HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2013 December 30.

Published in final edited form as:

Nature. ; 482(7385): 400–404. doi:10.1038/nature10755.

Cancer Exome Analysis Reveals a T Cell Dependent Mechanism of Cancer Immunoediting

Hirokazu Matsushita[1,*,†], **Matthew D. Vesely**[1,*], **Daniel C. Koboldt**[2], **Charles G. Rickert**[1], **Ravindra Uppaluri**[3], **Vincent J. Magrini**[2,4], **Cora D. Arthur**[1], **J. Michael White**[1], **Yee-Shiuan Chen**[1], **Lauren K. Shea**[1], **Jasreet Hundal**[2], **Michael C. Wendl**[2,4], **Ryan Demeter**[2], **Todd Wylie**[2], **James P. Allison**[5,6], **Mark J. Smyth**[7,8], **Lloyd J. Old**[9], **Elaine R. Mardis**[2,4], and **Robert D. Schreiber**[1]

[1]Department of Pathology and Immunology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, USA

[2]The Genome Institute, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, USA

[3]Department of Otolaryngology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, USA

[4]Department of Genetics, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, USA

[5]Department of Immunology, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

[6]Howard Hughes Medical Institute, Peter MacCallum Cancer Centre, East Melbourne, 3002, Victoria, Australia

[7]Cancer Immunology Program, Peter MacCallum Cancer Centre, East Melbourne, 3002, Victoria, Australia

Correspondence should be addressed to: Dr. Robert Schreiber, Department of Pathology and Immunology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, (314) 362-8787 (office), (314) 362-3939 (fax), schreiber@immunology.wustl.edu.

*These authors contributed equally to this work

†Present address: Department of Immunotherapeutics (Medinet), The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.

Author Contributions

H.M. and M.D.V. were involved in all aspects of this study including planning and performing experiments, analyzing and interpreting data, and writing the manuscript. C.G.R., R.U., C.D.A., J.M.W., Y.S.C., and L.K.S. also performed experiments and analyzed data. V.J.M., R.D., and members of The Genome Institute performed Illumina library preparation, cDNA capture and sequencing as well as validation Roche/454 pyrosequencing and 3730 sequencing. D.C.K. analyzed and interpreted sequencing data from this study and previously published cancer genome data. J.H. and T.W. analyzed cDNA CapSeq data for potential MHC class I epitopes. M.C.W. performed the phylogenetic analysis on the tumour cells. J.P.A., M.J.S., and L.J.O. interpreted data and contributed to the preparation of the final manuscript. E.R.M. and R.D.S oversaw all the work performed, planned experiments, interpreted data and wrote the manuscript.

The authors declare no competing financial interests.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

[8]Department of Pathology, University of Melbourne, Parkville, 2010 Victoria, Australia

[9]New York Branch of The Ludwig Institute for Cancer Research at Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

Abstract

Cancer immunoediting, the process whereby the immune system controls tumour outgrowth and shapes tumour immunogenicity, is comprised of three phases: elimination, equilibrium and escape[1–5]. Although many immune components that participate in this process are known, its underlying mechanisms remain poorly defined. A central tenet of cancer immunoediting is that T cell recognition of tumour antigens drives the immunologic destruction or sculpting of a developing cancer. However, our current understanding of tumour antigens comes largely from analyses of cancers that develop in immunocompetent hosts and thus may have already been edited. Little is known about the antigens expressed in nascent tumour cells, whether they are sufficient to induce protective anti-tumour immune responses or whether their expression is modulated by the immune system. Here, using massively parallel sequencing, we characterize expressed mutations in highly immunogenic methylcholanthrene-induced sarcomas derived from immunodeficient *Rag2*$^{-/-}$ mice which phenotypically resemble nascent primary tumour cells[1,3,5]. Employing class I prediction algorithms, we identify mutant spectrin-β2 as a potential rejection antigen of the d42m1 sarcoma and validate this prediction by conventional antigen expression cloning and detection. We also demonstrate that cancer immunoediting of d42m1 occurs via a T cell-dependent immunoselection process that promotes outgrowth of pre-existing tumour cell clones lacking highly antigenic mutant spectrin-β2 and other potential strong antigens. These results demonstrate that the strong immunogenicity of an unedited tumour can be ascribed to expression of highly antigenic mutant proteins and show that outgrowth of tumour cells that lack these strong antigens via a T cell-dependent immunoselection process represents one mechanism of cancer immunoediting.

For this study, we chose two representative, highly immunogenic, unedited methylcholanthrene (MCA)-induced sarcoma cell lines, d42m1 and H31m1, derived from immunodeficient *Rag2*$^{-/-}$ mice[1]. Both grow progressively when transplanted orthotopically into *Rag2*$^{-/-}$ mice, but are rejected when transplanted into naive wild type (WT) mice (Supplementary Fig. 1 and 2). Using a modified form of exome sequencing involving cDNA capture by mouse exome probes and Illumina deep sequencing (i.e., cDNA Capture Sequencing (“cDNA CapSeq”)), we identified 3,737 somatic, non-synonymous mutations in d42m1 cells (3,398 missense, 221 nonsense, 2 nonstop and 116 splice site mutations) and 2,677 non-synonymous mutations in H31m1 cells (2,391 missense, 160 nonsense, 3 nonstop and 123 splice site mutations) (Fig. 1a, Supplementary Fig. 3 and Supplementary Table 1). The mutations in each cell line were largely distinct—d42m1 and H31m1 share only 119 identical missense mutations (Fig. 1b and Supplementary Table 2)—a result that potentially explains the unique antigenicity of each cell line (Supplementary Fig. 4). Although d42m1 and H31m1 display mutations in known cancer genes[6], the functional effects of these novel mutations remain undefined. Nevertheless, both tumours have cancer-causing mutations in *Kras* (codon 12) and *Trp53* that are frequently observed in human and mouse cancers[7,8,9] (Supplementary Table 3). The mutation calls were confirmed by independent Roche/454

pyrosequencing of 22 genes using tumour genomic DNA and by documenting their absence in normal cells from the same mouse that developed the tumour (Supplementary Table 4).

Comparing cDNA CapSeq data of d42m1 and H31m1 cells to human cancer genomes[10–17] revealed two similarities. First, 46–47% of mutations in d42m1 and H31m1 are C/A or G/T transversions which represent chemical-carcinogen signatures[7,13,14] similar to those of lung cancers from smokers (44–46%) but not seen in human cancers induced by other mechanisms (8–16%) (Fig. 1c). Second, the mutation rates of d42m1 and H31m1 are about 10-fold higher than those of lung cancers from smokers, but within 3-fold of hypermutator smoker lung cancers with mutations in DNA repair pathway genes (Fig. 1d). Interestingly, d42m1 and H31m1 also display mutations in DNA repair genes (Supplementary Table 3), although these novel mutations have not been functionally characterized. Thus, mouse MCA-induced sarcomas display qualitative and quantitative genomic similarities to carcinogen-induced human cancers.

