bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Targeting the PD-1/PD-L1 pathway potentiates immunoediting to counterbalance neutral evolution in a mouse model of colorectal cancer

Mirjana Efremova[1], Victoria Klepsch[2], Pornpimol Charoentong[1], Francesca Finotello[1], Dietmar Rieder[1], Hubert Hackl[1], Natascha Hermann-Kleiter[2], Gottfried Baier[2], Anne Krogsdam[1], Zlatko Trajanoski[1]*

[1]Biocenter, Division of Bioinformatics, Medical University of Innsbruck, Innsbruck, Austria

[2]Division of Translational Cell Genetics, Medical University of Innsbruck, Innsbruck, Austria

*To whom correspondence should be addressed:

zlatko.trajanoski@i-med.ac.at

Zlatko Trajanoski, PhD

Biocenter, Division for Bioinformatics

Medical University of Innsbruck

Innrain 80, 6020 Innsbruck Austria

Email: zlatko.trajanoski@i-med.ac.at

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ABSTRACT (250 words):

Background: The cancer immunoediting hypothesis postulates a dual role of the immune system: protecting the host by eliminating tumor cells, and shaping the developing tumor by editing the cancer genome. However, to what extent immunoediting is shaping the cancer genome in common malignancies is still a matter of debate. Moreover, the impact of cancer immunotherapy with checkpoint blockers on modulating immunoediting remains largely unexplored.

Results: Here we employed a mouse model of colorectal cancer (CRC), next-generation sequencing, and computational analyses to elucidate the impact of evolutionary and immune-related forces on editing the tumor. We first carried out genomic and transcriptomic analyses of a widely-used model, MC38 cell line and show that this is a valid model for hypermutated and microsatellite-unstable CRC. Analyses of the data from longitudinal samples of wild type and immunodeficient RAG1 knockout mice transplanted with MC38 cells revealed that upregulation of checkpoint molecules and infiltration of Tregs are the major tumor escape mechanisms. Strikingly, the impact of neutral evolution on sculpting the tumor outweighed immunoediting by T cell dependent and T cell independent mechanisms in the progressing tumors. We also show that targeting the PD-1/PD-L1 pathway potentiated immunoediting and rendered tumors more homogeneous.

Conclusions:

In summary, our study demonstrates that neutral evolution is the major force that sculpts the tumor during progression, and that checkpoint blockade effectively enforces T cell dependent immunoselective pressure in this model. The results have important implication for basic research studies on the mechanisms of resistance to checkpoint blockade and for clinical translation.

Keywords: neutral evolution, immunoediting, checkpoint blocker

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

BACKGROUND

The concept of cancer immunosurveillance, i.e. that the lymphocytes can recognize and eliminate tumor cells was proposed almost 50 years ago [1], but the definitive work supporting the existence of this process was published 30 years later by the Schreiber lab [2]. In this seminal work an elegant experiment was carried out using a mouse model lacking recombination activating gene 2 (RAG2), i.e. a gene that encodes a protein that is involved in the initiation of V(D)J recombination during B and T cell development. RAG2 deficient mice are viable but fail to produce mature B or T lymphocytes [3]. RAG2 deficient mice developed sarcomas more rapidly and with greater frequency than genetically matched wild type controls and tumors derived from those mice were more immunogenic than those from wild type mice [2]. These findings led to the development of the refined cancer immunosurveillance concept: the cancer immunoediting hypothesis [4]. The cancer immunoediting postulates a dual role of the immunity in the complex interactions of the tumor and the host: the immune system, by recognizing tumor-specific antigens, can not only protect the host by eliminating tumor cells, but can also sculpt the developing tumor by editing the cancer genome and producing variants with reduced immunogenicity.

Cancer immunoediting is more difficult to study in humans, but clinical data from patients with severe immunodeficiencies is supporting the notion that this process exists also in humans [5]. Indirect evidence for the existence of immunoediting for some cancers was provided by calculating the ratio of observed and predicted neoantigens, i.e. tumor antigens derived from mutated proteins [6]. Using similar approach, we recently provided additional data that support the existence of immunoediting in microsatellite instable (MSI) colorectal cancer (CRC) [7]. However, as we recently showed in a pan-cancer genomic analyses, the composition of the intratumoral immune infiltrates is highly heterogeneous and changing during tumor progression [8], making it difficult to distinguish between genetic, immune, and other evasion mechanisms. Over and above these mechanistic questions on tumor progression, there is an urgent need to investigate cancer immunoediting also in the context of cancer immunotherapy. Cancer immunotherapy with checkpoint inhibitors like anti-CTLA-4 or anti-PD-1/-PD-L1 are showing remarkable clinical responses [9]. However, one of the biggest challenges is intrinsic resistance to immunotherapy and the development of resistant disease after therapy, i.e. acquired resistance to immunotherapy. As many patients with advanced cancers are now receiving immunotherapy, elucidating the role of cancer immunoediting as a potential mechanism of acquired resistance to immunotherapy [10] is of utmost importance.

Surprisingly, despite the recognition of the cancer immunoediting process and the widespread use of both, mouse models and next-generation sequencing (NGS) technologies, the impact of immunoediting on the cancer genome has not been well characterized. Cancer immunoediting was investigated in a mouse model of sarcoma using next-generation sequencing (NGS) of the tumor exome and algorithms for predicting neoantigens [11]. This sarcoma model showed that immunoediting can produce tumor cells that lack tumor-specific rejection antigens, but how this finding translates into common malignancies remained unclear. Later, two widely used tumor models, a CRC cell line MC38 and a prostate cancer cell line TRAMP-C1 were used to identify immunogenic tumor mutations by combining NGS and mass spectrometry [12]. However, since neither longitudinal samples of wild type and/or immunodefficient mice nor checkpoint blockade was applied, two major questions remained unanswered: 1) To what extent is T cell dependent immunoselection sculpting the cancer genome?, and 2) How is immunotherapy with checkpoint blockers modulating immunoediting? Quantitative evaluation of immunoediting during tumor progression and as well as following therapeutic intervention using checkpoint blockers could not only provide novel mechanistic insights, but might also inform immunotherapeutic strategies that could potentially be translated into the clinic.

We therefore designed a study to investigate immunoediting of an epithelial cancer genome using wild type and immunodeficient mice, NGS, and analytical pipelines to process and analyze the data. We first

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

characterized the genomic and transcriptomic landscape of the mouse CRC cell line MC38 and show that this cell line is valid model for hypermutated CRC. We then carried out experiments with wild type and immunodeficient RAG1$^{-/-}$ mice with transplanted tumors and analyzed longitudinal samples with respect to the genomic landscape and the immunophenotypes of the tumors. The results show the extent of immunoediting of the cancer genome in this model in relation to other selection processes. Finally, we performed experiments with anti-PD-L1 antibodies and show how the targeting the PD-1/PD-L1 pathway modulates immunoediting.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

RESULTS

Genomic, transcriptomic, and immunogenomic characterization of MC38 cell line

Functional studies on immunoediting require genetic tools and controls afforded by mouse studies. Since immunoediting has not been quantified using mouse epithelial cancers, we designed experiments with transplanted tumors using the murine MC38 cell line. The MC38 murine CRC cell line is a grade III adenocarcinoma that was chemically induced in a female C57BL/6 mouse and used since then as a transplantable mouse tumor model [13]. Several studies have shown that the cell line is immunogenic and can be used as a model for investigating anticancer immunity and immunotherapy [14], [15], [16], [17]. To characterize the genome and transcriptome of the MC38 cell line, we performed whole-exome sequencing, SNP array analysis, and RNA-sequencing (Figure 1A). We identified 7581 somatic mutations of which 3099 were nonsynonymous (2917 missense, 179 stop-gained, 3 stop-loss) and 240 indels (Figure 1B). Of the 7581 SNVs, the majority (6037) were transversions, of which most (3252) were C>A/G>T. Human hypermutated CRC tumors containing POLE mutations show increased proportions of C>A/G>T and T>G/A>C transversions [18], [19]. In contrast, it has been shown that the mouse CT26 cell line shows predominantly C>T/G>A SNVs [20], similar to the primary human non-hypermutated CRC tumors [21].

We investigated whether known CRC driver mutations are also present in MC38. We found missense mutations in TP53, BRAF, PTEN, and mutations in the TGF beta pathway (SMAD2, SMAD4, ACVR2A, TGFB2, but not TGFBR2). KRAS was not mutated and there was only one intron mutation in APC, however there was a truncating mutation in AXIN2 which is known to regulate ß-catenin in the Wnt signaling pathway. Recently discovered frequent mutations in SOX9, and ARID1A [21] were also present in the MC38 cell line. SOX9 is a transcription factor that inhibits Wnt signaling [22] and has a role in regulating cell differentiation in the intestinal stem cell niche [23], whereas ARID1A is involved in suppressing MYC transcription [24].

A large-scale genomic analysis of human colorectal samples [21] identified three subtypes of colorectal cancer: (1) microsatellite stable tumors (MSS), (2) tumors with microsatellite instability (MSI) due to a DNA mismatch repair system deficiency and (3) hypermutated group tumors that harbor mutations in the exonuclease (proofreading) domain of the DNA polymerases Pol δ (POLD1) and Pol ε (POLE). The MC38 data also showed mutations in two mismatch repair genes MSH3 and MUTYH, as well as in POLD1, indicating that the MC38 cell line is a valid model to study human MSI and hypermutated CRC. Both MSI and hypermutated CRC are reported to have better prognosis, higher infiltration of CD8+ T cells and respond well to checkpoint blockade therapy [25], likely due to the high number of neoantigens.

We then characterized copy number variants (CNV) of the MC38 cell line using exome sequencing and SNP arrays. The analysis of the copy number profiles inferred from the exome sequencing data using hidden Markov model algorithm (see Methods) and from the SNP array data were concordant and showed mostly diploid genome with some regions of amplifications and deletions (Figure 1A). We identified amplifications in the regions that contain the genes MYC and ERBB2. Finally, we carried out transcriptomic analysis of the MC38 cell line in comparison to normal skin tissues. The transcriptomic data was used to: 1) identify pathways that were up- or downregulated in the cell line, and 2) to identify expressed tumor antigens including neoantigens (identified using exome sequencing data and prediction algorithm as previously described [26]) and cancer-germline antigens (CGAs). The latter are tumor antigens that are considered to be tumor specific since these molecules are expressed only in germline cells and in tumor cells. Pathway enrichment analysis identified pathways related to cell cycle, DNA replication, DNA repair, and metabolism of nucleotides (Supplementary Figure S1).

