Multiple Myeloma associated DIS3 mutations drive AID-dependent IgH Translocations

Tomasz M. Kuliński[1, 2], Olga Gewartowska[1, 2, 3], Mélanie Mahé[4], Karolina Kasztelan[1,2], Janina Durys[2], Bertrand Séraphin[4], and Andrzej Dziembowski[1, 2, 3,*]

Affiliations:

[1] Laboratory of RNA Biology, International Institute of Molecular and Cell Biology, Trojdena 4, Warsaw, 02-109, Poland.

[2] Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawinskiego 5a, Warsaw, 02-106, Poland.

[3] Faculty of Biology, University of Warsaw, Pawinskiego 5a, Warsaw, 02-106, Poland.

[4] Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France; Centre National de Recherche Scientifique, UMR 7104, Illkirch, France; Institut National de Santé et de Recherche Médicale, U964, Illkirch, France; Université de Strasbourg, Illkirch, France.

*Correspondence to: adziembowski@iimcb.gov.pl

ABSTRACT

DIS3 is mutated in multiple myeloma (MM), but the mechanism by which oncogenesis occurs specifically in B cell lineage remains to be established. Somatic DIS3 variants have substitutions enriched around its RNB domain, which in a dominant negative way, inactivate or reduce the exoribonucleolytic activity of this enzyme responsible for nuclear RNA degradation. Here using knock-in mice with a clinical Dis3 G766R variant, we demonstrate a B cell-specific mutagenic effect that induces aberrant chromosomal translocations, increasing the incidence of plasmacytoma, a mouse model of early-stage MM. Dis3 G766R ¯dependent translocations display characteristics typical to aberrant activation-induced deaminase (AID) activity sites. Indeed, analysis of MM clinical samples revealed that in MM driver genes, DIS3 alleles lead to increased AID-dependent DNA lesions. Mechanistically, mutated DIS3 accumulates on chromatin-associated RNA substrates, including aberrant AID action sites, fostering oncogenic chromosomal rearrangements during immunoglobulin class switch recombination. In conclusion, MM DIS3 mutations lead to a gain-of-function phenotype and drive MM development, enhancing driver translocations.

INTRODUCTION

Multiple myeloma (MM), the second most prevalent hematologic malignancy among adults, arises during plasma cell differentiation. Plasma cells are antibody-producing, bone marrow-residing, terminally differentiated B lymphocytes (Manier *et al*, 2017; Kumar *et al*, 2017). These cells have already undergone B cell activation coupled with activation-induced deaminase (AID) expression and activity. AID enables antibody rearrangements and diversification by evoking several rounds of somatic hypermutation (SHM) and class switch recombination (CSR) (Chaudhuri *et al*, 2003; Chen & Wang, 2019). There are two main classes of MM: hyperdiploid (HRD), characterized by trisomies of several chromosomes, and non-HRD, characterized by chromosomal translocations (Manier *et al*, 2017; Castaneda & Baz, 2019).

The majority of translocations in MM involve the immunoglobulin heavy chain (IGH), which is one of the most heavily transcribed loci in plasma cells located on chromosome 14. These translocations lead to the overexpression of various oncogenes, including FGFR3, MMSET (NSD2), CCND3, CCND1, MYC, and MAF (Affer *et al*, 2014; Castaneda & Baz, 2019; Manier *et al*, 2017; Lohr *et al*, 2014; Chapman *et al*, 2011). Similarly, pristane-induced plasmacytoma, a murine model of MM, is driven by translocations to the Myc locus, which are strictly dependent on the aberrant activity of AID (Ramiro *et al*, 2004; Takizawa *et al*, 2008). Clinical data suggest that AID-induced point mutations in a subset of MM driver genes also play a crucial role in the development of human MM (Maura *et al*, 2020). Still, mechanistically, this process is poorly understood.

DIS3, a gene involved in RNA metabolism (Dziembowski *et al*, 2007), is frequently mutated (Walker *et al*, 2018; Lohr *et al*, 2014; Chapman *et al*, 2011) in MM. *DIS3* encodes a catalytic subunit of the major eukaryotic nuclear ribonuclease exosome involved in the processing of stable RNA species, such as ribosomal RNA (rRNA), and clearing pervasive transcription products, such as promoter upstream transcripts (PROMPTs) and enhancer RNA (eRNA) (Dziembowski *et al*, 2007; Tomecki *et al*, 2014; Szczepińska *et al*, 2015). DIS3 mutations were assumed to alter posttranscriptional gene expression regulatory programs leading to enhanced cell proliferation or cell death (Ohkura *et al*, 1988; Tomecki *et al*, 2014; Snee *et al*, 2016; Mukarami *et al*, 2007; Segalla *et al*, 2015). However, as we described in the accompanying paper, MM-associated DIS3 allies, which inhibit or, in case of recurrent mutations (D479, D488, and R780), completely abolish the exoribonucleolytic activity of this enzyme, have dominant-negative effects (Kuliński *et al*, 2023 submitted). Furthermore, they arise early during the MM evolution, presumably driving MM, but are counter-selected at the later stages of the disease because of their inhibitory effects on cell proliferation (Kuliński *et al*, 2023 submitted).