When parental d42m1 sarcoma cells were transplanted into naïve WT mice, approximately 20% of recipients developed escape tumours (Supplementary Fig. 5a, c). Cell lines made from three escape tumours (d42m1-es1, d42m1-es2 and d42m1-es3) formed progressively growing sarcomas when transplanted into naïve WT recipients (Fig. 2a). In contrast, parental d42m1 tumour cells passaged through *Rag2*$^{-/-}$ mice maintained high immunogenicity (Supplementary Fig. 5b, d). Additional analyses revealed that whereas 8 of 10 clones of d42m1 were rejected in WT mice, two clones (d42m1-T3 and d42m1-T10) grew with kinetics similar to d42m1 escape tumours (Fig. 2a and Supplementary Fig. 6). Thus, the d42m1 cell line consists mostly, but not entirely, of highly immunogenic clones and undergoes immunoediting in WT mice. cDNA CapSeq of parental d42m1 cells, clones and escape tumours revealed that all expressed similar numbers of mutations (Supplementary Fig. 7a and Supplementary Table 1) and phylogenetic analysis revealed that all d42m1-derived cells were genomically related to one another but distinct from H31m1 and normal fibroblasts (Supplementary Fig. 7b). However, regressor clones clustered more closely to parental d42m1 cells while progressor clones clustered more closely to cells from escape tumours. Thus, the d42m1 tumour cell line consists of a related, but heterogeneous population of tumour cells.

Tumour-specific mutant proteins presented on mouse or human MHC class I molecules are known to represent one class of tumour-specific antigens for CD8$^+$ T cells[18,19]. Therefore, we used *in silico* analysis[20] to assess the theoretical capacities of missense mutations from d42m1-related tumour cells to bind MHC class I proteins. Each d42m1-related cell type expressed many potential high affinity (IC_{50} < 50 nM; Affinity Value > 2) epitopes that could bind to H-2D^b or H-2K^b (Fig. 2b). Of these, 39–42 were expressed only in the regressor subset of d42m1-related cells (7–9 for H-2D^b, 30–35 for H-2K^b), including 31 expressed in all regressor cells (Supplementary Table 5). Thus, ~1% of the missense mutations in d42m1 are selectively expressed in rejectable d42m1 clones.

Whereas parental and regressor d42m1 cells stimulated IFN-γ release *in vitro* when incubated with a specific CD8$^+$ cytotoxic T lymphocyte (CTL) clone (C3) derived from a WT mouse that had rejected parental d42m1 tumour cells (Fig. 3a, b), progressor d42m1

clones, cells from escape tumours or unrelated MCA sarcomas were not. This result demonstrated that all regressor d42m1 tumour cells share a mutation that forms the epitope recognized by C3 CTLs. Since recognition of d42m1 regressor cells by C3 CTLs is restricted by H-2D^b (Fig. 3c), we postulated that an R913L mutation in spectrin-β2 produced the most likely target for C3 CTLs because its expression was restricted to d42m1 regressor clones and it formed an epitope that showed high affinity binding potential to H-2D^b in contrast to the WT sequence predicted to bind with low affinity (Fig. 3d and Supplementary Table 5).

To verify the importance of mutant spectrin-β2 on d42m1 antigenicity, we independently identified the tumour antigen recognized by the C3 CTL clone using a T cell-based expression cloning approach[21]. After three screening rounds, a single positive cDNA was identified encoding a sequence identical to the R913L spectrin-β2 mutant (Fig. 3e). Thus, conventional antigen expression cloning identified the same mutation predicted by the genomic sequencing.

Mutation-specific qRT-PCR revealed the presence of mutant *spectrin-β2* mRNA in parental d42m1 tumour cells and regressor d42m1 clones, but not in progressor d42m1 clones or escape tumours (Fig. 3f), nor in normal tissue of the mouse from which the d42m1 tumour was derived (Supplementary Table 4 and Supplementary Fig. 8). Additionally, C3 CTLs discriminated between mutant and WT spectrin-β2 peptide sequences when presented on an unrelated H-2D^b expressing cell line (Fig. 3g). Whereas the mutant (VAVVNQIAL) peptide stimulated C3 CTLs in a dose-dependent manner, the WT (VAVVNQIAR) peptide did not, even when added in 1000-fold excess. Using labeled H-2D^b tetramers generated with mutant peptide, mutant spectrin-β2 specific CD8$^+$ T cells accumulated over time in parental d42m1 tumours developing *in vivo* and draining lymph nodes (DLNs) prior to tumour rejection (Fig. 4a, b). In contrast, no mutant spectrin-β2 specific CD8$^+$ T cells were detected in progressively growing escape tumours or DLNs. These data demonstrate that mutant spectrin-β2 expressed selectively in a high proportion of unedited d42m1 tumour cells evokes a T cell response in naïve WT mice that promotes the elimination of antigen expressing tumour cells.

To test whether expression of mutant spectrin-β2 was sufficient to drive rejection of d42m1 tumour cells, we enforced expression of either mutant or WT spectrin-β2 in d42m1-es3 cells that lack this mutation (Supplementary Fig. 9a) and followed their growth in WT mice. Whereas d42m1-es3 tumour cell clones transduced with either control retrovirus or retrovirus encoding WT spectrin-β2 (WT.1 and WT.3) grew progressively with growth kinetics similar to unmanipulated d42m1-es3 cells, d42m1-es3 clones expressing mutant spectrin-β2 (mu.6 and mu.14) were rejected in WT mice, but not in *Rag2*$^{-/-}$ mice (Fig. 4c and Supplementary Fig. 9b, c, d). CD8$^+$ T cells specific for mutant spectrin-β2 did not infiltrate d42m1-es3 tumours expressing WT spectrin-β2 (WT.3), but were present in rejecting d42m1-es3 tumours expressing mutant spectrin-β2 (mu.14) (Fig. 4d). Thus, mutant spectrin-β2 is indeed a major rejection antigen of d42m1 sarcoma cells and d42m1 escape from immune control is the consequence of outgrowth of d42m1 clones that lack expression of dominant rejection antigens.

The possibility that the lack of dominant rejection antigen(s) in a small subset of d42m1 cells was due to epigenetic silencing was ruled out because no spectrin-β2 mutation was (a) found by sequencing genomic DNA from progressor d42m1 clones or escape tumours (Supplementary Table 4) or (b) expressed in d42m1 progressor clones or escape tumours following treatment with inhibitors of methyltransferases or histone deacetylases (Supplementary Fig. 10). We therefore asked whether T cell-dependent immunoselection explained outgrowth of escape tumours. Specifically, we examined the *in vivo* growth behavior of a tumour cell mixture containing a vast majority of highly immunogenic, mutant spectrin-β2^{+} d42m1-T2 cells and a minority of mutant spectrin-β2^{-} d42m1-T3 progressor cells. To distinguish between the two cell types, we labeled d42m1-T2 with RFP (modified to eliminate class I epitopes) and d42m1-T3 with GFP and documented that the labeling did not alter their *in vivo* growth characteristics. We found that we could recapitulate the tumour growth phenotype of parental d42ml at a ratio of 95% d42m1-T2 cells to 5% d42m1-T3 cells (Fig. 4e). At this ratio, 100% of *Rag2*$^{-/-}$ mice and WT mice depleted of either CD4^{+} or CD8^{+} T cells developed progressively growing tumours (Fig. 4f). In contrast, 5/20 (25%) WT mice injected with the tumour cell mixture developed escape tumours, a result that recapitulated the behavior of parental d42m1. Tumours harvested from *Rag2*$^{-/-}$ mice were comprised of 84% d42m1-T2 cells and 14% d42m1-T3 cells (Fig. 4h) and expressed mutant *spectrin-β2* (Fig. 4g), i.e., they resembled the initial 95:5 cell mixture. In contrast, tumours that grew out in WT mice consisted of 98% d42m1-T3 tumour cells and lacked mutant *spectrin-β2* (Fig. 4g, h). Thus, d42m1 escape tumours develop as a consequence of T cell-dependent immunoselection favoring the outgrowth of tumour cells that lack major rejection antigens.