With respect to the tumor antigens, we identified a large number of expressed neoantigens (Figure 1C) and expressed CGAs (Figure 1D), which provides evidence for the immunogenicity of this model. Of the 3096 amino acid changes (missense and stop codon) in MC38, 1529 neoantigens were predicted to strongly

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

bind to the C57BL/6 MHC class I molecules H2-K^b and H2-D^b with < 500 nM, and of these, 564 were in expressed genes. Additionally, several CGAs were highly expressed in MC38 including ATAD2, RQCD1, SPAG9, PBK, CTAGE5, CASC5, CEP55, which were also found to be expressed in the CT26 cell line [20]. It is noteworthy that these CGAs were also expressed in the skin samples.

Thus, the characterization of the genomic and transcriptomic landscape of the CRC MC38 cell line demonstrates its validity as a model for hypermutated and/or MSI colorectal cancer.

Upregulation of checkpoint molecules and infiltration of Tregs are the major tumor escape mechanisms in MC38 model of CRC

In our mouse model used to recapitulate the process of cancer immunoediting, MC38 cells were subcutaneously injected into wild type C57Bl/6 and immunodeficient $RAG1^{-/-}$ mice. The tumor growth was monitored regularly and samples were taken at predefined time points and subjected to detailed analysis using FACS, exome and RNA sequencing, and SNP array analysis (Figure 2A). As expected, the tumor growth was significantly accelerated in the $RAG1^{-/-}$ mice compared to the wild type mice (Figure 2B).

FACS analysis revealed infiltration of both innate and adaptive immune cells including CD8+ T cells, NK cells, and M1 macrophages in the wild type mice, that increased with time, although not significantly (Figure 2C). RNA expression profiles revealed higher expression of chemoattractant molecules such as CXCL9 and CCL5 in the wild type mice in comparison to the immunodeficient mice (Supplementary Figure S2). However, despite the presence of tumor infiltrating lymphocytes and the slower growth of the tumors, the adaptive immune system failed to eliminate the tumors. Tumors may utilize several mechanisms of escape such as antigen loss, upregulation of inhibitory molecules, downregulation of major histocompatibility molecules (MHC), or by creating an immunosuppressive environment. The CD8/Tregs ratio, which is a surrogate marker for suppressive tumor microenvironment, was higher in the skin samples compared to the tumor samples at day 23 (Figure 2C) suggesting that one escape mechanism in this model is the presence of immunosuppressive cells. The numbers of MDSCs and Tregs were comparable in both time points in the wild type mice, whereas the M2 macrophages were significantly reduced. The tumor progression in the wild type samples was associated with upregulation of immunoinhibitory genes, including PD-1, CTLA-4, TIM3, and LAG3 (Supplementary Figure S3). MC38 expressed low levels of PD-L1, whereas PD-L1 was slightly upregulated in $RAG1^{-/-}$ and more in the wild type mice. Analysis of the differentially expressed genes with respect to overrepresented pathways in the wild type vs $RAG1^{-/-}$ tumors showed upregulation of several immune processes related to activation of adaptive immune system response such as costimulation by the CD28, PD-1 signaling, antigen processing and presentation, NK cell mediated cytotoxicity, TCR signaling and interferon gamma signaling (Figure 2D and Supplementary Figure S4A). Downregulated pathways and GO terms included processes related to cell cycle, DNA replication and TNF signaling (Supplementary Figure S4B).

These data indicate that two tumor escape mechanisms are activated in this model: infiltration of immunosuppressive Treg cells and upregulation of inhibitory genes.

Neutral evolution outweighs T cell dependent and T cell independent immunoselection during tumor progression

Tumor progression is an evolutionary process under Darwinian selection [27], a characteristic that has been attributed as the primary reason of therapeutic failure, but also as a feature that holds the key to more effective control. At the time of detection, a tumor has acquired novel somatic mutations of which only a small subset (drivers) has an evolutionary advantage. The immune system exerts also an evolutionary pressure through a T cell dependent immunoselection process by acting on a tumor cell population that displays strong rejection antigens [11], and to some extent by T cell independent immunoselection through

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

M1 macrophages, IFNγ, and NK cells [28]. In addition to the ongoing evolutionary and immune-related clonal selection, recent study using a theoretical model demonstrated the occurrence of neutral evolution during tumor development [29]. According to this model, tumor heterogeneity in some cancers including CRC can be explained by neutral expansion and the accumulation of passenger mutations without selective sweeps.

To elucidate the impact of immunoselection on the progressing tumor, we used NGS to identify nonsynonymous mutations and MHC class I binding algorithm to predict the corresponding neoantigens. Analysis of exome sequencing data showed high number of mutations that were shared between the MC38 cell line and the two consecutive time points in both, wild type (2919) and RAG1$^{-/-}$ (2942) samples (Figure 3A). The number of newly generated mutations was about eight- to ten-fold higher than the number of potentially targeted mutations in both wild type and RAG1$^{-/-}$ mice (Figure 3A). For example, in the wild type sample at day 23 there were 386 newly generated mutations compared to 50 mutations shared only with MC38 cell line, which are potentially targeted in the wild type samples at day 46.

According to the cancer immunediting hypothesis, the immune system can sculpt the developing tumor by editing the cancer genome and thereby modifying the heterogeneity of the tumor: strong immunoediting would render tumors more homogeneous by eradicating immunogenic clones. In order to analyze the heterogeneity of the tumor during progression, exome sequencing data and SNP array data was used to estimate cancer cell fractions (CCF) of all point mutations and subsequently tumor heterogeneity. Analyses of the tumor heterogeneity did not reveal large differences during progression in both, wild type and RAG1$^{-/-}$ samples (Figure 3B). Strikingly, the analyses showed that the variant allele frequencies (VAF) of the majority of the mutations did not change with time in both the wild type and in the RAG1$^{-/-}$ mouse. On average, 95% of the mutations in the wild type and in the RAG1$^{-/-}$ samples did not change their VAF (Supplementary Table 1), suggesting that neutral evolution rather than Darwinian evolution is driving the tumor growth in this model.

We then characterized the neoantigens using exome sequencings data (to derive somatic mutations), RNA-sequencing data (for filtering expressed mutations) and an algorithm for predicting MHC binding (see Methods). In order to identify immunogenic mutations, we filtered the expressed neoantigens with the highest binding affinity (IC_{50}<500 nM). In a previous study with MC38 cell line seven mutant peptides were identified using mass spectrometry, of which two elicited a T cell response [12]. In our analysis six out the seven peptides were predicted and four of them were detectable from the RNA-expression data (Figure 3C). The large impact of neutral evolution was evident also in the Venn diagrams for the neoantigens (Figure 3D). The number of newly generated neoantigens was comparable in all the samples (126 and 129 for the wild type samples at day 23 and 46) and was much higher than the potentially lost or targeted neoantigens (17 in both wild type samples).

We then focused our analysis on the tumor samples taken at the same time point, day 23 for the wild type and RAG1$^{-/-}$ samples and considered neoantigens found both in the MC38 cell line and in at least one of the RAG1$^{-/-}$ tumors (Figure 3E). There were 530 neoantigens that were shared in the wild type, RAG1$^{-/-}$, and the MC38 cell line samples. About 3% of the neoantigens (17 out of 530) were detectable only in RAG1$^{-/-}$ tumors (Supplementary Table 2), out of which 16 were derived from mutations not detected or eliminated in the wild type tumors. Only one out of the 17 neoantigens was lost because of low expression. The small number of lost neoantigens imply that the impact of the T cell dependent immunoediting in this model is rather modest. Additionally, similar number of neoantigens (19) was detectable only in wild type tumors, suggesting that these neoantigens were edited by T cell independent mechanisms. Upregulation of genes related to NK cell mediated toxicity and IFN signaling further supports this observation (Supplementary Figure S5A). Analysis of the downregulated transcripts revealed genes related to DNA replication and cell cycle (Supplementary Figure S5B).

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Heterogeneity analysis showed that all samples including the MC38 cell line were similarly heterogeneous (Figure 3F and 3G). To infer how the clonal composition changes between samples, we used a Bayesian Dirichlet process to cluster clonal and subclonal mutations. The results showed that the clonal and subclonal clusters were on the leading diagonal of the plots indicating there was no change in the mutational profile and the clonal and subclonal composition between any two samples (Figure 3F). As expected, there was a large percentage of subclonal mutations both in the MC38 and in the individual tumor samples (Figure 3H).

Overall, the results suggest that neutral evolution outweighs both, T cell dependent and T cell independent immunoselection. Moreover, the temporal variation in subclonal architecture is largely determined by neutral evolution and to a small extent by Darwinian selection pressure.

Targeting the PD-1/PD-L1 pathway potentiates immunoediting and renders the tumors more homogeneous

We next investigated the impact of the strong immunological pressure induced by targeting the PD-1/PD-L1 axis on the cancer genome, on the neoantigen landscape, and on the tumor heterogeneity. It has been previously shown that MC38 responds to different immunotherapies [15, 17, 30, 31]. In order to identify neoantigens that would be potential targets of T cells activated by checkpoint blockade therapy, wild type C57Bl/6 mice were treated with antiPD-L1 antibodies or IgG2b antibodies as control. Treatment was started one day after tumor inoculation and then every three to four days. Samples from six tumors treated with anti-PD-L1 and six tumors treated with IgG2b were taken on day 14. Three samples of each group were used for exome sequencing, and three for RNA-sequencing.

Treatment with anti-PD-L1 antibodies reduced tumor growth in the treated mice by 65% compared to the controls (Figure 4A), which is in line with previous studies showing that MC38 responds well to PD-1/PD-L1 blockade therapy [32, 33]. This was reflected also from the RNA-sequencing data by a strong upregulation of IFNγ, perforin (PRF1), and granzyme A and B (GZMA and GZMB), as well as a number of immunomodulators and MHC molecules (Supplementary Figure S6). GO and pathway analysis showed upregulation of immune related processes such as PD-1 signaling, chemokine signaling, cytokine-cytokine receptor interaction, and NK cell mediated cytotoxicity (Figure 4B). Hence, blocking of the PD-1/PD-L1 pathway induces very strong adaptive and to a lesser extent innate mediated anti-tumor activity in this mouse model.

Analysis of the exome sequencing data showed a large fraction of mutations that were shared in all samples (2555) and 305 mutations that were detectable in the control sample and in the MC38 cell line, but absent from the anti-PD-L1 treated samples (Figure 4C). These mutations are potentially targeted by the immune system following blockade of the PD-1/PD-L1 pathway. A smaller number of mutations were detectable only in the anti-PD-L1 treated samples and the MC38 cell line (52). Overall, in the anti-PD-L1 treated samples the fraction of mutations resulting in expressed antigens was similar to the control sample (about 25%). Analysis of the peptides did not show any obvious patterns that could pinpoint rules defining immunogenicity of the mutations (Supplementary Table 3).