DIS3 mutations are specific to non-HRD MM (Lohr *et al*, 2014; Lionetti *et al*, 2015; Todoerti *et al*, 2021) and do not impact other cancer types. Moreover, DIS3 is currently the only subunit of the RNA exosome complex linked to cancer, even though depletions of exosome catalytic subunits in various model systems have some mutagenic effects (Domingo-Prim *et al*, 2019; Marin-Vicente *et al*, 2015; Laffleur *et al*, 2022; Dziembowski *et al*, 2007; Milbury *et al*, 2019). Actually, dysfunctions of nuclear proteins involved in RNA 3' end processing commonly leads to DNA damage (Skaar *et al*, 2015; Cevher *et al*, 2010; Kamp *et al*, 2022) without notable examples of their mutations causing cancer (Murphy & Kleiman, 2020). The exosome and, more recently DIS3, in addition to their essential role in the physiology of every cell type, were proposed to target AID to DNA during CSR and SHM in B cells (Laffleur *et al*, 2021). AID is expressed only in activated B cells, where it deaminates cytosine residues in DNA, leading to the production of uracil and generating G:U mismatches. Those recruit DNA repair machinery that removes uracil residues, which leaves lesions that can be filled in by error-prone DNA polymerases, resulting in point mutations/indels (in SHM). Mismatches on both DNA strands lead to double-stranded DNA breaks, followed by non-homologous end joining and large deletions (during CSR). Targeting AID, necessary for SHM and CSR (Revy *et al*, 2000; Muramatsu *et al*, 2000), is highly complex process involving transcription, rapid RNA degradation, and many additional protein factors (Hwang *et al*, 2015). Target regions are generally GC-rich, contain tandemly repeated DGYW sequence motifs, have the potential to form G-quadruplexes, and are prone to RNA polymerase stalling and the formation of RNA-DNA hybrids (R-loops) (Carpenter *et al*, 2010; Pefanis *et al*, 2015; Yu *et al*, 2003; Shinkura *et al*, 2003; Chaudhuri *et al*, 2003; Qiao *et al*, 2017; Zarrin *et al*, 2004). *In vitro*, AID predominantly targets the non-template strand during transcription. *In vivo*, AID mutates both DNA strands, necessitating a special mechanism involving the exosome to displace the nascent RNA (Basu *et al*, 2011; Pavri *et al*, 2010). Full mechanistic details have not yet been established, and the effect of MM-associated DIS3 point mutations in SHM and CSR has not been experimentally explored.

In this study, we generated a transgenic knock-in mouse model of DIS3-dependent MM and analyzed genomes of clinical samples from MM patients to investigate the mechanism of DIS3-mediated carcinogenesis. We found that DIS3 mutations lead to the accumulation of the mutant protein at sites of AID activity, driving oncogenic chromosomal translocations. At the same time, in contrast to DIS3 KO, MM DIS3 variants do not abrogate CSR.

RESULTS

Heterozygous MM-associated DIS3^{G766R} mutation has a mild growth inhibitory effect in mice

Analysis of clinical samples and MM cell lines in a study accompanying this paper did not reveal the oncogenic mechanism of DIS3 mutations (Kuliński *et al*, 2023 submitted). Moreover, attempts to generate mouse lines with recurrent MM DIS3 alleles failed because of strong dominant negative effects leading to early embryonic lethality (Kuliński *et al*, 2023 submitted). Therefore we decided to introduce a milder clinically relevant DIS3 variant. We constructed a knock-in mouse line with MM-specific DIS3^{G766R} mutation, which inhibit but does not inactivate the exoribonucleolytic activity of DIS3 (Chapman *et al*, 2011). This change has a predicted high probability of pathogenicity (Figure EV1A) and leads to DIS3 stalling on structured RNA substrates (Tomecki *et al*, 2014, 2010).

The homozygous Dis3^{G766R} allele constructed in a C57BL/6 genetic background is embryonically lethal (Figure EV1B, C), similar to constitutive DIS3 knockout. However, in the heterozygous state, no developmental phenotypes or hematologic abnormalities were observed (Figure EV1D, E). Notably, unlike heterozygous knockout, which exhibited no indications of a growth phenotype, Dis3$^{G766R/+}$ mice were smaller compared to their WT littermates (Figure 1A, Figure EV2A-D). Consistent with the inhibitory effect of MM DIS3 variants on cell growth, mouse embryonic fibroblasts (MEFs) isolated from heterozygous Dis3^{G766R} mice proliferated slower and entered senescence at a lower number of passages (Figure 1B). At the same time, there were no significant alterations of the annotated transcriptome compared with wild-type (WT) MEFs (Figure 1C, Dataset EV1), as revealed by the RNA-sequencing (RNA-seq) experiment. There was also no indication of transcriptional responses to DNA damage contrasting previously shown effects of DIS3 depletions (Gritti *et al*, 2022) or MM cell line with homozygous DIS3 mutation (Kuliński *et al*, 2023 submitted).

Finally, Dis3$^{G766R/+}$ MEFs did not have a higher frequency of spontaneous immortalization (Figure 1D), arguing against a generalized tendency toward cancerous transformation.

