Recurrent samples have significantly more transversions than primary samples ($P < 0.0005$).

Address correspondence to: Jaclyn Biegel, The Children's Hospital of Philadelphia, ARC Room 1002, 3615 Civic Center Boulevard, Philadelphia, Pennsylvania 19104, USA. Phone: 215.590.3856; Fax: 215.590.3764; E-mail: biegel@mail.med.upenn.edu. Or to: Gad Getz, Broad Institute, 301 Binney St., Cambridge, Massachusetts 02142, USA. Phone: 617.714.7471; Fax: 617.714.8102; E-mail: gadgetz@broadinstitute.org. Or to: Charles Roberts, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. Phone: 617.632.6497; Fax: 617.582.8096; E-mail: Charles_Roberts@dfci.harvard.edu.

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646–674.
2. TCGA. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature.* 2008;455(7216):1061–1068.
3. TCGA. Integrated genomic analyses of ovarian carcinoma. *Nature.* 2011;474(7353):609–615.
4. Roberts CWM, Biegel JA. The role of SMARCB1/INI1 in development of rhabdoid tumor. *Cancer Biol Ther.* 2009;8(5):412–416.
5. McKenna ES, et al. Loss of the epigenetic tumor suppressor SNF5 leads to cancer without genomic instability. *Mol Cell Biol.* 2008;28(20):6223–6233.
6. Eaton KW, Tooke LS, Wainwright LM, Judkins AR, Biegel JA. Spectrum of SMARCB1/INI1 mutations in familial and sporadic rhabdoid tumors. *Pediatr Blood Cancer.* 2011;56(1):7–15.
7. Versteege I, et al. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature.* 1998;394(6689):203–206.
8. Wilson BG, Roberts CW. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer.* 2011; 11(7):481–492.
9. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 2011;12(4):R41.
10. Durinck S, et al. Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov.* 2011; 1(2):137–143.
11. Stransky N, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science.* 2011;333(6046):1157–1160.
12. Bozic I, et al. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci U S A.* 2010;107(43):18545–18550.
13. White FV, et al. Congenital disseminated malignant rhabdoid tumor: a distinct clinicopatho-

logic entity demonstrating abnormalities of chromosome 22q11. *Am J Surg Pathol.* 1999; 23(3):249–256.

14. Roberts CW, Galusha SA, McMenamin ME, Fletcher CD, Orkin SH. Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice. *Proc Natl Acad Sci U S A.* 2000;97(25):13796–13800.

15. Roberts CW, Leroux MM, Fleming MD, Orkin SH. Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. *Cancer Cell.* 2002;2(5):415–425.

16. Wang X, et al. TCR-dependent transformation of mature memory phenotype T cells in mice. *J Clin Invest.* 2011;121(10):3834–3845.

17. Jones S, et al. Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science.* 2010;330(6001):228–231.

18. Wiegand KC, et al. ARID1A mutations in endometriosis-associated ovarian carcinomas. *N Engl J Med.* 2010;363(16):1532–1543.

19. Varela I, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature.* 2011;469(7331):539–542.

20. Smida J, et al. Genomic alterations and allelic imbalances are strong prognostic predictors in osteosarcoma. *Clin Cancer Res.* 2010;16(16):4256–4267.

21. Zhang J, et al. A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature.* 2012;481(7381):329–334.

22. Ding L, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature.* 2012;481(7382):506–510.

23. Carter SL, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotech.* 2012;30(5):413–421.

24. Berger MF, et al. The genomic complexity of primary human prostate cancer. *Nature.* 2011; 470(7333):214–220.

25. Chapman MA, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature.* 2011;471(7339):467–472.