A major shift was observed in the fraction of expressed neoantigens from clonal origin in the anti-PD-L1 treated samples (Figure 4D). The fraction of clonal neoantigens was 8.8, 26.8, and 10.8 in the MC38, anti-PD-L1 treated, and the control tumors, respectively. Tumor heterogeneity analysis revealed more homogenous tumors undergoing treatment with checkpoint blockers compared to the control tumors and the MC38 cell line (Figure 4E). The same pattern can be observed in the 2-d density plots which show a shift of subclonal mutations in MC38 towards clonality in the anti-PD-L1 samples (Figure 4F), suggesting clonal expansion because of a selective advantage of subclones.

The analyses of this experimental data suggest that targeting the PD-1/PD-L1 pathway potentiates immunoediting and counterbalances neutral evolution in this mouse model. Moreover, this

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

immunotherapeutic intervention renders the tumors more homogeneous, which could possibly explain the development of resistance to checkpoint blockers.

Immunoediting and acquired resistance to PD-1 blockade in melanoma

In order to test the relevance of our findings in human cancer, we analyzed genomic data from a recent study on acquired resistance to PD-1 blockade in melanoma [34]. In this work pretreatment and relapse samples from four patients with metastatic melanoma, which were subjected to anti PD-1 blockade therapy, were analyzed by exome sequencing. Sequencing data showed that two of the tumors developed loss-of-function mutations in JAK1 and JAK2, respectively, which resulted into lack of response to IFNγ. The third tumor had a mutation in the antigen-presenting protein ß2M which prevented the immune system to recognize the tumor, whereas the fourth tumor had no defined mutations which could be associated with the relapse [34].

Using exome sequencing data, we analyzed the samples taken before therapy and after relapse with respect to the changes of the mutational landscape, the tumor heterogeneity and the clonal architecture. As can be seen in Figure 5A, large fraction of the mutations was detectable in baseline samples and in the relapse samples in all four cases, implicating that the bulk of the mutations were not efficiently targeted. Newly generated mutations ranged between 5% (case 1) and 33% (case 2). Mutations that were potentially immunoedited following PD-1 blockade, i.e. mutation detectable only in the baseline samples ranged between 4% (case 2) and 58% (case 3). Specifically, case 3 appeared to have strong immunoediting effects on the cancer genome.

With respect to the tumor heterogeneity, targeting the PD-1/PD-L1 pathway showed similar trend: relapsed tumors that acquired larger number of mutations became more heterogeneous (case 2 and case 4), whereas the tumor with lower number of acquired mutations became more homogeneous (case 3) (Figure 5B). The analysis for case 1 did not reveal changes in the tumor heterogeneity likely due to the high number of mutations in both, baseline and relapse sample (1045). Thus, in this case the impact of newly generated mutations on the tumor heterogeneity is rather small. The analyses of the clonal architecture revealed that in all tumors there was a loss of clonal mutations in the relapsed samples compared to the baseline, ranging from 1% (Case 2) to 24% (Case 3) (Figure 5C). Tumors that became more heterogeneous had increased number of subclonal mutations compared to the baseline (case 2 and case 4). In accordance with the immunoediting hypothesis, the relapsed sample showing strong immunoediting effect (case 3) had the largest decrease of both, clonal and subclonal mutations, and hence, was more homogeneous.

Overall, these results indicate that immunoediting can be associated with acquired resistance to PD-1 blockade in melanoma in specific mutational phenotypes. Targeting the PD-1 pathways in these phenotypes seems to broaden the T cell repertoire in a way that both, clonal and subclonal mutations are targeted and subsequently render the tumor more homogeneous. Hence, a clone which is resistant to immune attack will ultimately dominate the population.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

DISCUSSION

With the development of immunotherapies with checkpoint blockers as well as other immunotherapeutic strategies including therapeutic vaccines and engineered T cells [35], the interaction of the tumor and the immune system, and the question how the cancer genome is edited came into focus. Our understanding of the process of cancer immunoediting and its relevance for therapeutic intervention is still incomplete and requires comprehensive genomic analyses of longitudinal samples. Here we characterized for the first time the extent of immunoediting that tumors undergo during progression or as a consequence of the targeting the PD-1/PD-L1 axis. The quantification of cancer immunoediting using a mouse model of common cancer suggests several biological conclusions and has also important implications for clinical translation.

First, neutral evolution outweighs the effects of T cell dependent and T cell independent immunoselection on the cancer genome during tumor progression in this model. Neutral tumor evolution was only recently identified using a theoretical model that determines the expected distribution of subclonal mutations, and implies that a large number of new mutations are generated in ever smaller subclones, resulting in many passenger mutations that are responsible for intratumor heterogeneity, but have minimal or no impact on tumor expansion [29]. In this neutral evolution model all the mutations responsible for expansion are present in the founding cell and subsequent mutations are neutral. Analysis of the TCGA data showed that CRC and other cancers were dominated by neutral evolution whereas other cancers were not [29]. Hence, tumor adaptation in neutral cancers is driven by cancer cell plasticity rather than clonal selection. Here we provide an experimental evidence for the impact of neutral evolution based on the genomic and immunogenomic analyses of a widely-used mouse model. While the model we have used has certain limitations since it uses a cell line and it does not recapitulate evolution of the tumor as it occurs naturally, it provides further support for the neutral evolution.

Second, targeting the PD-1/PD-L1 pathway effectively potentiates immunoediting. Currently, we can only speculate on the underlying mechanisms driving the strong immune response. It has been previously shown that immunotherapy with anti-CTLA-4 antibodies leads to a significant number of newly detected T cell responses [36], which can be assigned to the broadening of the T cell receptor repertoire [37]. Our data support this model also in therapeutic strategy that blocks the PD-1/PD-L1 axis. The broadening of the T cell receptor repertoire might be one of the mechanisms of action of anti-PD-1 treatment and could explain the success of immunotherapy in a number of malignancies. Since CTLA-4 and PD-1 have differing immunological effects on circulating T cells, further mouse and human studies are necessary in order to test the hypothesis that broadening of the T cell receptor repertoire is a mechanism that potentiates immunoediting also in a therapeutic strategy that blocks the PD-1/PD-L1 axis.

And third, targeting the PD-1/PD-L1 pathway renders the tumors more homogeneous. While we did not carry out long-term experiments with different dosages and treatment schedules, one implication of this data is that the tumors might eventually become resistant to immunotherapy. We provide also additional data from a human study showing that in some cases tumors that relapse after PD-1 blockade are more homogeneous. Hence, cancer immunoediting represents one mechanism of acquired immunotherapy resistance.

Our findings have important implications for basic research studies on mechanisms of resistance to checkpoint blockade and for clinical translation. Most importantly, given that neutral evolution, T cell dependent immunoediting, and T cell independent immunoediting are sculpting the tumor, it is of utmost importance to carry out comprehensive genomic and immunogenomic analyses of pre- and post-treatment samples. Since conventional cancer therapy as well as cancer immunotherapy are altering the genomic landscape, clones that are resistant to therapy might arise and outcompete other clones. Thus, it is an imperative to characterize the used mouse models and the evolutionary forces driving the tumor in order to dissect the contribution of individual components on shaping the cancer genome.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Finally, our results have important implications also for clinical research. Given the fact that some cancers including CRC, stomach, lung, and bladder are dominated by neutral evolution [29], it will be important to study tumors over time to dissect out the impact of the immunological selection following checkpoint blockade. Neutral evolution theoretically generates greater tumor heterogeneity and hence, may facilitate adaptation after the initiation of immunotherapy. However, investigating evolutionary dynamics within human cancer is challenging since longitudinal observations are unfeasible and both, the genetic and immune landscape of cancer are highly dynamic and interwoven [8]. Use of new technologies such as single cell sequencing, as well as multi-region sequencing and better sequencing depth together with improved computational methods will provide better understanding of the interplay between the clonal architecture of a tumor and the antitumor response of the immune system. In this context, advances of organoid technologies and gene editing will open new avenues of research and ultimately lead to the development of precision immune-oncology.

CONCLUSIONS

In summary, we demonstrated that neutral evolution is the major force that sculpts the tumor during progression, and that checkpoint blockade effectively enforces T cell dependent immunoselective pressure in mouse model of CRC. Our study adds another layer of complexity of the tumor evolution and the dynamic nature of clonal selection driven by immunological and non-immunological mechanisms. An improved understanding of how the immune system shapes the tumor progression will be fundamental to improving response to immunotherapies and combating resistance, and will require comprehensive genomic and immunogenomic analyses of both, mouse models and human samples.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

METHODS

Mouse experiments

Wild type C57BL/6N mice *RAG1*$^{-/-}$ (B6.129S7-*RAG1*^{tm1Mom}/J) mice were purchased from Charles River. Mice were maintained under SPF conditions. All animal experiments were performed in accordance with the Austrian "Tierversuchsgesetz" (BGBI. Nr.501/1989 i.d.g.F. and BMWF-66.011/0061-II/3b/2013) and were approved by the Bundesministerium für Wissenschaft und Forschung (bm:wf).

5×10^4 MC38 colon carcinoma cells were injected subcutaneously (s.c.) into the left flank of 8- to 12-week-old female wild type or *RAG1*$^{-/-}$ mice. Tumor growth was monitored three times per week by measuring tumor length and width. Tumor volume was calculated according to the following equation: ½(length × width^2). For survival analysis, mice with tumors greater than the length limit of 15 mm were sacrificed and counted as dead.

Wild type mice were injected s.c. with $5x10^5$ MC38 melanoma cells and administered with 0.5mg of an anti-mouse PD-L1 (Clone10F.9G2; BE0101) or corresponding IgG2b (LTF-2; BE0090) control antibody (all from BioXCell, USA) every 3 to 4 days starting from day 1 of MC38 challenge according to [38]. Tumor growth was monitored as described above.

Immunophenotyping

Mononuclear infiltrating cells were isolated from both subcutaneous tumors and skin tissue at the indicated time points [38]. Briefly, tumor and skin tissues from sacrificed mice were prepared by mechanical disruption followed by digestion for 45 min with collagenase D (2.5 mg/ml; Roche, 11088858001) and DNase I (1 mg/ml; Roche, 11284932001) at 37°C. For skin tissue Liberase (5mg/ml; Roche, 5401020001) was added to the above described digestion mix. Digested tissues were incubated 5 min at 37°C with EDTA (0.5 M) to prevent DC/T cell aggregates and mashed through a 100-µm filter and a 40-µm filter. Cells were washed, and resuspended in PBS+2% FCS.