MM-specific DIS3^{G766R} mutation drives murine plasmacytoma

To check the potential general predisposition of Dis3$^{G766R/+}$ mice to develop cancer, 33 individuals of Dis3$^{G766R/+}$ genotype and 30 WT mice were aged under standard conditions for over two years. However, no effect of the MM-associated Dis3$^{G766R/+}$ allele was observed (Figure 2A). Since DIS3 mutations are specific to MM and not commonly found in other types of cancer, we decided to specifically investigate the development and progression of

malignancies associated with a B cell lineage. We employed a pristane-induced murine plasmacytoma model (Takizawa *et al*, 2008; Ramiro *et al*, 2004). The C57BL/6 mouse line is naturally resistant to plasmacytoma. Both Dis3$^{G766R/+}$ and WT mice exhibited similar responses to the pristane injection, indicated by the incidence of diffuse alveolar hemorrhage (Figure 2B). However, Dis3$^{G766R/+}$ mice exhibited a drastic acceleration and a significantly higher incidence of plasmacytoma compared with WT controls (Figure 2C). Plasmocytoma cases in Dis3$^{G766R/+}$ displayed typical characteristics with atypical plasma cells in an ascites fluid (Figure EV3A-B). They were also accumulating more ascites fluid (Figure EV3C-F).

To test if the higher incidence of plasmacytoma can at least partially be explained by transcriptome changes, we have performed RNA-seq experiments of *in vitro* activated B cells isolated from WT and DIS3$^{G766R/+}$ mice. However, the effect of the mutation on annotated transcriptome was negligible (Figure 2D-F, Dataset EV2-EV4). Also, the most prominent DIS3 substrates, such as rRNA processing by product and PROMPTs, showed very moderate or no molecular phenotypes (Figure EV4A,B). These findings are in concordance with results obtained from the analysis of MM patients described in the accompanying manuscript, which indicate no specifically deregulated RNA that could influence cancerogenesis. This suggests that the oncogenic effect of Dis3 mutations is exerted through the mutant protein impact on DNA rather than RNA.

DIS3-dependent translocations in mice have an AID footprint

We established that the Dis3^{G766R} mutation in mice leads to a greater predisposition to plasmacytoma, known to depend on aberrant AID-dependent translocations during CSR (Takizawa *et al*, 2008; Ramiro *et al*, 2004). Considering that transcription of the IGH switch locus is essential for AID-dependent recombination (Pefanis *et al*, 2015, 2014; Basu *et al*, 2011), we aimed to characterize translocations in Dis3$^{G766R/+}$ plasmacytomas.

To obtain a sufficient amount of material and better compare plasmacytomas from Dis3$^{G766R/+}$ and WT mice, we have backcrossed Dis3$^{G766R/+}$ mouse line into plasmacytoma prevalent BALB/C background (more than 6 crosses) and used WT and Dis3$^{G766R/+}$ littermates to induce plasmacytoma. Then we isolated DNA from the cells pelleted form the ascit and used deep DNA sequencing to look globally at chromosomal translocations. The analysis of mate-pair and whole-genome DNA sequencing data revealed that plasmacytomas from Dis3$^{G766R/+}$ mice had significantly higher rates of chromosomal translocations involving IGH locus (Figure 3A, H and Dataset EV5-EV6). In addition, more sequencing chimeric reads originating from two different genomic locations were identified in Dis3$^{G766R/+}$ (Figure 3B), indicating increased genome structural instability. These results

imply that the Dis3$^{G766R/+}$ mutator effect impacts genomic arrangement. A meta-analysis of genomic regions surrounding chromosomal breaks revealed striking enrichment of AID occupancy detected by chromatin immunoprecipitation (ChIP-seq) (Figure 3C). Further analysis revealed that both translocation sites and regions of AID activity are GC-rich (Figure 3D), are enriched with AID-specific motifs (Figure 3E), and are in close proximity to regions predicted to form G-quadruplexes (Figure 3F). When transcribed (Figure 3G), such regions are prone to RNA polymerase stalling and R-loop formation, which are, in turn, susceptible to AID-dependent translocations (Casellas *et al*, 2016; Yu *et al*, 2003). DIS3 mutants, including the MM DIS3^{G766R} variant, have been shown to less efficiently degrade structured RNA (Tomecki *et al*, 2014; Szczepińska *et al*, 2015), suggesting significant difficulties in processing G-quadruplexes. A snapshot of the immunoglobulin genomic region (the best studied in the context of AID activity) exemplifies the correlations mentioned above (Figure 3H).

We conclude that MM-associated DIS3 G766R mutation increases the rate of chromosomal translocations, which display an AID footprint.

Mutational profiles revealed signatures of promiscuous AID activity and a prevalence of translocations in patients with MM DIS3 alleles

Having established that specifically in B cell lineage MM DIS3 alleles lead to increased genome instability, we aimed to correlate this observation with the CoMMpass MM clinical data comparing patients with and without DIS3 mutations (Kuliński *et al*, 2023 submitted)

First, analysis of translocations involving IgH locus revealed their significantly higher occurrence in patients with DIS3 mutations (64.3% of non-recurrent DIS3 mutated patients, 64.5 % of recurrent DIS3 mutated patients, and 33.2 in the rest of the patients) (Figure 4A, B), agreeing with the specificity of DIS3 mutations for non-HD MM and the role of DIS3 in AID-dependent translocations.