This report shows that the combination of cancer exome sequencing and *in silico* epitope prediction algorithms can identify highly immunogenic, tumour-specific mutational antigens in unedited carcinogen-induced cancers that serve as targets for the elimination phase of cancer immunoediting. To our knowledge, this is the first study to use a genomics approach to experimentally identify a tumour antigen, to specifically identify an antigen from an unedited tumour and to demonstrate that T cell-dependent immunoselection is a mechanism underlying outgrowth of tumour cells that lack strong rejection antigens. This mechanism most likely also produces other types of escape tumours, such as those that develop inactivating mutations in antigen presentation genes (e.g., those encoding MHC class I proteins), which are frequently observed in clinically apparent human cancers[22,23]. Developing carcinogen-induced tumours (e.g., mouse MCA sarcomas or human smoker lung cancers) may be the preferred targets of cancer immunoediting because they express the greatest number of mutations that might function as neoantigens. However, since ~1% of the mutations in d42m1 are selectively expressed in regressor tumour clones, it is possible that spontaneous tumours arising by other means that harbour as few as 100–200 mutations could still be susceptible to immunological sculpting as they develop.

The immunodominance of mutant spectrin-β2 in driving tumour rejection in many ways resembles that of certain viral antigens[24] and is likely due to the presence in d42m1 of 4 copies of chromosome 11 that carries the spectrin-β2 gene thereby producing a highly abundant neoepitope that binds to H-2D^{b} 750-fold stronger than that of the WT sequence.

More work is needed to determine which of the other mutations, if any, selectively expressed in d42m1 regressors function as rejection antigens. Immunoepitope analysis of parental H31m1 reveals that it expresses multiple potential strong neoantigens (19 potential strong binders to H-2D^b and 58 to H-2K^b) (Supplementary Fig. 11a) and induces both H-2D^b and H-2K^b restricted CD8$^+$ T cell responses during rejection (Supplementary Fig. 11b). This result suggests that H31m1 displays an even more complex antigenicity than d42m1 and probably explains why H31m1 never produces escape tumours in WT mice (Supplementary Fig. 11c).

Chemically-induced tumours have played a critical role in the history of tumour immunology, providing the first unequivocal demonstration of tumour-specific antigens[25,26] and, subsequently the first evidence of cancer immunoediting[1,2]. It is therefore significant that this same model has now provided new insights into the antigenic targets of cancer immunoediting and some of the key molecular mechanisms that drive the process. While more work is needed to determine whether and how frequently this process occurs during development of spontaneous and carcinogen-induced human cancers, it is tempting to speculate that a genomics approach to tumour antigen identification could, in the future, facilitate the development of individualized cancer immunotherapies directed at tumour-specific—rather than cancer-associated—antigens.

ONLINE-ONLY METHODS

Mice

Ifngr1$^{-/-}$ mice[27] and *Ifnar1*$^{-/-}$ mice[28] on a 129/Sv background were originally provided by Dr. Michel Aguet and were bred in our specific pathogen-free animal facility. Wild type and *Rag2*$^{-/-}$ mice were purchased from Taconic Farms. All mice were male and on a 129/Sv background and were housed in our specific pathogen-free animal facility. For all experiments, male mice were 8–12 weeks of age and performed in accordance with procedures approved by the AAALAC accredited Animal Studies Committee of Washington University in St. Louis.

Tumour transplantation

3-methylcholanthrene (MCA)-induced sarcomas used in this study were generated in male 129/Sv strain wild type or *Rag2*$^{-/-}$ mice and banked as low passage tumour cells as previously described[1]. Tumour cells derived from frozen stocks were propagated *in vitro* in RPMI media (Hyclone, Logan, UT) supplemented with 10% FCS (Hyclone) and injected subcutaneously in 150 μl of endotoxin-free PBS into the flanks of recipient mice. Tumour cells were >90% viable at the time of injection as assessed by trypan blue exclusion and tumour size was quantified as the average of two perpendicular diameters. For antibody depletion studies, 250 μg of control Ig (PIP), anti-CD4 (GK1.5), or anti-CD8α (YTS169.4) were injected intraperitoneally into mice at day −1 and every 7 days thereafter.

Isolation of normal skin fibroblasts

Skin fibroblasts were isolated from three independent male 129/Sv *Rag2*$^{-/-}$ pups by harvesting skin and incubating in 0.25% trypsin (Hyclone) at 37°C for 30 minutes prior to

washing in DMEM media (Hyclone). After washing, chunks of skin were filtered to achieve single cell suspensions and cultured *in vitro* with DMEM media. After 3 passages, skin fibroblasts were harvested to isolate genomic DNA and total RNA.

Extraction of genomic or complementary DNA

Genomic DNA from sarcoma cells and normal skin fibroblasts was extracted using DNeasy Blood & Tissue Kit (Qiagen). For cDNA isolation, total RNA from sarcoma cells and normal skin fibroblasts was isolated using RNeasy Mini kit (Qiagen) and cDNA was synthesized using oligo (dT) primers and SuperScript II Reverse Transcriptase (Invitrogen).

cDNA capture, sequencing, and alignment (cDNA CapSeq)

cDNA samples from each tumour (100 ng) were constructed into Illumina libraries according to the manufacturer's protocol (Illumina Inc, San Diego, CA) with the following modifications: 1) cDNA was fragmented using Covaris S2 DNA Sonicator (Covaris, Inc. Woburn, MA) in 1X end-repair buffer followed by the direct addition of the enzyme repair cocktail (Lucigen, Madison, WI). Fragment sizes ranged between 100 and 500 bp. 2) Illumina adapter-ligated DNA was amplified in four 50 μl PCRs for five cycles using 4 μl adapter-ligated cDNA, 2X Phusion Master Mix and 250 nM forward and reverse primers, 5′AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC GAT C and 5′ CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCG AT C, respectively. 3) Solid Phase Reversible Immobilization (SPRI) bead cleanup was used to purify the PCR-amplified library and to select for 300–500 bp fragments. 500 ng of the size-fractionated Illumina library was hybridized with the Agilent mouse exome reagent. After hybridization at 65°C for 24 hrs, we added 50 μl of DynaBeads M-270 Streptavidin-coated paramagnetic beads (10 mg/ml) to selectively remove the biotinylated Agilent probes and hybridized cDNA library fragments. The beads were washed according to manufacturer's protocol (Agilent) and the captured library fragments were released into solution using 50 μl of 0.125 N NaOH and neutralized with an equal volume of neutralization buffer (Agilent). The recovered fragments then were PCR amplified according to the manufacturer's protocol using 11 cycles in the PCR. Illumina library quantification was completed using the KAPA SYBR FAST qPCR Kit (KAPA Biosystems, Woburn, MA). The qPCR result was used to determine the quantity of library necessary to produce 180,000 clusters on a single lane of the Illumina GAIIx. One lane of 100 bp paired-end data was generated for each captured sample (since cDNA was used as the source for sequencing, we refer to this process as cDNA Capture Sequencing or cDNA CapSeq). Illumina reads were aligned to the NCBI build 37 (Mm9) mouse reference sequence using BWA[29] v0.5.5 (with −q 5 soft trimming). Alignments from multiple lanes for the same sample were merged together using SAMtools r599, and duplicates were marked using Picard v1.29.