Tumor and skin infiltrating immune cells were incubated with FcR Block (BD Biosciences, 553142) to prevent nonspecific antibody binding before staining with appropriate surface antibodies for 30 min at 4°C, washed with PBS+2% FCS, and used for FACS analysis. For intracellular cytokine staining, cells were stimulated with 50 ng/ml Phorbol 12,13-dibutyrate (PDBu, Sigma, P1269), 500 ng ionomycin (Sigma, I0634), and GolgiPlug (BD Biosciences, 555029) for 4–5h. After fixation with the FoxP3 staining buffer set (eBiosciences, 00-5523) for at least 30 minutes at 4°C, cells were permeabilized with the fixation/permeabilization buffer (eBiosciences, 00-5523) and incubated with FcR Block (BD Biosciences, 553142) before staining with specific cell surface or intracellular marker antibodies. Data acquisition was performed on a LSR Fortessa cell analyzer (Becton Dickinson). Data analysis was conducted using the Flowlogic software (eBioscience, version 1.6.0_35).

The following antibodies were used for flow cytometry: CD4-V500 (BD, 560783), CD45-V500 (BD, 561487), CD8a-PerCP Cy5.5 (eBiosciences, 45-0081-82), CD3-PE (eBiosciences, 12-0031-83), CD11c-PerCP Cy5.5 (eBiosciences, 45-0114-80), CD11b-PE (BD, 557397), CD45-APC (eBiosciences, 17-0451-81), F4/80-PE-Cy7 (BioLegend, 123113), CD49b-FITC- (eBiosciences, 11-5971-81), Foxp3-FITC (eBiosciences, 11-5773-82), IFNγ-PE-Cy7 (eBiosciences, 25-7311-82), CD25-bv421 (BioLegend, 102034), Gr-1-APC (eBiosciences, 17-5931-81), MHCII-bv421 (BD, 561105).

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Exome- and RNA sequencing

DNA sequencing from the tumor, skin and MC38 cell samples was performed by exome capture using SureSelectXT Mouse All Exon capture probes (Agilent Technologies Österreich GmbH, Vienna, Austria) followed by sequencing with the Ion Proton™ System (Ion Torrent, Thermo Fisher Scientific). For RNA sequencing, total RNA was extracted, quality validated with the Agilent Bioanalyzer, and submitted to Lexogen 3' QantSeq library preparation, following the manufacturers instructions (Lexogen, Vienna Biocenter, Austria). Resulting libraries were sequenced with the Ion Proton™ System.

Exome-sequencing data analysis

The sequence reads were aligned to the mm10 reference genome with tmap (https://github.com/iontorrent/TS/tree/master/Analysis/TMAP) and preprocessed using GATK. Somatic point mutations were identified with Mutect [39] by comparing each tumor sample with the two skin samples and taking the intersection of the mutations. Insertions/deletions were called with Strelka [40] in the same way. The somatic mutations were annotated using the Ensembl Variant Effect Predictor tool [41]. Somatic copy number estimations were derived from exome-sequencing data using EXCAVATOR [42] by calculating $\log_2$ ratios between the read depth of the tumor and two germline skin samples using the "pool" mode. The estimated $\log_2$ ratios were then segmented by their novel heterogeneous shifting level model (HSLM). The CNAs identified using exome-sequencing data were concordant to those in the same samples by using Affymetrix SNP Array.

SNP arrays

Genome-wide copy number profiles of two wt samples (day 23 and day 46), two $RAG1^{-/-}$ samples (day 13 and day 23), all six aPD-L1 and IgG2b samples, MC38 and skin germline DNA were obtained using the Affymetrix Mouse Diversity Array. The genotyping analyses were carried out at Eurofins Genomics (Ebersberg, Germany) using the Affymetrix Mouse Diversity Array. The SNP arrays were processed, quantile-normalized, and median-polished using the Aroma Affymetrix CRMAv2 algorithm [43] together with 351 publically available Mouse Diversity Genotyping Array CEL files which were downloaded from the Center for Genome Dynamics at The Jackson Laboratory (http://cgd.jax.org/datasets/diversityarray/CELfiles.shtml). Copy number alterations (CNAs) for each probe were computed as $\log_2$-ratios between the probe signal intensities of each sample and the reference skin sample and then those ratios were segmented using the circular binary segmentation algorithm implemented in the R package DNAcopy [44].

Tumor heterogeneity

Normal contamination estimates were calculated using the homozygous point mutations in the cell line MC38. Considering that the purity of the cell line is 1, we checked the variant allele frequency of the homozygous mutations in MC38 in all the samples together with the estimated copy numbers of the corresponding region. The expected VAF of these mutations should be 1 in all samples assuming that there is no normal contamination and no new mutations appearing in the mouse samples at the same genomic position. As an estimate of the purity of the tumor, we took the mean of the VAF of those mutations found in a diploid region. These estimates were used to correct the mutation VAFs or copy number estimates in the rest of the analyses.

Mutations were filtered so that only mutations with at least 10 total reads and at least 5 alternative reads were considered. The CCF of each mutation was calculated using the approach of Yates et al [45]. Briefly, for each mutation the observed mutation copy number, n_{mut} (the fraction of tumor cells carrying a given mutation multiplied by the number of chromosomal copies at that locus) was calculated as:

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$n_{mut} = VAF\frac{1}{p}[pCN_t + CN_n(1-p)]$$

where VAF is the variant allele frequency of the mutation, p is the tumor purity, and CN_t and CN_n are the tumor and the normal locus specific copy number. Since mutations that are present of multiple chromosomal copies will have a mutation copy number higher than 1, we determined the number of chromosomes that the mutations is residing on. This was done so that for all mutations in amplified regions with a copy number of CNt, the observed fraction of mutated reads is compared to the expected fraction of mutated reads resulting from a mutation present on 1,2,3,…,CNt copies, considering a binomial distribution. The cancer-cell fraction was then calculated as the mutation copy number divided by the value of *C* with the maximum likelihood.

Mutations were defined as clonal if the CCF was > 0.95, and subclonal otherwise. Subclonal clusters of mutations were identified using a previously described statistical modelling of the distribution of clonal and subclonal mutations by a Bayesian Dirichlet process [46-48].

RNA-seq data analysis

The sequencing reads were first preprocessed through a quality control pipeline consisting of adapter removal with Cutadapt (DOI: http://dx.doi.org/10.14806/ej.17.1.200) and quality trimming with Trimmomatic [49] to remove bases with bad quality scores and reads shorter than 22 nucleotides. The quality trimmed reads were then mapped to the mm10 reference genome using a two-step alignment method; alignment with STAR [50] followed by alignment of the unmapped reads with Bowtie2. From the reads that mapped to multiple locations in the genome only the primary alignment was retained. Reads that mapped to ribosomal RNA locations in the genome were removed from further analysis using the *split_bam.py* script from the quality control package RSeQC [51]. Gene-specific read counts were calculated using HTSeq-count [52]. The R package DESeq2 [53] was used for differential expression analysis. The p-values were adjusted for multiple testing based on the false discovery rate (FDR) using the Benjamini-Hochberg approach.

Neoantigens and cancer-germline antigens

All possible 8-11 mer mutated peptides generated from all the nonsynonymous mutations (missense and nonsense) were used as an input to netMHCpan to predict their binding affinity to the C57BL/6 MHC class I alleles H-2K^b and H-2D^b. Amongst the candidate antigenic peptides, only the strong binders with binding affinity < 500 nM, and peptides arising from expressed genes were retained for further analysis. A mutation was considered expressed if the normalized counts of the corresponding gene were greater than 5.

The list of cancer-germline antigens (CGA) was downloaded from the Cancer-Testis database [54]. Their expression level was estimated using the normalized counts from DESeq2.

Human data

Mutational data from the melanoma patients [34] were provided by Dr. Antoni Ribas. Heterogeneity analyses were performed as described above.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Data accessibility

The mouse data sets were deposited in the GEO (GSE93018) and the Sequence Read Archive (SRP095725).

Statistical analysis

For comparison of two sample groups two-tailed unpaired Student's t-test was performed. Analysis and visualization of Gene Ontology terms and pathways associated with differentially expressed genes was performed using ClueGO [56]. A p-value of <0.05 was considered statistically significant: *p<0.5; **p < 0.01; ***p < 0.001.

SUPPLEMENTAL INFORMATION

Supplemental information includes 3 tables and 6 figures.

AUTHOR CONTRIBUTIONS

Z.T conceived the project. M.E. analyzed the mouse data. P.C. M.E., F.F., and H.H. analyzed the human data. D.R. organized and managed the data transfer and storage. A.K. carried out NGS. V.K. and N.H.-K. carried out the mouse experiments and the FACS analysis. M.E., G.B. A.K., and Z.T interpreted the results. M.E. and Z.T. wrote the manuscript.

ACKNOWLEDGMENTS

Mutational data from the melanoma patients were kindly provided by Dr. Antoni Ribas. This work was supported by the Austrian Science Fund (DK Molecular Cell Biology and Oncology), the Tiroler Standortagentur (Bioinformatics Tyrol), the European Commission (Horizon2020 project APERIM: Advanced bioinformatics tools for personalised cancer immunotherapy), and the Austrian National Bank (Jubiläumsfondsprojekt Nr. 16534).