Then to characterize the effect of MM DIS3 alleles on point mutations, we analyzed 96 COSMIC (Catalogue of Somatic Mutations in Cancer) mutational classes of single-base substitution (SBS) profiles (Rustad *et al*, 2021), estimating the contribution of each of them (Figure 4C-G, Figure EV5C, D). Exome-wide, age-associated SBS5 was identified as the primary mutational signature in patients of all DIS3 genotypes (Figure 4C, Figure EV5C), in agreement with the late onset of MM and previous analysis (Hoang *et al*, 2020; Maura *et al*, 2020; Alexandrov *et al*, 2015). Similarly, MM-specific mutations detected in DIS3 in CoMMpass study patients and mutations reported in the COSMIC database show a predominance of age-associated SBS5 (Figure 4D). Despite the exosome complex

involvement in SHM, only a small fraction of mutations across the exome could be attributed to mutation signatures associated with AID activity across all the DIS3 genotypes (Figure 4C), specifically SBS84 (reflecting canonical AID according to the most recent classification) and SBS85 (concordant with indirect effects of AID activity) (Kasar *et al*, 2015).

Interestingly, transcribed PROMPT regions display distinct somatic mutation patterns than the entire exome, having a significantly higher proportion of substitutions attributed to DNA damage, especially SBS3, associated with the erroneous repair of double-strand DNA breaks in a BRCA-dependent mechanism (Figure 4E). This, in principle, could be explained by the reported DIS3-dependent increase in genome instability accompanying RNA:DNA hybrid accumulation (Gritti *et al*, 2022). However, despite PROMPTs being the most prominent DIS3 substrates, the somatic MM DIS3 alleles do not affect the proportion (Figure 4E). Notably, the contribution of AID-dependent signatures in all transcribed PROMPT regions is very limited (Figure 4E).

Looking at known AID targets, which also include a subset of PROMPTs, the contribution of AID-dependent signatures is significant and higher in patients with MM DIS3 alleles, especially the recurrent ones (Figure 4F). Further analysis on even more narrow genomic regions of known AID-dependent MM driver genes revealed a very strong overrepresentation of signatures related to both AID-dependent mutagenesis and defects in the repair of DNA lesions typical for CSR and SHM (namely: BER SBS10b, SBS30, and MMR SBS6, SBS26) (Figure 4F-G). Notably, patients with MM-specific DIS3 variants exhibit a drastic increase in the contribution of these signatures, especially SBS26 (35% in recurrent DIS3 mutated patients, 23% in non-recurrent DIS3 mutated patients, and 15% in the rest of the patients; Figure 4E) supporting the AID-DIS3 cooperation.

All results presented above indicate that MM DIS3 alleles enhance AID-related mutagenesis occurring within a very limited time window of naive and memory B cell activation rounds at very specific loci, which are positively selected in MM.

DIS3 stalling on chromatin-bound substrates

Given the correlation between MM DIS3 alleles with AID-dependent mutations and translocations in both patients' samples and the murine model, we next wanted to gain insights into the mechanism, which we hypothesize to be related to DIS3 chromatin association.

Our attempts to immunoprecipitate DIS3 with chromatin were unsuccessful, possibly due to the indirect and transient nature of the DIS3 interaction with chromatin. Thus, to experimentally verify the direct involvement of MM DIS3 protein variants, we analyzed chromatin occupancy of the G766R mutant and WT DIS3 using the DamID methodology.

This approach is based on fusing a protein of interest to DNA adenine methyltransferase (Dam), which modifies adjacent adenines in DNA, providing sufficient sensitivity to detect even weak interactions between DNA and its associated factors (Vogel *et al*, 2007). We modified the established mouse B cell lymphoma CH12F3-2A cell line capable of cytokine-induced IGH CSR. Even before analyzing the sequencing data, we noticed 2.5-fold higher library yields from cells that express DIS3 G766R compared with WT DIS3 (Figure 5A). Such an increase can be explained by a higher chromatin occupancy of the mutated protein and agrees with the previously observed stalling of DIS3 mutant on RNA substrates (Szczepińska *et al*, 2015). DIS3 has a very broad pattern of association with chromatin. However, PROMPT regions, which give rise to the most potent DIS3 substrates (Figure 5B), and sites of translocations mapped in plasmacytomas show significant enrichment for DIS3 of both genotypes, coinciding with native transcription (Figure 5B,C). Notably, known AID off-target genes are among the most highly DIS3-occupied loci in the genome (Figure 5B). Moreover, DIS3 occupies the vicinity of switch regions in the IGH locus (Figure EV6A), agreeing with the role of DIS3 in CSR.