Mutation detection and annotation

Putative somatic mutations were identified using VarScan 2 (v2.2.4)[30] with the parameters "*--min-coverage 3 --min-var-freq 0.08 --p-value 0.10 --somatic-p-value 0.05 --strand-filter 1*" and specifying a minimum mapping quality of 10. Variants whose supporting reads

exhibited read position bias (average read position <10 or >90), strand bias (>99% of reads on one strand), or mapping quality (score difference >30, or mismatch quality sum difference >100) relative to reference supporting reads were removed as probable false positives. We also required that the variant allele be present in at least 10% of tumour reads and no more than 5% of normal reads. The single nucleotide variants (SNVs) meeting these criteria were annotated using an internal database of Genbank/Ensembl transcripts (v58_73k). In the event that a variant was annotated using multiple transcripts, the annotation of most severe effect was used. Non-silent coding mutations (missense, nonsense/nonstop, or splice-site) were prioritized for downstream analysis.

Mutation rate and overlap comparisons

Mutation rates were estimated for each tumour sample using the number of putative "tier 1" SNVs (missense, nonsense/nonstop, splice site, silent, or noncoding RNA). To account for variability in coverage between samples, the SNV count for each tumour sample (*S*) was divided by a coverage factor (*F*), computed as the fraction of all tier 1 SNVs identified in any tumour sample (n=16,991) that were covered by at least 4 reads in a given sample. For example, in the d42m1 parental sample, 15,852 of 16,991 tier 1 SNV positions were covered, for a coverage factor of 93.30%. The number of coverage-adjusted mutations in each sample was divided by the total size of tier 1 space in the mouse genome (43.884 Mbp) to determine the number of coding mutations per megabase (R).

$$R=(S/F)/(43.884\ Mbp)$$

For the mutation overlap comparisons and relatedness-to-parental-tumour analysis, only high-confidence missense mutations were used (i.e., 20X or above). A mutation was considered "shared" between two samples if both samples had a predicted mutation at the same genomic position. For the comparison of mutated genes between d42m1 and H31m1 parental lines, a gene was considered "shared" if both d42m1 and H31m1 samples had a predicted missense mutation in that gene, even if the mutations did not occur at the same position.

Roche/454 sequencing and validation

PCR primers were designed for 11 SNVs predicted to be somatic in d42m1 tumour samples, as well as 11 control sites which were H31m1-specific, low-confidence, or removed by the false-positive filter. All 22 SNVs were PCR amplified individually in 11 samples (SK1.1, d42m1, H31m1, T2, T3, T5, T9, T10, es1, es2, and es3) using MID-tailed primers to enable sample identification. PCR products were pooled together prior to sequencing on ¼ run of the Roche/454 Titanium platform. Read sequences and quality scores were extracted from 454 data files using *sffinfo* (454 proprietary software) then aligned to the mouse build 37 reference sequence (Mm19) using SSAHA2 v2.5.3[31] with the SAM output option. Alignments were imported to BAM format and a "pileup" assembly file generated using SAMtools v0.1.18[32]. The average 454 sequence depth for targeted positions was 1,216x per sample. Validation read counts and allele frequencies in each sample at each variant position were determined using the *pileup2cns* command of VarScan v2.2.7[30]. At least 20 reads with

base quality of 20 or higher were required to confirm or refute a variant. 454 sequencing data and the primers used are presented in Supplementary Table 4.

3730 sequencing and validation

Eight SNVs predicted to be somatic were selected for validation by PCR and 3730 sequencing in flow-sorted $CD45^+$ and $CD45^-$ cells from the original d42m1 tumour. Genomic DNA and cDNA from $CD45^-$ (tumour) cells, and cDNA from $CD45^+$ (normal immune) cells were used for PCR amplification and then PCR products were sequenced individually on ABI 3730 using universal primers. Manual review was performed using amplicon-based assembly in the Integrative Genomics Viewer (IGV)[33] to determine the somatic status for each site. Data is presented in Supplementary Table 4.

MHC class I epitope prediction

All missense mutations for each d42m1-related tumour or H31m1 were analyzed for the potential to form MHC class I neoepitopes that bind to either $H\text{-}2D^b$ or $H\text{-}2K^b$ molecules. The artificial neural network (ANN) algorithm provided by the Immune Epitope Database and Analysis Resource (www.immuneepitope.org) was used to predict epitope binding affinities[20] and the results were ultimately expressed as "Affinity Values" (Affinity Value = $1/IC_{50}$ X 100). Predicted strong affinity epitopes expressed in d42m1 regressor tumours are listed in Supplementary Table 5.

Phylogenetic analysis of tumour samples

Sequencing data from normal *Rag2*$^{-/-}$ fibroblasts, d42m1 parental cells, d42m1 regressor clones, d42m1 progressor clones, d42m1 escape tumours, and H31m1 tumour cells were compared using PHYLogeny Inference Package (PHYLIP)[34] to generate a phylogenetic tree displaying the relatedness of each sample.

Antibodies

Anti-$H\text{-}2K^b$ (B8-24-3) and anti-$H\text{-}2D^b$ (B22/249) mAbs were generously provided by Dr. Ted H. Hansen (Washington University School of Medicine). Anti-CD4 (GK1.5), anti-CD8α (YTS169.4) mAbs and control immunoglobulin (PIP, a mAb specific for bacterial glutathione S-transferase) were produced from hybridoma supernatants and purified in endotoxin-free form by Protein G affinity chromatography (Leinco Technologies, St. Louis, MO). Purified Rat IgG was purchased from Sigma (St. Louis, MO). CD45-FITC, CD45-PE, CD8-APC, and purified anti-CD16/32 were purchased from BioLegend (San Diego, CA).

cDNA library construction and screening

To generate a d42m1 tumour cell cDNA library, mRNA was isolated from parental d42m1 tumour cells using a QuickPrep mRNA Purification kit (Amersham), converted into cDNA using SuperScript II First Strand Synthesis System (Invitrogen) and inserted into the EcoRI site of the expression vector pcDNA3 (Invitrogen). The cDNA library was divided into pools of 100 bacterial colonies with 200–300 ng of DNA from each pool transfected into 2.5 x 10^4 monkey COS cells engineered to ectopically express mouse $H\text{-}2D^b$ ($COS\text{-}D^b$) cells using Lipofectamine 2000. After 48 hr, 5 x 10^3 C3 CTL cells were added, and supernatants

were assayed for IFN-γ release 24 hrs later by ELISA. A single positive cDNA clone was isolated after screening 120,000 cDNA colonies. The putative H-2D^b-binding peptide VAVVNQIAL was predicted using the algorithm available at the Immune Epitope Database and Analysis Resource, http://www.immuneeptiope.org/. The peptides were kindly produced by Dr. Paul Allen and Steve Horvath (Washington University School of Medicine).