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

REFERENCES

1. Burnet FM: **The concept of immunological surveillance.** *Prog Exp Tumor Res* 1970, **13:**1-27.
2. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, Schreiber RD: **IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity.** *Nature* 2001, **410:**1107-1111.
3. Shinkai Y, Rathbun G, Lam KP, Oltz EM, Stewart V, Mendelsohn M, Charron J, Datta M, Young F, Stall AM, et al.: **RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement.** *Cell* 1992, **68:**855-867.
4. Schreiber RD, Old LJ, Smyth MJ: **Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion.** *Science* 2011, **331:**1565-1570.
5. Teng MW, Galon J, Fridman WH, Smyth MJ: **From mice to humans: developments in cancer immunoediting.** *J Clin Invest* 2015, **125:**3338-3346.
6. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N: **Molecular and genetic properties of tumors associated with local immune cytolytic activity.** *Cell* 2015, **160:**48-61.
7. Mlecnik B, Bindea G, Angell HK, Maby P, Angelova M, Tougeron D, Church SE, Lafontaine L, Fischer M, Fredriksen T, et al: **Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability.** *Immunity* 2016, **44:**698-711.
8. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z: **Pan-cancer immunogenomic analyses reveals genotype-immunophenotype relationships and predictors of response to checkpoint blockade.** *Cell Rep* 2017, **18:**248-262.
9. Page DB, Postow MA, Callahan MK, Allison JP, Wolchok JD: **Immune modulation in cancer with antibodies.** *Annu Rev Med* 2014, **65:**185-202.
10. Restifo NP, Smyth MJ, Snyder A: **Acquired resistance to immunotherapy and future challenges.** *Nat Rev Cancer* 2016, **16:**121-126.
11. Matsushita H, Vesely MD, Koboldt DC, Rickert CG, Uppaluri R, Magrini VJ, Arthur CD, White JM, Chen YS, Shea LK, et al: **Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting.** *Nature* 2012, **482:**400-404.
12. Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, Franci C, Cheung TK, Fritsche J, Weinschenk T, et al: **Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing.** *Nature* 2014, **515:**572-576.
13. Corbett TH, Griswold DP, Jr., Roberts BJ, Peckham JC, Schabel FM, Jr.: **Tumor induction relationships in development of transplantable cancers of the colon in mice for chemotherapy assays, with a note on carcinogen structure.** *Cancer Res* 1975, **35:**2434-2439.
14. Woo SR, Turnis ME, Goldberg MV, Bankoti J, Selby M, Nirschl CJ, Bettini ML, Gravano DM, Vogel P, Liu CL, et al: **Immune inhibitory molecules LAG-3 and PD-1 synergistically regulate T-cell function to promote tumoral immune escape.** *Cancer Res* 2012, **72:**917-927.
15. Zippelius A, Schreiner J, Herzig P, Muller P: **Induced PD-L1 expression mediates acquired resistance to agonistic anti-CD40 treatment.** *Cancer Immunol Res* 2015, **3:**236-244.
16. Pajtasz-Piasecka E, Szyda A, Rossowska J, Krawczenko A, Indrova M, Grabarczyk P, Wysocki P, Mackiewicz A, Dus D: **Loss of tumorigenicity of murine colon carcinoma MC38/0 cell line**

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

after transduction with a retroviral vector carrying murine IL-12 genes. *Folia Biol (Praha)* 2004, **50:**7-14.

17. Zhang T, Sentman CL: **Cancer immunotherapy using a bispecific NK receptor fusion protein that engages both T cells and tumor cells.** *Cancer Res* 2011, **71:**2066-2076.

18. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, et al: **Signatures of mutational processes in human cancer.** *Nature* 2013, **500:**415-421.

19. Mouradov D, Sloggett C, Jorissen RN, Love CG, Li S, Burgess AW, Arango D, Strausberg RL, Buchanan D, Wormald S, et al: **Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer.** *Cancer Res* 2014, **74:**3238-3247.

20. Castle JC, Loewer M, Boegel S, de Graaf J, Bender C, Tadmor AD, Boisguerin V, Bukur T, Sorn P, Paret C, et al: **Immunomic, genomic and transcriptomic characterization of CT26 colorectal carcinoma.** *BMC Genomics* 2014, **15:**190.

21. Cancer Genome Atlas N: **Comprehensive molecular characterization of human colon and rectal cancer.** *Nature* 2012, **487:**330-337.

22. Topol L, Chen W, Song H, Day TF, Yang Y: **Sox9 inhibits Wnt signaling by promoting beta-catenin phosphorylation in the nucleus.** *J Biol Chem* 2009, **284:**3323-3333.

23. Mori-Akiyama Y, van den Born M, van Es JH, Hamilton SR, Adams HP, Zhang J, Clevers H, de Crombrugghe B: **SOX9 is required for the differentiation of paneth cells in the intestinal epithelium.** *Gastroenterology* 2007, **133:**539-546.

24. Nagl NG, Jr., Zweitzig DR, Thimmapaya B, Beck GR, Jr., Moran E: **The c-myc gene is a direct target of mammalian SWI/SNF-related complexes during differentiation-associated cell cycle arrest.** *Cancer Res* 2006, **66:**1289-1293.

25. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, et al: **PD-1 Blockade in Tumors with Mismatch-Repair Deficiency.** *N Engl J Med* 2015, **372:**2509-2520.

26. Hackl H, Charoentong P, Finotello F, Trajanoski Z: **Computational genomics tools for dissecting tumour-immune cell interactions.** *Nat Rev Genet* 2016, **17:**441-458.

27. Greaves M, Maley CC: **Clonal evolution in cancer.** *Nature* 2012, **481:**306-313.

28. O'Sullivan T, Saddawi-Konefka R, Vermi W, Koebel CM, Arthur C, White JM, Uppaluri R, Andrews DM, Ngiow SF, Teng MW, et al: **Cancer immunoediting by the innate immune system in the absence of adaptive immunity.** *J Exp Med* 2012, **209:**1869-1882.

29. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A: **Identification of neutral tumor evolution across cancer types.** *Nat Genet* 2016, **48:**238-244.

30. Deng L, Liang H, Burnette B, Beckett M, Darga T, Weichselbaum RR, Fu YX: **Irradiation and anti-PD-L1 treatment synergistically promote antitumor immunity in mice.** *J Clin Invest* 2014, **124:**687-695.

31. Ramakrishnan R, Assudani D, Nagaraj S, Hunter T, Cho HI, Antonia S, Altiok S, Celis E, Gabrilovich DI: **Chemotherapy enhances tumor cell susceptibility to CTL-mediated killing during cancer immunotherapy in mice.** *J Clin Invest* 2010, **120:**1111-1124.

32. Ngiow SF, Young A, Jacquelot N, Yamazaki T, Enot D, Zitvogel L, Smyth MJ: **A Threshold Level of Intratumor CD8+ T-cell PD1 Expression Dictates Therapeutic Response to Anti-PD1.** *Cancer Res* 2015, **75:**3800-3811.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

33. Homet Moreno B, Zaretsky JM, Garcia-Diaz A, Tsoi J, Parisi G, Robert L, Meeth K, Ndoye A, Bosenberg MW, Weeraratna AT, et al: **Response to programmed cell death-1 blockade in a murine melanoma syngeneic model requires costimulation, CD4, and CD8 T cells.** *Cancer Immunol Res* 2016.

34. Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, Torrejon DY, Abril-Rodriguez G, Sandoval S, Barthly L, et al: **Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma.** *N Engl J Med* 2016, **375:**819-829.

35. Schumacher TN, Schreiber RD: **Neoantigens in cancer immunotherapy.** *Science* 2015, **348:**69-74.

36. Kvistborg P, Philips D, Kelderman S, Hageman L, Ottensmeier C, Joseph-Pietras D, Welters MJ, van der Burg S, Kapiteijn E, Michielin O, et al: **Anti-CTLA-4 therapy broadens the melanoma-reactive CD8+ T cell response.** *Sci Transl Med* 2014, **6:**254ra128.

37. Robert L, Tsoi J, Wang X, Emerson R, Homet B, Chodon T, Mok S, Huang RR, Cochran AJ, Comin-Anduix B, et al: **CTLA4 blockade broadens the peripheral T-cell receptor repertoire.** *Clin Cancer Res* 2014, **20:**2424-2432.

38. Hermann-Kleiter N, Klepsch V, Wallner S, Siegmund K, Klepsch S, Tuzlak S, Villunger A, Kaminski S, Pfeifhofer-Obermair C, Gruber T, et al: **The Nuclear Orphan Receptor NR2F6 Is a Central Checkpoint for Cancer Immune Surveillance.** *Cell Rep* 2015, **12:**2072-2085.

39. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G: **Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.** *Nat Biotechnol* 2013, **31:**213-219.

40. Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK: **Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs.** *Bioinformatics* 2012, **28:**1811-1817.

41. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F: **The Ensembl Variant Effect Predictor.** *Genome Biol* 2016, **17:**122.

42. Magi A, Tattini L, Cifola I, D'Aurizio R, Benelli M, Mangano E, Battaglia C, Bonora E, Kurg A, Seri M, et al: **EXCAVATOR: detecting copy number variants from whole-exome sequencing data.** *Genome Biol* 2013, **14:**R120.

43. Bengtsson H, Wirapati P, Speed TP: **A single-array preprocessing method for estimating full-resolution raw copy numbers from all Affymetrix genotyping arrays including GenomeWideSNP 5 & 6.** *Bioinformatics* 2009, **25:**2149-2156.

44. Olshen AB, Venkatraman ES, Lucito R, Wigler M: **Circular binary segmentation for the analysis of array-based DNA copy number data.** *Biostatistics* 2004, **5:**557-572.

45. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, Jamal-Hanjani M, Wilson GA, Birkbak NJ, Hiley CT, et al: **Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade.** *Science* 2016, **351:**1463-1469.

46. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, Aas T, Alexandrov LB, Larsimont D, Davies H, et al: **Subclonal diversification of primary breast cancer revealed by multiregion sequencing.** *Nat Med* 2015, **21:**751-759.

47. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, Raine K, Jones D, Marshall J, Ramakrishna M, et al: **The life history of 21 breast cancers.** *Cell* 2012, **149:**994-1007.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

48. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, Dawson KJ, Iorio F, Nik-Zainal S, Bignell GR, et al: **Heterogeneity of genomic evolution and mutational profiles in multiple myeloma.** *Nat Commun* 2014, **5:**2997.

49. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data.** *Bioinformatics* 2014, **30:**2114-2120.

50. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013, **29:**15-21.

51. Wang L, Wang S, Li W: **RSeQC: quality control of RNA-seq experiments.** *Bioinformatics* 2012, **28:**2184-2185.

52. Anders S, Pyl PT, Huber W: **HTSeq--a Python framework to work with high-throughput sequencing data.** *Bioinformatics* 2015, **31:**166-169.

53. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biol* 2014, **15:**550.

54. Almeida LG, Sakabe NJ, deOliveira AR, Silva MC, Mundstein AS, Cohen T, Chen YT, Chua R, Gurung S, Gnjatic S, et al: **CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens.** *Nucleic Acids Res* 2009, **37:**D816-819.

55. Nielsen M, Lundegaard C, Blicher T, Lamberth K, Harndahl M, Justesen S, Roder G, Peters B, Sette A, Lund O, Buus S: **NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence.** *PLoS One* 2007, **2:**e796.

56. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J: **ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks.** *Bioinformatics* 2009, **25:**1091-1093.