The data presented above indicate that DIS3 acts together with AID to drive translocations. To support cooperation with an orthogonal methodology, we have analyzed interactions between DIS3 and AID using Proximity ligation assay (PLA) (Figure 5D). Indeed in CH12F3-2A cells, we see PLA positive spots, which number is higher upon *in vitro* activation (Figure 5E and Figure EV6B). We also examined DNA damage accompanying B cell activation in vitro using γ-H2AX, an established marker dsDNA breaks (Mah *et al*, 2010). Naïve B cells present a low level of DNA damage, which, however, increased significantly in cells 3 days post-activation (Figure 5F, G). In contrast to naïve, activated B-cells demonstrated a moderate, yet statistically significant, increase of DNA damage in Dis3$^{G766R/+}$ compared to Dis3$^{+/+}$ (Figure 5F, G), further supporting the role of MM DIS3 alleles in B cell specific cancerogenesis.

MM DIS3 variants do not affect B cell activation and Class Switch Recombination

The nuclear exosome has been shown to target AID to switch regions during CSR, and a complete knockout of the nuclear exosome complex (Basu *et al*, 2011) and DIS3 itself (Laffleur *et al*, 2021) abrogate class switching. Therefore, we wanted to analyze the effect of DIS3^{G766R} mutation on CSR using *in vitro*-activated primary mouse B cells. As revealed by cytometric analysis, on average 20% of cells have underwent the switch to the IgG1 heavy chain class at day 3 post activation, in both genotypes, proving that heterozygous DIS3^{G766R} mutation does not affect CSR (Figure 6A, Figure EV6C). RNA-seq experiment WT and

heterozygous DIS3^{G766R} B cells also revealed normal transcriptional responses to *in vitro* activation (Figure 6B,C Dataset EV7-S10). Accordingly, analysis of CoMMpas MM datasets revealed that class switch recombination is unaffected in patients with DIS3 mutations (Figure 6D). Altogether our results prove that MM DIS3 alleles in a heterozygous state, prevalent in patients, lead to gain-of-function phenotypes inducing oncogenic translocations and mutations without abrogating exosome function essential for CSR.

DISCUSSION

A common property of cancer cells is alteration of gene expression. This was initially attributed mainly to changes in transcriptional regulation but more recently also to posttranscriptional regulatory mechanisms. In hematologic malignancies, mutations of spliceosome components, such as SF3b (Yoshida *et al*, 2011; Wang *et al*, 2011; Papaemmanuil *et al*, 2011; Cieśla *et al*, 2023) and U2AF (Graubert *et al*, 2011; Yoshida *et al*, 2011), are leading to changes in alternative splicing patterns. Still, mechanistic links are missing, despite many clinical and functional analyses employing several transgenic mouse models that show effects on RNA metabolism (Taylor & Lee, 2019). In MM, two genes involved in RNA metabolism DIS3 (RRP44) (Dziembowski *et al*, 2007) and TENT5C (FAM46C) (Mroczek *et al*, 2017) are specifically mutated (Walker *et al*, 2018; Lohr *et al*, 2014; Chapman *et al*, 2011), jointly being affected in almost 30% of MM patients. TENT5C is a MM specific oncosupressor (Lohr *et al*, 2014; Mroczek *et al*, 2017; Chapman *et al*, 2011). In normal B cells, TENT5C enhances the expression of immunoglobulins by stabilizing their mRNA through cytoplasmic polyadenylation, which at the same time increases the protein load into the endoplasmic reticulum (Manfrini *et al*, 2020). TENT5C disruption decreases ER stress and enhances proliferation in antibody-producing malignant plasma cells.

The present study provides a mechanism by which DIS3 drives oncogenic transformation in MM. Contrary to previous suggestions, dysregulation of the transcriptome is not the cause of the neoplastic transformation. MM DIS3 alleles are also not causing global genome destabilization. Instead, MM DIS3 mutations lead to a gain-of-function which introduces B cell-specific destabilization of the genome, enabling the acquisition of genetic lesions and giving rise to oncogenic genome rearrangements, such as IgH translocations, known to be the primary drivers of MM (Figure 7).

DIS3 is a main nuclear exoribonuclease that acts from the RNA 3’ end. It is targeted to many types of RNA species. At the chromatin, these are mainly pervasive transcription products that are rapidly cleared. In B cells, DIS3 additionally targets sites of SHM and CSR, facilitating AID recruitment. MM DIS3 variants stalls at their substrates during SHM and CSR, generating a gain-of-function phenotype of AID promiscuous activity leading to higher

levels of translocations. In a mouse model, a relatively mild heterozygous DIS3^{G766R} mutation drastically increases the rate of plasmacytoma development, and no detectable changes in the annotated transcriptome do accompany such a mutagenic effect. In MM samples, DIS3 alleles display detectable transcriptional signatures, which are, however, reversible as recurrent DIS3 alleles leading to an increased rate of translocations, are often couterselected in the late stages of MM (Kuliński *et al*, 2023 submitted). Interestingly, despite the overall increase in driver translocations, there is a striking specificity for specific translocations depending on the DIS3 genotype (Figure 4A, B). The t(4;14) translocation is depleted in patients with recurrent DIS3 somatic mutations. At the same time, in patients with the coexistence of both lesions, the recurrent MM DIS3 variants always represent minor clones only. Notably, the t(4,14) translocation activating pro-proliferative FGFR3 expression is accompanied by the del(13q) deletion in 90% of patients (Keats *et al*, 2006; Fonseca *et al*, 2001). Thus, its coexistence with the defective MM-specific DIS3 allele may lead to the faster and more effective counter-selection of recurrent DIS3 mutation in patients with the t(4,14) translocation.