Expression vectors

Full length cDNA encoding wild type spectrin-β2 and mutant spectrin-β2 were cloned from parental d42m1 tumour cells by RT-PCR using primer pairs 5′-TGAGACAGTCAAGATGACGACCACGGTAGCCACA-3′ and 5′-CGGGACAACAGGGAAGTTCACTTCTTCTTGCCGA-3′. Wild type and mutant spectrin-β2 cDNA were subcloned from the TOPO-XL vector (Invitrogen) into the RV-GFP vector[35]. To generate the RV-RFP vector, full length cDNA encoding RFP was cloned from the pTurboRFP-C vector (Evrogen) by RT-PCR using primer pairs 5′-ATCTCAGAATTCATGAGCGAGCTGATCAAGGA - 3′ and 5′-ATCTCAGGATCCTTATCTGTGCCCCAGTTTGCTAG - 3′. RFP cDNA was then cloned into the RV vector. To remove candidate T cell epitopes in RFP, the nucleotide A was replaced by G at position 334 in the cDNA, resulting in amino acid substitution N112D. Coding sequences of the constructs were verified by DNA sequencing (Big Dye method; Applied Biosciences). The dominant negative version of the IFNGR1 subunit (IFNGR1ΔIC) was expressed into H31m1 and d42m1 tumour cells as previously described[36].

Establishment of CTL lines and clones

To generate the d42m1 specific C3 CTL clone, wild type mice were injected with 1 x 10^6 parental d42m1 tumour cells. Fourteen days later, the spleen was harvested from a mouse that rejected the tumour and a CTL line was established by stimulating 40 x 10^6 splenocytes with 2 x 10^6 parental d42m1 tumour cells pre-treated for 48 hr with 100 U/ml of recombinant murine IFN-γ and irradiated (100 Gy). After CD8$^+$ T cell purification using magnetic-beads (Miltenyi Biotec) and limiting dilution, the CTL clone C3 was obtained.

Measurement of IFN-γ production

To generate target cells, tumour cells were treated with 100 U/ml IFN-γ for 48 hrs and irradiated with 100 Gy prior to use. The C3 CTL clone was co-cultured at the indicated ratios with target tumour cells (10,000 or 5,000 cells) in 96-well round-bottomed plates overnight. IFN-γ in supernatants was quantified using an IFN-γ ELISA kit (eBioscience). For blocking assays, 10 μg/ml of α-CD8 (YTS-169.4), α-CD4 (GK1.5), or control immunoglobulin (PIP) were added to the cell culture of effector (C3 CTL clone) and target cells (tumours).

Cytotoxicity assay

To generate target cells, tumour cells were treated with 100 U/ml rMuIFN-γ for 48 hrs prior to use. One million tumour cells were labeled with 25 μCi of $Na_2{}^{51}CrO4$ (PerkinElmer, Boston, MA) for 90 minutes at 37ºC, washed and 10,000 cells seeded per well in 96-well round-bottom plates. The C3 CTL clone was co-cultured with the tumour target cells at the

indicated effector/target cell ratios and incubated for 4 hrs at 37ºC in 5% CO_2. Radioactivity was detected in the supernatants and percent specific killing was defined as (experimental condition cpm - spontaneous cpm)/(maximal (detergent) cpm - spontaneous cpm) x 100. Data points were obtained in duplicate.

Fluorescence-activated cell sorting analysis

For flow cytometry, cells were stained for 20 minutes at 4°C with 500 ng of Fc block (anti-CD16/32) and 200 ng of CD45, CD4, or CD8α in 100 μL of staining buffer (PBS with 1% FCS and 0.05% NaN_3 (Sigma)). Propidium iodide (PI) (Sigma) was added at 1 μg/mL immediately before FACS analysis. For quantitative analysis of tumour-infiltrating lymphocytes/leukocytes (TIL) and lymph node populations, a $CD45^+PI^-$ gate was used and gated events were collected on a FACSCalibur (BD Biosciences) and analyzed using FloJo software.

Tumour, draining lymph node, and spleen harvest

After tumour cell transplantation, established tumours were excised from mice, minced and treated with 1 mg/ml type IA collagenase (Sigma) in HBSS (Hyclone) for 2 hrs at room temperature. The ipsilateral inguinal tumour-draining lymph nodes and spleen were also harvested and crushed between two glass slides and vigorously resuspended to make single-cell suspensions.

Tetramers

H-$2D^b$ tetramers conjugated to phycoerythrin (PE) were prepared with mutant spectrin-β2 peptides and produced by the NIH Tetramer Core Facility (Emory University, Atlanta, GA).

Mutation specific RT-PCR and real-time RT-PCR

Total RNA from tumour cells was isolated by RNeasy Mini kit (Qiagen) and cDNA was synthesized from the total RNA using oligo (dT) primers and SuperScript II Reverse Transcriptase (Invitrogen). Real-time PCR specific for wild type spectrin-β2, mutant spectrin-β2 and GAPDH using the SYBR Green Mastermix kit (Applied Biosystems) were performed on ABI 7000. The primer sequences for used for mutant spectrin-β2 are 5′-GGTGAACCAGATTGCACT-3′ and 5′-TGTCCACCAGTTCTCTGAACT-3′.

Detection of mutation in *spectrin-β2* cDNA

The point mutation in the *spectrin-β2* gene creates a *PstI* restriction site (C*G*GCAG to C*T*GCAG, underlined italic letters indicate the site of mutation). To amplify *spectrin-β2* cDNA we used a forward primer (ACCCTGGCCCTGTACAAGAT) and reverse primer (TAGACTCGATGACCTTGGTCT). The PCR conditions used were 94°C for 2 min, followed by 35 cycles of 94°C for 30s, 55°C for 30s and 72°C for 30s. The PCR products were digested for 2 hrs at 37°C with *PstI* restriction enzyme, which cleaved mutant *spectrin-β2*, but not wild-type *spectrin-β2*, and generates a 200 bp fragment from cDNA. The products were resolved by electrophoresis on a 1.2% agarose gel and visualized by ethidium bromide staining.