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIGURE LEGENDS:

Figure 1. Genomic, transcriptomic, and immunogenomic characterization of mouse MC38 cell line. A) Circos plot showing (outer to inner): cytogenetic bands in black, grey and white. Track 1: DNA copy number log ratio values. Dark grey: diploid; Red: amplification (log ratio > 0.25); Blue: deletion (log ratio < -0.25). Track 2: Point mutations, plotted based on the variant allele frequency. Inner is frequency 0, outer is 100. Colours are purple (0-40), green (40-60), yellow (60-80), and pink (80-100). Track 3: Predicted MHC binding IC50 scores for the nonsynonymous mutations. Mutations with the highest binding affinity are coloured orange (IC50 < 50). Track 4: insertions and deletions coloured according to their allele frequency. B) Number of mutations in MC38 classified by type. C) Number of predicted and expressed neoantigens in MC38. D) Known germline antigens with the highest expression in MC38. The expression values are in normalized counts.

Figure 2. Tumor progression and tumor-infiltrating immune cells in wild type and RAG$^{-/-}$ mice. A) Schematic diagram of the experimental setup. B) Tumor growth curves of 5×10^4 MC38 cells inoculated into C57Bl/6 wild type (n=10) and RAG1$^{-/-}$ mice (n=4). The data are expressed as the mean ± SEM, statistically analyzed by a Student's t test. C) Tumor-infiltrating lymphocytes in wild type and RAG1$^{-/-}$ mice analyzed by flow cytometry * $P > 0.01$, ** $P < 0.01$, *** $P < 0.001$. D) Enriched functions and pathways of the significantly differentially expressed genes in tumors of the wild type vs RAG1$^{-/-}$ mice taken at day 23. The network is created using ClueGO. The pathways are functionally grouped based on the kappa score and the most significant term of each group is highlighted. The size of the nodes shows the term significance after Bonferroni correction.

Figure 3. Genomic and immunogenomic analyses of progressing tumors in wild type and RAG$^{-/-}$ mice. A) Shared nonsynonymous mutations between MC38, and wild type and RAG1$^{-/-}$ samples during progression. Mutations found in at least one sample from the same type are considered. B) Two-dimensional density plots showing the clustering of the cancer cell fractions of all mutations shared between two samples; increasing intensity of red indicates the location of a high posterior probability of a cluster. C) Immunogenic neoantigens that were experimentally validated in a previous work [12] and detected in this study. Asterisks show epitopes that were predicted but not expressed. D) Shared expressed neoantigens between MC38, and wild type and RAG1$^{-/-}$ samples during progression. Neoantigens found in at least one sample from the same type are considered. E) Shared nonsynonymous mutations and expressed neoantigens between MC38, and wild type and RAG1$^{-/-}$ samples at day 23. Mutations/neoantigens found in at least one sample from the same type are considered. F) Two-dimensional density plots showing the clustering of the cancer cell fractions of all mutations shared between two samples. G) Violin plots showing tumor heterogeneity estimated from the cancer cell fractions. H) Fractions of clonal and subclonal expressed neoantigens in MC38 and all tumor samples. Error bars represent standard error of the mean.

Figure 4. Genomic and immunogenomic impact of targeting the PD-1/PD-L1 axis. A) Tumor growth curve in C57Bl/6 wild type mice (n=11) inoculated with 5×10^5 MC38 cells and administered with 0.5 mg of an anti-mouse PD-L1 blocking antibody as immune checkpoint inhibitor is compared to tumor growth curve seen in mice (n = 12) injected with IgG2b control. The data are expressed as the mean ± SEM, statistically analyzed by Student's t test. B) Enriched functions and pathways of the significantly differentially expressed genes in tumors of the aPD-L1 therapy vs the IgG2b control samples. The network is created using ClueGO. The pathways are functionally grouped based on the kappa score and the most significant term of each group is highlighted. The size of the nodes shows the term significance after Bonferroni correction. C) Shared

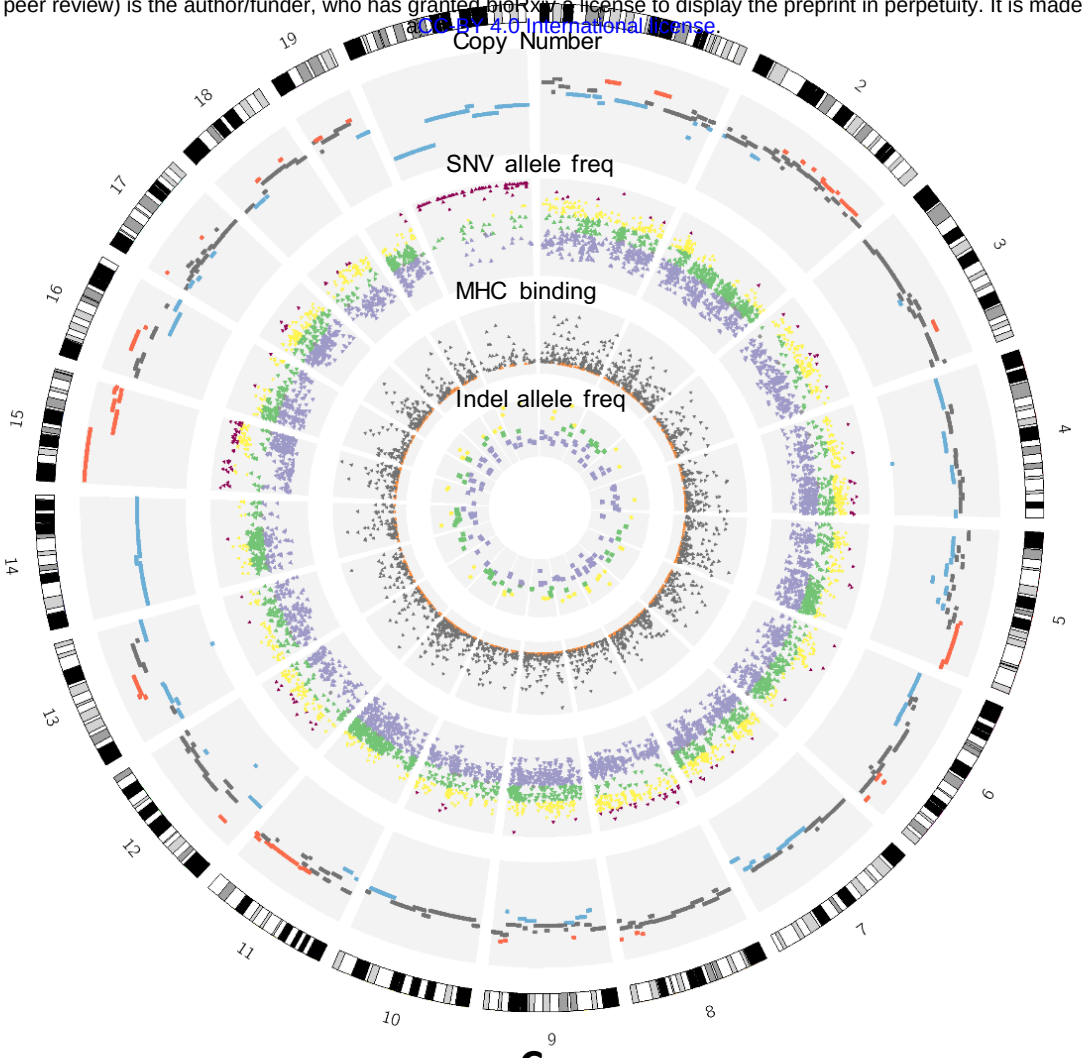
bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

nonsynonymous mutations and expressed neoantigens between MC38, and anti-PD-L1 treated and control samples. Mutations/neoantigens found in at least one sample from the same type are considered. D) Fractions of clonal and subclonal expressed neoantigens in MC38 and all tumor samples. Error bars represent standard error of the mean. E) Violin plots showing tumor heterogeneity estimated from the cancer cell fractions. F) Two-dimensional density plots showing the clustering of the cancer cell fractions of all mutations shared between two samples; increasing intensity of red indicates the location of a high posterior probability of a cluster.

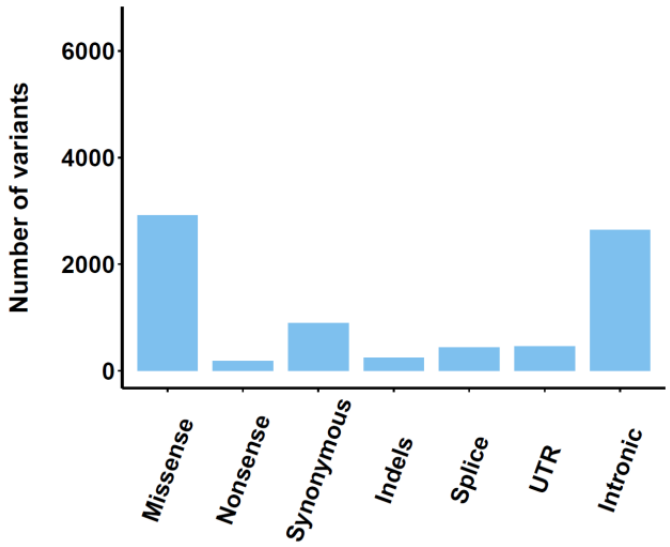
Figure 5. Immunoediting and acquired resistance to PD-1 blockade in melanoma. A) Number of mutations at onset of therapy and after relapse from a study by Zaretsky et al. [34]. B) Violin plots showing tumor heterogeneity estimated from the cancer cell fractions. C) Relative variation in the number of nonsynonymous mutations detected in the relapse samples compared to the baseline in the study by Zaretsky et al [34].