In physiological situations, there are activation-dependent interactions between DIS3 and AID in the B-cell lineage confirmed by us and restricted to limited spots, comparable to the numbers of DNA damage foci measured by γ-H2AX B-cell activation *in vitro*. The statistically significant but moderate increase of DNA damage in Dis3$^{G766R/+}$ cells is in contrast to DIS3 depletions, which dramatically affect both cell physiology and genome integrity (Gritti *et al*, 2022). Moreover, MM DIS3 mutations display no CSR impairments both in mouse models and in MM samples contrasting DIS3 inactivation, which leads to the rearrangement of tertiary chromatin architecture and inhibits CSR in murine B cells (Laffleur *et al*, 2021). A very restricted spectrum of DIS3-dependent mutagenesis in MM is supported by the analysis of mutational signatures in MM samples, clearly showing the association of DIS3 mutations with AID-related lesions in MM driver genes and not genome-wide, which can occur only in a short window of opportunity CSR and SHM during B-cell activation. All this explains the specificity of DIS3 mutations for MM or, in the case of the mouse model, plasmacytoma.

In summary, our study provides a mechanistic explanation for the role of a DIS3 mutation in MM. Together with the accompanying paper, we describe a very unusual oncogene that drives MM but is counter-selected later because of its toxicity.

METHODS

CH12F3 cell line

The mouse CH12F3 B-cell lymphoma cell line was acquired from the Ricken BRC cell bank. Cells were cultured in RPMI 1640 ATCC's modified medium (Invitrogen; A1049101) supplemented with 10% FBS (Invitrogen; 10270-106), 5% NCTC-109 medium (Gibco A/Thermo; 21340039), 50 µM 2-mercaptoethanol, penicillin/streptomycin (Sigma; P4083) at density between 0.5-1 x 10^6 cells/ml. Media were changed every 3 days.

To obtain lentiviral constructs for the DamID experiment, mouse WT and S766R *DIS3* coding sequences were fused to the Dam methylase coding sequence. WT *DIS3* cDNA was amplified and cloned into pLgw V5-EcoDam lentiviral constructs (Addgene; 59210) using SLIC. The S766R DIS3 construct was obtained by site-directed mutagenesis.

For the negative control construct, mEGFP was amplified using PCR and cloned into pLgw V5-EcoDam lentiviral constructs (Addgene; 59210) using SLIC. Lentiviral transductions of CH12F3 cells were carried out exactly as described previously. Transduced cells were selected using Zeocin at a concentration of 250 µg/ml for 14 days and subsequently, cell cultures were expanded and preserved in 50% FBS, 40% RPMI medium and 10% DMSO.

CH12F3 cells were activated for 3 days by supplementing the standard culture medium with 1 ng/ml TGF-β1, IL-4 5 ng/ml, 1 µg/ml αCD40.

(*)Primers

WT *DIS3* cDNA F	aagtatatatttttaataagggtgggcgcgatgctcaggtccaagacgttc
WT *DIS3* cDNA R	cactttgtacaagaaagctgggtcggcgcgcttctcaagcttcctcttc
site-directed mutagenesis F	catcactatCgcttagcctccccc
site-directed mutagenesis R	gggggaggctaagcGatagtgatg
mEGFP F	tccgaaaacctgtacttccaaggaaccggtatggtgagcaagggcgagg
mEGFP R	atcaccctgaaaatacaaattctcgctagccttgtacagctcgtccatgc

DamID-seq protocol

The DIS3 DamID-seq procedure and library preparations were performed as described previously(Kind *et al*, 2015). In short: DNA from CH12F3 cells expressing DIS3 DamID constructs or not, as the control, were isolated using a Bioline Isolate II genomic DNA kit (BIO-52067). First, 0.5 µg of DNA was digested with DpnI (NEB; R0176) in 10 µl volume at 37°C for 16h. Next, DpnI was heat-inactivated, and the DpnI adaptor was ligated in 20 µl volume with 2.5 units of T4 DNA ligase (NEB; M0202). This was followed by DpnII digestion in a total volume of 50 µl with 10U of DpnII (NEB; R0543). DNA was amplified by

PCR with Taq polymerase using the following program: samples were denatured at 95 for 8 minutes, then aplyfied for 19 cycles, with: 20'' of denaturation at 94℃, 30'' of anneling at 58℃ and 20'' elongation 72℃ followed by a 2 min incubation at 72℃.

PCR products were cleaned up using a Bioline Isolate II PCR and Gel kit according to the manufacturer's instructions (BIO-52060). DNA was repaired with Blunt-end ending 3′ overhang using End-It DNA End-Repair kit in 25 µl volume (Epicentre; ER81050), and 3′-A overhang was added with Klenow fragment (NEB; M0212) and dATP. This was followed by purification using CleanPCR beads (CleanNA; CPCR-0005). DNA concentration was measured by Nanodrop and adjusted to 40 ng/µl. Illumina Y-adaptor was ligated, and the libraries were PCR amplified using standard protocols and sequenced with an average read number of 40M.