Isolation of non-transformed cells from frozen primary d42m1 tumour biopsy

A frozen d42m1 tumour biopsy from the original d42m1 tumour was thawed and treated with 1 mg/ml type IA collagenase (Sigma) in HBSS for 2 hrs at room temperature. After filtration, single-cell suspensions were stained for 20 minutes at 4°C with 500 ng of Fc block (anti-CD16/32) and 200 ng of CD45-PE in 100 μL of staining buffer. Propidium iodide was added at 1 μg/mL immediately before sorting. A $CD45^{+}PI^{-}$ gate was used and the top 15% percent and the bottom 15% of gated events were collected using a FACSAria II (BD Biosciences). Sorted $CD45^{+}$ cells (host leukocytes) and $CD45^{-}$ cells (primary d42m1 tumour cells) were collected and genomic DNA as well as RNA was isolated to synthesize cDNA for 3730 sequencing to validate that the mutation calls detected by Illumina were somatic and tumour-specific.

Statistical Analysis

Samples were compared using an unpaired, two-tailed Student's *t* test, unless specified.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Jessica Archambault for expert technical assistance, Dr. Ted H. Hansen (Washington University) for providing MHC class I antibodies, Steve Horvath and Dr. Paul M. Allen (Washington University) for synthesizing MHC class I peptides, NIH Tetramer Core Facility for producing MHC class I tetramers, and Dr. Thaddeus S. Stappenbeck (Washington University) for technical help in recovering frozen tumour samples. We also thank Dr. Emil Unanue (Washington University), Dr. Paul M. Allen (Washington University) and Dr. Jack Bui (University of California at San Diego) for helpful criticisms and comments, all members of the Schreiber laboratory for helpful discussions, and the many members of The Genome Institute at Washington University School of Medicine, especially Dr. Li Ding for her insights into our analytical approaches. This work was supported by grants to R.D.S from the National Cancer Institute, the Ludwig Institute for Cancer Research, the Cancer Research Institute, and the WWWW Foundation, to E.R.M from the National Human Genome Research Institute. M.D.V. is supported by a pre-doctoral fellowship from the Cancer Research Institute. J.P.A. is supported by the Howard Hughes Medical Institute and the Ludwig Center for Cancer Immunotherapy, M.J.S. by the National Health and Medical Research Council of Australia (NH&MRC) and from the Association for International Cancer Research, and L.J.O. by the Ludwig Institute for Cancer Research and the Cancer Research Institute.

References

1. Shankaran V, et al. IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. Nature. 2001; 410:1107–11. [PubMed: 11323675]
2. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol. 2002; 3:991–8. [PubMed: 12407406]
3. Koebel CM, et al. Adaptive immunity maintains occult cancer in an equilibrium state. Nature. 2007; 450:903–7. [PubMed: 18026089]
4. Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ. Natural Innate and Adaptive Immunity to Cancer. Annu Rev Immunol. 2011; 29:235–271. [PubMed: 21219185]
5. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science. 2011; 331:1565–70. [PubMed: 21436444]
6. Futreal PA, et al. A census of human cancer genes. Nat Rev Cancer. 2004; 4:177–83. [PubMed: 14993899]

7. Chen AC, Herschman HR. Tumorigenic methylcholanthrene transformants of C3H/10T1/2 cells have a common nucleotide alteration in the c-Ki-ras gene. Proc Natl Acad Sci U S A. 1989; 86:1608–11. [PubMed: 2646640]
8. Tuveson DA, et al. Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. Cancer Cell. 2004; 5:375–87. [PubMed: 15093544]
9. Kirsch DG, et al. A spatially and temporally restricted mouse model of soft tissue sarcoma. Nat Med. 2007; 13:992–7. [PubMed: 17676052]
10. Ley TJ, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature. 2008; 456:66–72. [PubMed: 18987736]
11. Ding L, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010; 464:999–1005. [PubMed: 20393555]
12. Shah SP, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature. 2009; 461:809–13. [PubMed: 19812674]
13. Lee W, et al. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature. 2010; 465:473–7. [PubMed: 20505728]
14. Pleasance ED, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature. 2010; 463:184–90. [PubMed: 20016488]
15. Totoki Y, et al. High-resolution characterization of a hepatocellular carcinoma genome. Nat Genet. 2011; 43:464–9. [PubMed: 21499249]
16. Puente XS, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature. 2011; 475:101–5. [PubMed: 21642962]
17. Pleasance ED, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature. 2010; 463:191–6. [PubMed: 20016485]
18. Boon T, Coulie PG, Van den Eynde BJ, van der Bruggen P. Human T cell responses against melanoma. Annu Rev Immunol. 2006; 24:175–208. [PubMed: 16551247]
19. Segal NH, et al. Epitope landscape in breast and colorectal cancer. Cancer Res. 2008; 68:889–92. [PubMed: 18245491]
20. Nielsen M, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci. 2003; 12:1007–17. [PubMed: 12717023]
21. van der Bruggen P, et al. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science. 1991; 254:1643–7. [PubMed: 1840703]
22. Khong HT, Restifo NP. Natural selection of tumor variants in the generation of “tumor escape” phenotypes. Nat Immunol. 2002; 3:999–1005. [PubMed: 12407407]
23. Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoediting. Annu Rev Immunol. 2004; 22:329–60. [PubMed: 15032581]
24. Yewdell JW. Confronting complexity: real-world immunodominance in antiviral CD8+ T cell responses. Immunity. 2006; 25:533–43. [PubMed: 17046682]
25. Prehn RT, Main JM. Immunity to methylcholanthrene-induced sarcomas. J Natl Cancer Inst. 1957; 18:769–78. [PubMed: 13502695]
26. Old LJ, Boyse EA. Immunology of Experimental Tumors. Annu Rev Med. 1964; 15:167–86. [PubMed: 14139934]
27. Huang S, et al. Immune response in mice that lack the interferon-gamma receptor. Science. 1993; 259:1742–5. [PubMed: 8456301]
28. Muller U, et al. Functional role of type I and type II interferons in antiviral defense. Science. 1994; 264:1918–21. [PubMed: 8009221]
29. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–60. [PubMed: 19451168]
30. Koboldt DC, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics. 2009; 25:2283–5. [PubMed: 19542151]
31. Ning Z, Cox AJ, Mullikin JC. SSAHA: a fast search method for large DNA databases. Genome Res. 2001; 11:1725–9. [PubMed: 11591649]
32. Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–9. [PubMed: 19505943]

33. Robinson JT, et al. Integrative genomics viewer. Nat Biotechnol. 2011; 29:24–6. [PubMed: 21221095]
34. Felsenstein J. Phylogeny Inference Package. Cladistics. 1989; 5:164–166.
35. Ranganath S, et al. GATA-3-dependent enhancer activity in IL-4 gene regulation. J Immunol. 1998; 161:3822–6. [PubMed: 9780146]
36. Dighe AS, Richards E, Old LJ, Schreiber RD. Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFN gamma receptors. Immunity. 1994; 1:447–56. [PubMed: 7895156]