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A



B



C

Neoantigens
Predicted (<500nM) : 1529
Expressed: 564

D

Cancer germline antigens				
Gene	**MC38**	**MC38**	**skin**	**skin**
Atad2	2284	1977	610	660
Rqcd1	1289	1409	508	439
Spag9	1363	1315	1494	1865
Pbk	1083	1065	378	382
Ctage5	921	962	914	1251
Odf2	892	764	510	382
Casc5	673	652	296	422
Tmeff1	535	634	238	167
Dcaf12	540	619	907	818
Cep55	272	364	54	126

Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

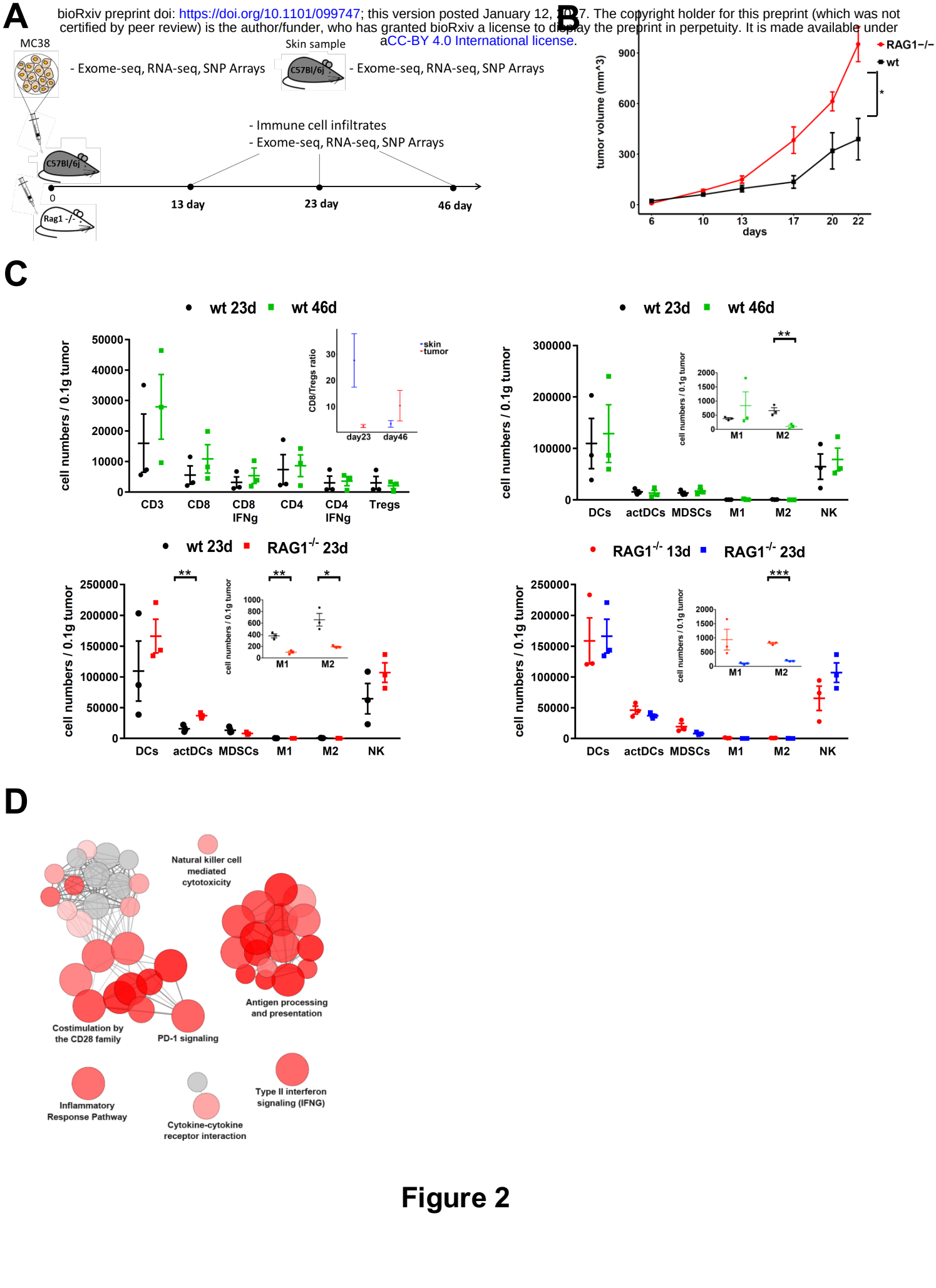
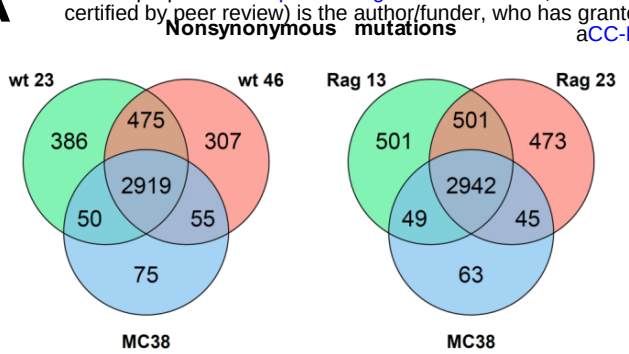


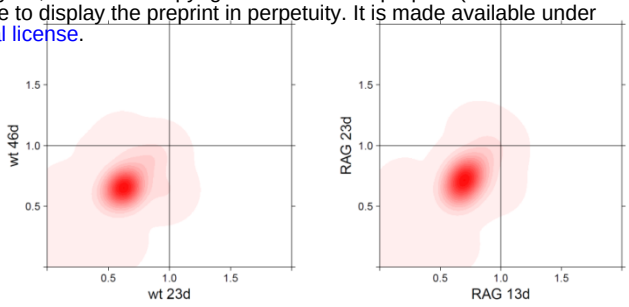
Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A



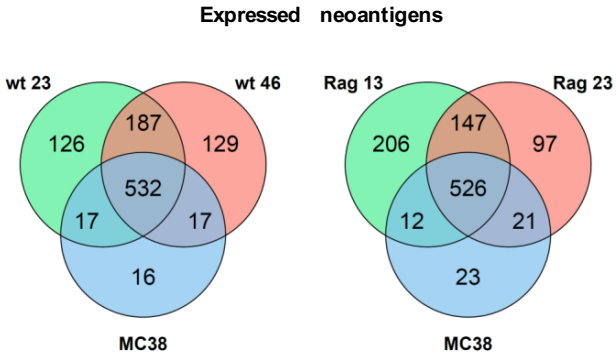
B



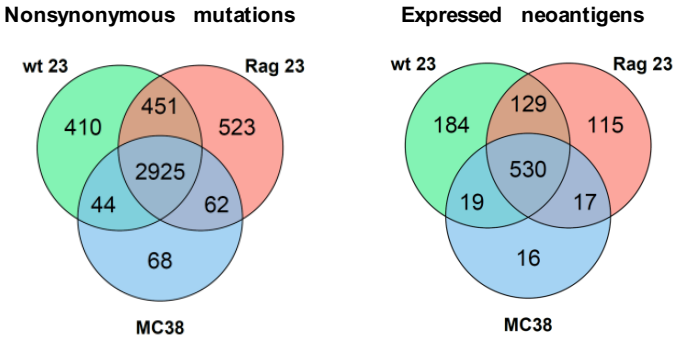
C

Gene	Peptide	MHC allele	IC_{50}(mut)	IC_{50}(WT)
Dpagt1	SIIVFNL[V/**L**]	H-2K^b	25	187
Reps1	AQL[P/**A**]NDVVL	H-2D^b	39	122
Adpgk	ASMTN[R/**M**]ELM	H-2D^b	4	6
Cpne1*	SSP[D/**Y**]SLHYL	H-2K^b	218	770
Aatf	MAPIDHT[A/**T**]M	H-2D^b	101	261
Med12*	SSSVLFE[D/**Y**]M	H-2K^b	314	1518