(*)Primers

DpnI PCR	nnnngtggtcgcggccgaggatc
Adaptor_top	ctaatacgactcactatagggcagcgtggtcgcggccgagga
DpnI adaptor bottom	Tcctcggccgcg
Y-adaptor top	Acactctttccctacacgacgctcttccgatct
Y-adaptor bottom	P-gatcggaagagcacacgtct (5′- phosphorylated)
P5- Illumina-2 PCR	aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatct

PLA procedure

In situ Proximity Ligation Assay was performed on activated CH12F3-2A cells and non- activated (naive) cells as a control. Activation was performed by seeding cells at 105 cells/ml in culture medium supplemented with 1 µg/ml anti-CD40 antibody (Invitrogen, MA1-81395), 5 ng/ml IL-4 (Peprotech, 214-14) and 1 ng/ml TGF-β1 (Peprotech, 100-21), and cells were grown for 72 h at 37°C in 5% CO2. For each experimental condition, 2×106 cells were spun (400 rcf, 4 min), washed with PBS and spun again. Next, cells were fixed with 3,7% formaldehyde in PBS for 25 min at room temperature, washed with PBS twice, permeabilized with 1% Tween-20 (v/v) in PBS for 15 min at room temperature and washed with PBS once. Then, cells were blocked with 3% BSA in PBS for 1 h at room temperature. In order to reduce Fc receptor-mediated binding by antibodies used in PLA procedure, Mouse BD Fc Block (BD Biosciences, 553142) was added to cells in amount of 1 µg/10^6 cells and incubated according to manufacturer's instructions. Primary anti-AID (1:100; Invitrogen, PA5-18913) and anti-DIS3 antibodies (1:100; Proteintech, 14689-1-AP) were added to cells in blocking solution and incubated overnight at 4°C. Additionally, three background controls were performed in which cells were incubated only with anti-AID antibody or only with anti-DIS3 antibody, or without primary antibodies. Next, cells were washed with PBS twice and the in-situ PLA was performed with the Duolink In Situ PLA kits (Sigma-Aldrich) according to manufacturer instructions with the exception that the reactions

and washings were carried out in 1,5 ml microcentrifuge tubes, and all the incubations were performed in thermoblock. The interaction between AID and DIS3 was detected using Duolink In Situ PLA Probe Anti-Goat MINUS kit (Sigma-Aldrich, DUO92006) and Duolink In Situ PLA Probe Anti-Rabbit PLUS kit (Sigma-Aldrich, DUO92002) against anti-AID antibody and anti-DIS3 antibody, respectively. Fluorescent signal was generated using Duolink In Situ Detection Reagents Red kit (Sigma-Aldrich, DUO92008). Cells were mounted and the nuclei were stained using the Duolink In Situ Mounting Medium with DAPI (Sigma-Aldrich, DUO82040) as follows: after the last wash, cells were spun, all the supernatant was removed carefully, cells were resuspended thoroughly but gently in 12 µl of Mounting Medium, the cell suspension was dripped onto a microscope slide, covered with a coverslip (Ø 18 mm), and gently pressed down. Edges of the coverslip were sealed with nail polish.

γH2A.X immunofluorescent staining

DNA double-strand breaks were stained in mouse B cells using γH2A.X as a marker. For each experimental condition 106 cells were spun (400 rcf, 4 min), washed with PBS and spun again. All the following steps were carried out in suspension in 1,5 ml microcentrifuge tubes. Washed cells were fixed with 3,7% formaldehyde in PBS for 25 min at room temperature, washed with PBS twice, permeabilized with 1% Tween-20 (v/v) in PBS for 15 min at room temperature and washed with PBS once. Then, cells were blocked with 3% BSA in PBS for 1 h at room temperature. Primary anti-γH2A.X (1:400; Abcam, ab2893) was added to cells in blocking solution and incubated overnight at 4°C. After being washed twice with PBS, cells were incubated with goat anti-rabbit IgG secondary antibody conjugated to Alexa Fluor 555 (1:800; Invitrogen, A-21429) for 1 h at room temperature. Next, cells were washed with PBS once and the nuclei were stained using 2 µg/ml Hoechst 33342 (Invitrogen, H3570) in PBS for 5 min at room temperature. After being washed twice with PBS, cells were spun, all the supernatant was removed carefully, cells were resuspended thoroughly but gently in 12 µL of ProLong Glass Antifade Mountant (P36980, Invitrogen), the cell suspension was dripped onto a microscope slide, covered with a coverslip (Ø 18 mm), and gently pressed down.

Microscopic analysis

For Fig. 5, cells were imaged using Zeiss LSM800 confocal microscope with 40×/1.3 oil immersion apochromatic objective and GaAsP PMTs detectors. ZEN software (Zeiss, version 2.6) was used for microscope control. Imaging was performed at a controlled temperature of 22□C. Z-stack images were processed using Fiji/ImageJ software (version 1.53f51). For each Z-stack, the maximum intensity Z-projection was applied and then

images of PLA channel were subjected to rolling-ball background subtraction with a radius set to 20 pixels and contrast enhancement with 0.025% saturated pixels.