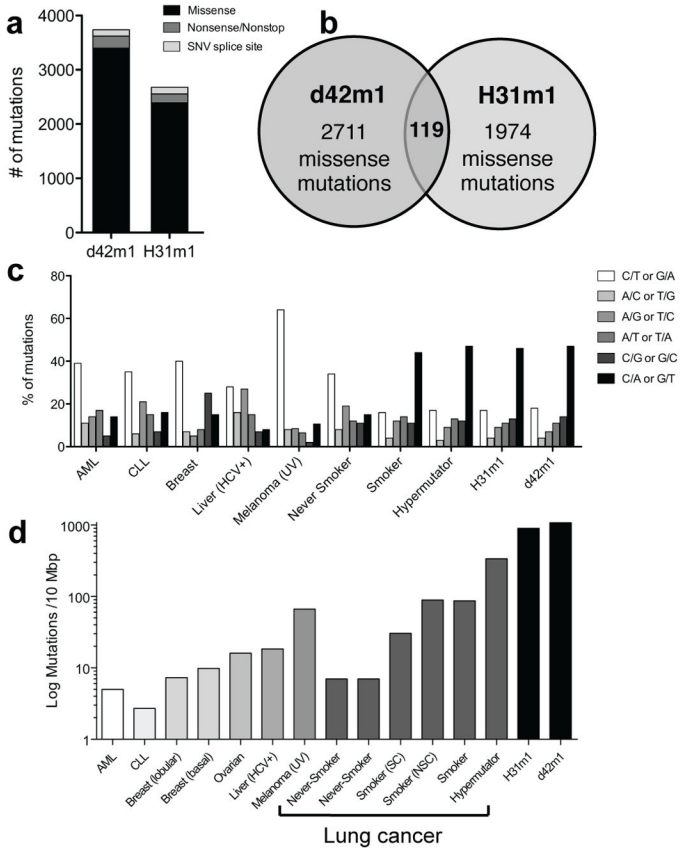

Figure 1. Unedited MCA-induced sarcomas, d42m1 and H31m1, genomically resemble carcinogen-induced human cancers

a, Number of non-synonymous mutations in d42m1 and H31m1 tumour cells as detected by cDNA CapSeq. **b**, Missense mutations compared between d42m1 and H31m1 that had at least 20x sequencing coverage. **c**, Spectrum of DNA nucleotide substitutions detected in d42m1 and H31m1 as compared to previously generated data from human cancers including acute myelogenous leukemia[10] (AML), chronic lymphocytic leukemia[16] (CLL), breast cancer (breast-lobular[12], breast-basal[11]), ovarian cancer (Mardis et al. manuscript in preparation), liver cancer (Hepatitis C Virus (HCV)-positive)[15], melanoma (ultraviolet (UV)-induced)[17], and lung cancers (non-small cell (NSC)[13], small cell (SC)[14], Never-Smoker, Smoker, and Hypermutator (Mardis et al. manuscript in progress). **d**, Mutation rate for d42m1 and H31m1 and human cancers described in **c**.

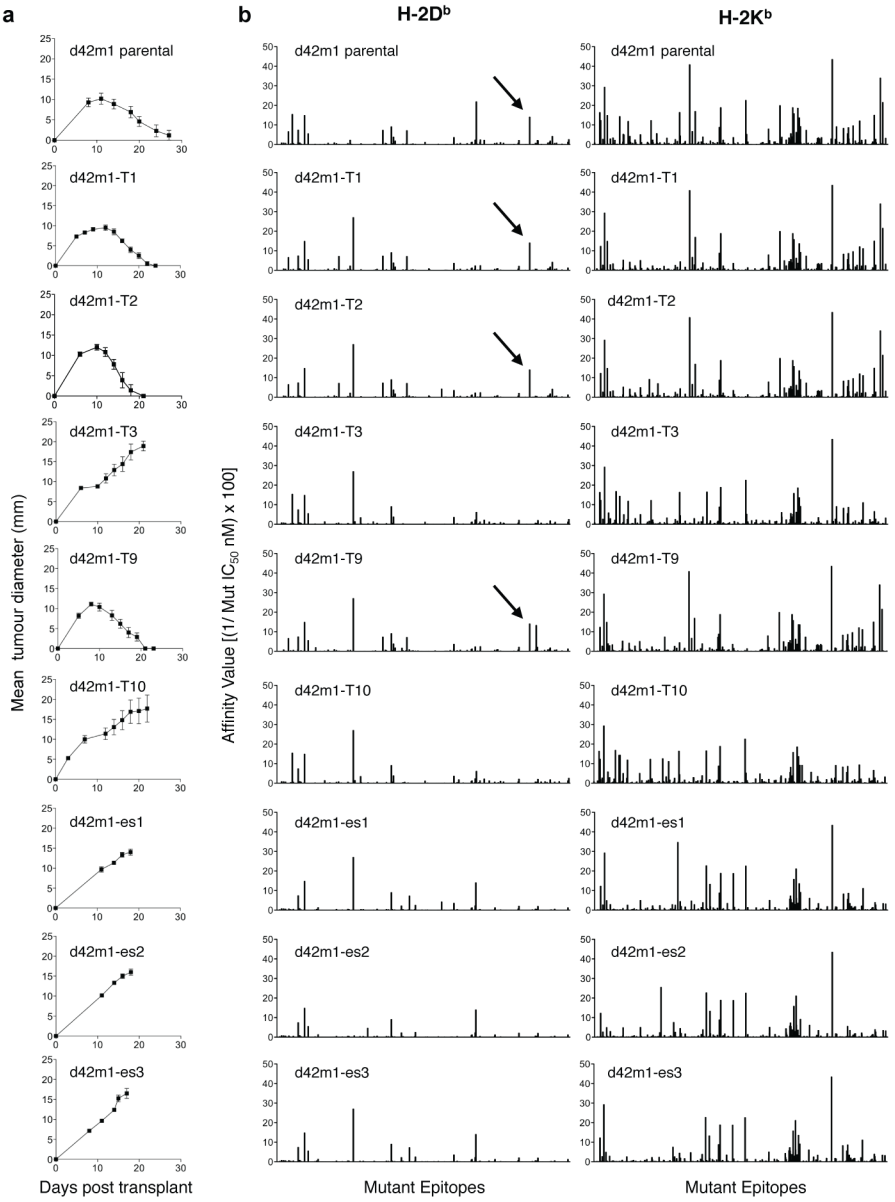

Figure 2. Affinity value profiles of predicted MHC class I epitopes from tumour-specific mutations

a, Growth of d42m1 parental cells, a representative sample of tumour clones, and three escape tumours following transplantation into WT mice (n=5, squares). Data are presented as average tumour diameter ± s.e.m. and are representative of three independent experiments. **b**, Missense mutations for each d42m1-related tumour examined in (**a**) were analyzed for potential MHC class I neoepitopes that bind either H-2D^b or H-2K^b. Predicted epitope binding affinities were ultimately expressed as "Affinity Values" (Affinity Value = 1/IC$_{50}$ X 100). Arrow is pointing to a H-2D^b epitope created by the R913L spectrin-β2 mutant.