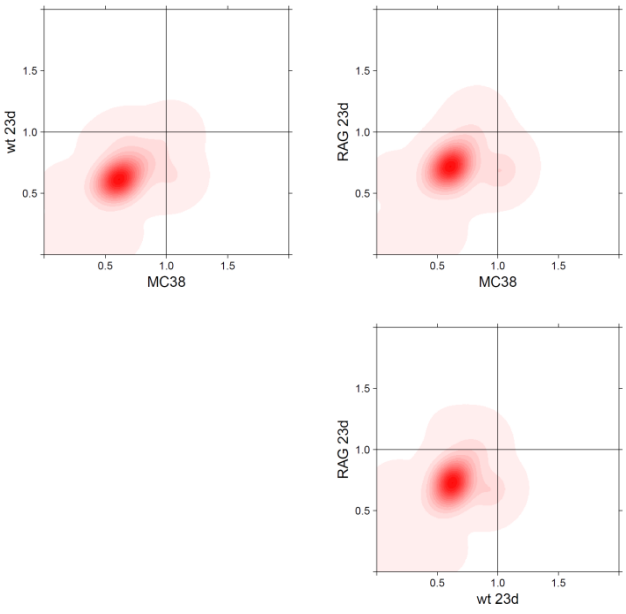
D



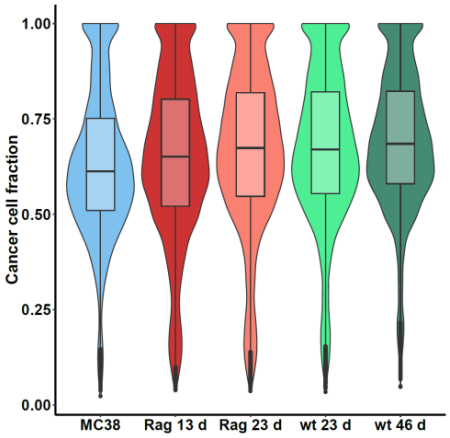
E



F



G



H

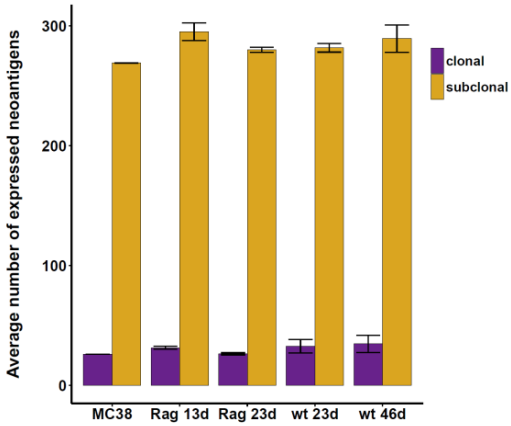
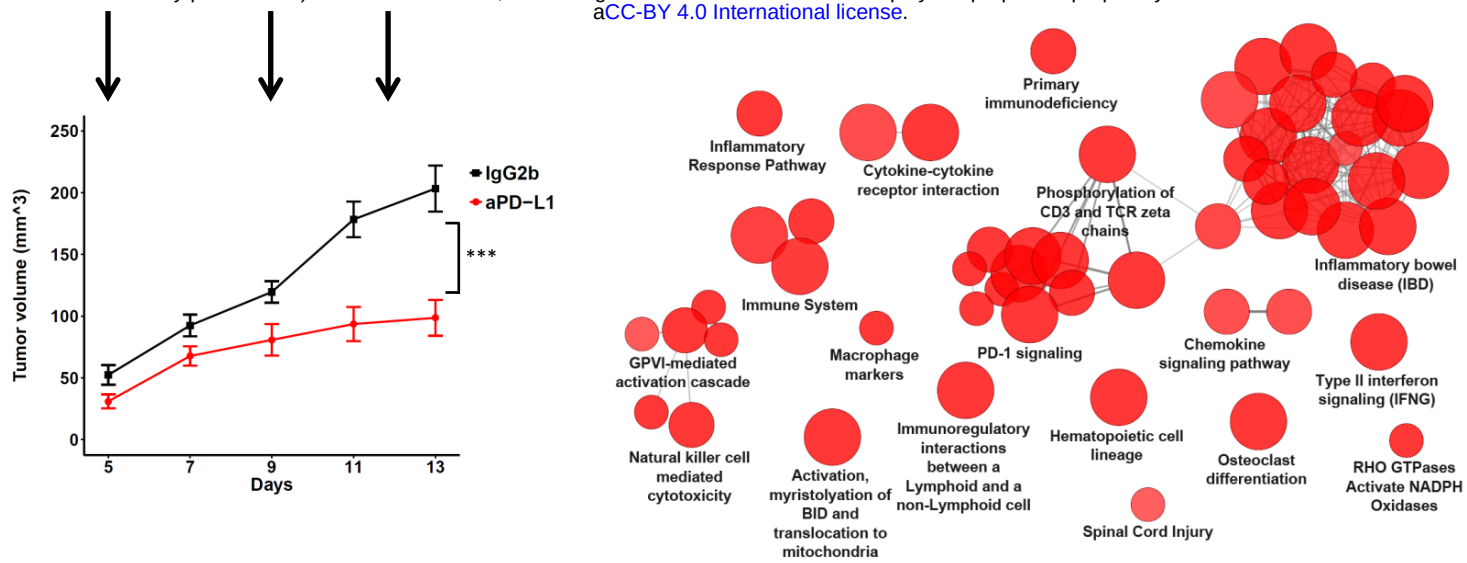
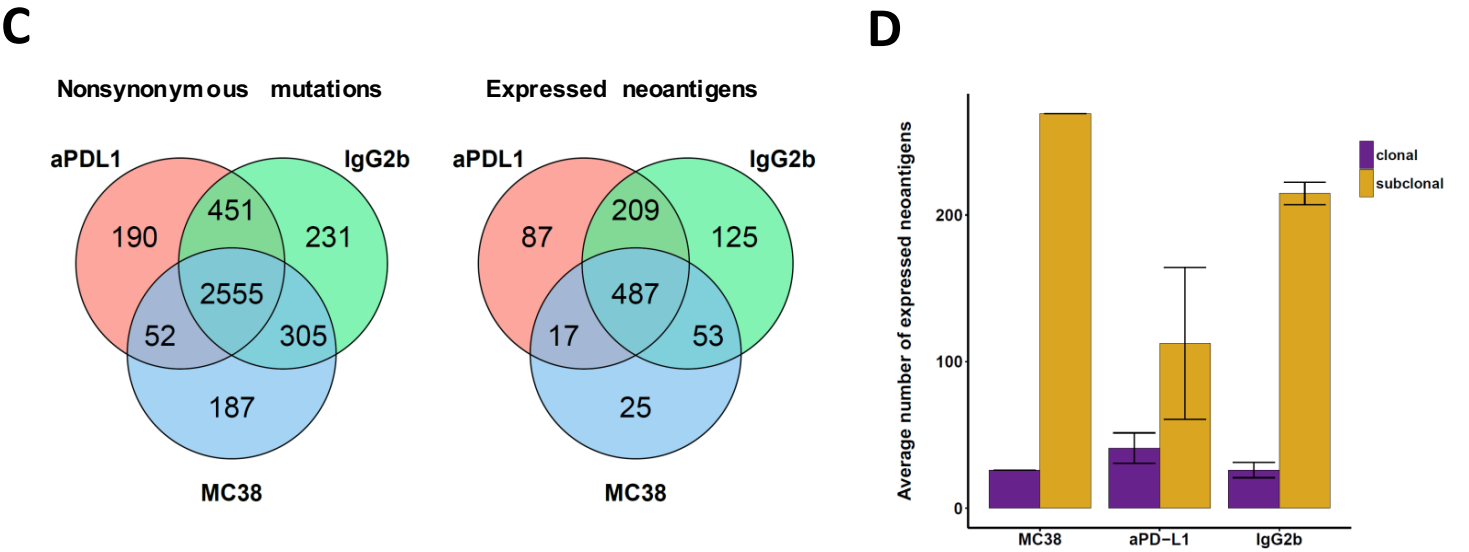


Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.





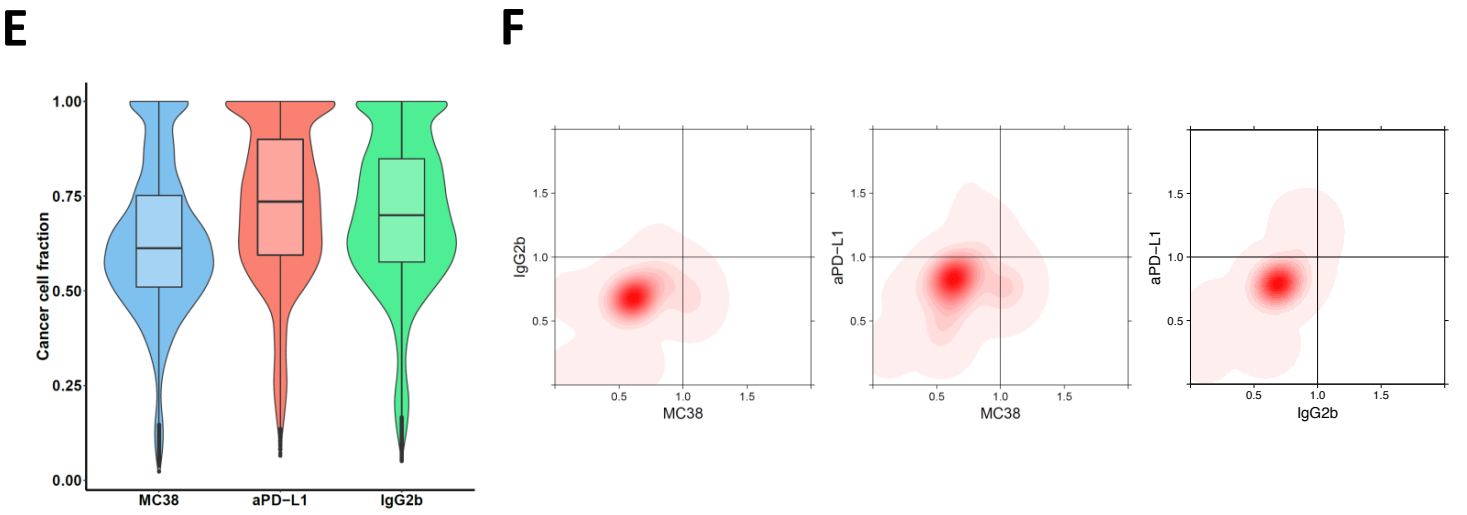


Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/099747; this version posted January 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

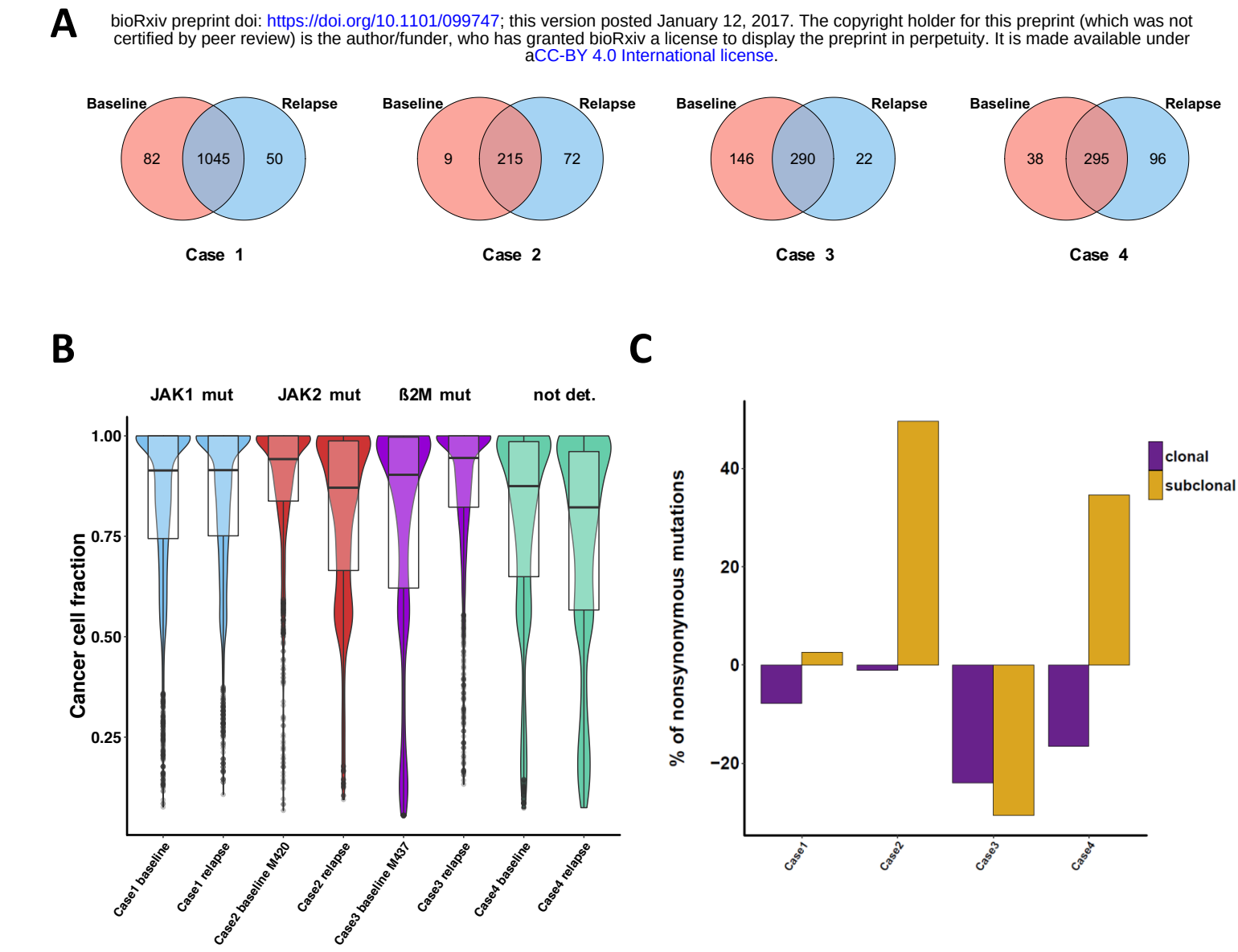


Figure 5