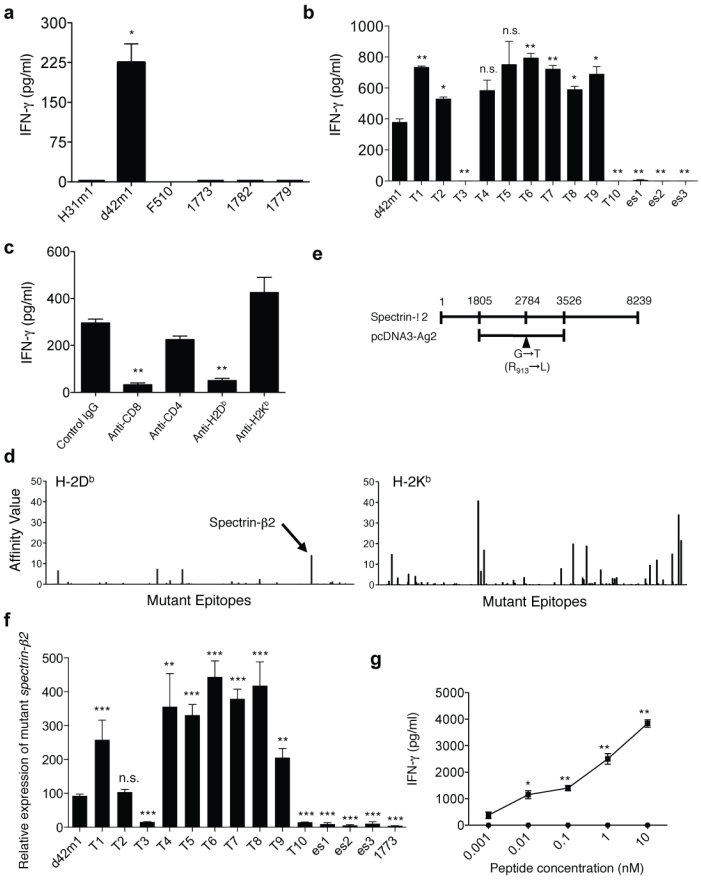

Figure 3. Identification of mutant spectrin-β2 as an authentic antigen of an unedited tumour

a, **b**, IFN-γ release by C3 CTLs against different unedited sarcomas (**a**) or against d42m1-related tumours (**b**). **c**, IFN-γ release by C3 CTLs is inhibited by mAbs that block CD8 and H-2D^b, but not CD4 or H-2K^b. **d**, MHC class I epitopes predicted to be shared in all of the regressor d42m1 tumours, but not in progressor d42m1 tumours. **e**, Representation of the cDNA clone that stimulated C3 CTLs encoding the spectrin-β2 R913L mutation. **f**, qRT-PCR for mutant *spectrin-β2* in d42m1-related tumours and 1773. **g**, IFN-γ release by C3 CTLs incubated with COS-D^b cells pulsed with WT (circles) or mutant (squares) spectrin-β2 peptides. Data are representative of three independent experiments. Samples were compared using an unpaired, two-tailed Student's *t* test (*p<0.05, **p<0.01, and ***p<0.001; n.s. is non-significant).

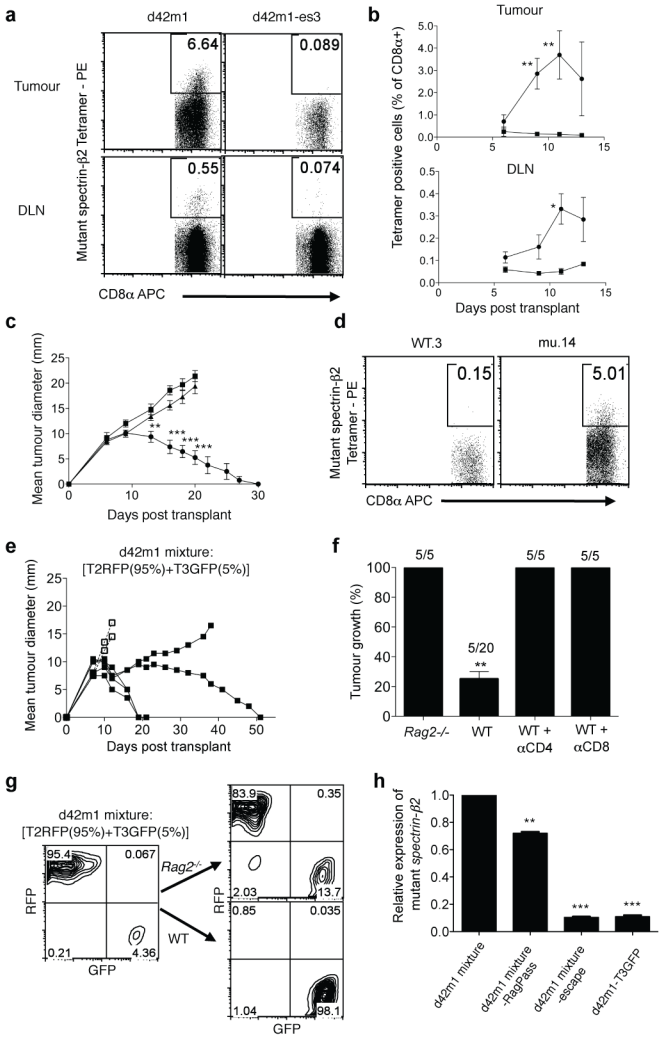

Figure 4. Mutant spectrin-β2 is a major rejection antigen of d42m1

a, Mutant spectrin-β2 specific $CD8^+$ T cells were detected by tetramer staining in tumours and DLNs from mice challenged with d42m1 parental cells, but not d42m1-es3 cells on day 11 post transplant. **b**, Quantification and kinetics of mutant spectrin-β2 tetramer staining in mice challenged with d42m1 parental cells (n=3, circles) or d42m1-es3 cells (n=3, squares). **c**, Growth of d42m1-es3 tumour cell clones transduced with WT (n=5, squares) or mutant spectrin-β2 (n=5, circles) and control d42m1-es3 cells (n=5, triangles) following transplantation (1 x 10^6 cells) into WT mice. Data are presented as average tumour diameter ± s.e.m. **d**, d42m1-es3 tumours reconstituted with WT (WT.3) or mutant spectrin-β2 (mu.14) were harvested at day 11 and $CD8\alpha^+$ T cells were stained with mutant spectrin-β2 tetramers. **e**, Growth of a mixture of d42m1-T2RFP (95%) and of d42m1-T3GFP (5%) following transplantation (1 x 10^6 total cells) into WT (n=5, solid lines, closed squares) or *Rag2*$^{-/-}$ (n=2, dashed lines, open squares) mice. **f**, Tumour outgrowth in *Rag2*$^{-/-}$ or WT mice treated or untreated with mAbs that deplete $CD4^+$ or $CD8^+$ T cells following challenge with 1 x 10^6 cells of a d42m1 mixture (95% d42m1-T2RFP and 5% d42m1-T3GFP). Data presented as percent tumour positive from 2–4 independent experiments (n=2–5 mice per group). **g**, **h**,

GFP and RFP expression (**g**) and mutant *spectrin-β2* expression (**h**) were analyzed in the d42m1-T2RFP/d42m1-T3GFP tumour cell mixture before injection and from tumours that grew out in *Rag2*$^{-/-}$ mice (RagPass) or escaped in WT mice by flow cytometry (**g**) or qRT-PCR (**h**). Data are representative of two independent experiments. Samples were compared using an unpaired, two-tailed Student's t test (*$p<0.05$, **$p<0.01$, and ***$p<0.001$; n.s. is non-significant).