For Fig.5 and Figure EV6, cells were imaged using Olympus IX81-ZDC wide-field fluorescence microscope with 40×/0.95 apochromatic objective and EM-CCD camera (Hamamatsu, C9100-02) equipped with ScanR modular imaging platform (Olympus). ScanR Acquisition software (version 2.2.0.8) was used for microscope control. Imaging was performed at a controlled temperature of 22□C. Image analysis was performed using ScanR Analysis software (version 3.0.0). Images were subjected to rolling-ball background subtraction with a radius set to 40 and 2 pixels for DAPI and PLA channel, respectively. Next, segmentation was performed on the DAPI channel using the Intensity algorithm. The identified objects were gated based on a scatter plot of the circularity factor relative to the area. The object that comprised the gate was classified as normal, single nucleus. PLA spots in a cell were searched for as sub-objects in the area encompassing a nucleus and a 7-pixel radius outside it, basing on PLA channel and using the Intensity algorithm. Each identified sub-object was assumed to be a PLA spot. Further numerical data processing and statistical analysis were performed using R environment. Cells containing at least one PLA spot were filtered out and for each experimental condition, a total number of spots in a cell and its mean were calculated. Statistical significance was calculated using the Wilcoxon test. All data were checked beforehand for normality using the Shapiro-Wilk test.

For Fig. 5F,G cells were imaged using Olympus IX81 wide-field fluorescence microscope with 60×/1.35 oil immersion apochromatic objective and CCD camera (Hamamatsu, Orca-R2/C10600) equipped with ScanR modular imaging platform (Olympus). ScanR Acquisition software was used for microscope control. Imaging was performed at a controlled temperature of 22□C. Image analysis was performed using ScanR Analysis software (version 3.0.0). Images were subjected to rolling-ball background subtraction with a radius set to 75 and 2 pixels for Hoechst and γH2A.X channel, respectively. Next, segmentation was performed on the Hoechst channel using the Intensity algorithm. The identified objects were gated based on a scatter plot of the circularity factor relative to the area. The object that comprised the gate was classified as normal, single nucleus. γH2A.X foci in a cell were searched for as sub-objects in the area encompassing a nucleus basing on γH2A.X channel and using the Intensity algorithm. Each identified sub-object was assumed to be a DNA double-strand break. Further numerical data processing and statistical analysis were performed using R environment. For each experimental condition a total number of DSB in a cell was calculated. Statistical significance was calculated using the Wilcoxon test.

For Fig. 5F representative images obtained by the abovementioned imaging procedure were selected. Images were processed using Fiji/ImageJ software (version 1.53f51). Images of both channels were cropped to provide a more accurate visualization of γH2A.X foci in the cells. Images of γH2A.X channel were subjected to rolling-ball background subtraction with a radius set to 3 pixels and contrast enhancement with 0.01% saturated pixels. Images of Hoechst channel were subjected to rolling-ball background subtraction with a radius set to 40 pixels

Mouse line generation

The G766R mice were constructed according to the strategy described in Figure EV4A in the Institut clinique de la Souris (ICS) Illkirch, France using homologues recombination technology in mES cells in the C57BL/6N background. Mice are genotyped by PCR using primers:

766seq_F	gtctcaaaaacacacggcagagactc
766seq_R	gaggatgtgtgagataagatacaccgg

Dis3 WT allele migrates as a 202 bp band, whereas the mutant G766R allele gives a 281 bp amplicon.

Mice breeding conditions

Mice were bred in the animal house of Faculty of Biology, University of Warsaw maintained under conventional conditions in a room with controlled temperature (22 ± 2°C) and humidity (55 ± 10%) under a 12 h light/12 h dark cycle in open polypropylene cages filled with wood chip bedding (Rettenmaier). The environment was enriched with nest material and paper tubes. Mice were fed ad libitum with a standard laboratory diet (Labofeed B, Morawski). Animals were closely followed-up by the animal caretakers and researchers, with regular inspection by a veterinarian, according to the standard health and animal welfare procedures of the local animal facility.

Mouse embryonic fibroblast primary cell cultures

Primary cell cultures of mouse embryonic fibroblasts (MEF) were isolated from E13.5 embryos as defined by the date of the vaginal plug and confirmed by characteristic features of the embryo morphology. Embryo heads were used for DNA isolation and genotyping. The corpses were gutted and mechanically disintegrated by chopping and passing the tissue fragments suspended in culture media (DMEM supplemented with 10% FBS, 50 μM 2-mercaptoethanol, penicillin/streptomycin) with a syringe through the gauge 19 needles. Disintegrated tissue fragments of a single individual were left to settle in 100 mm petri dishes and form outgrowths. After 3 days, cells were treated with 0.25% Trypsin in 1 mM EDTA for

3 min. Cells were passaged for at least 3 passages before they were collected or used for further applications. Cells were passaged every 3 days. Cells isolated from 3 independent litters in pairs, including $DIS3^{+/+}$ and $DIS3^{G766R/+}$ were used for the RNA-seq library preparation.

For the 3T3 protocol, MEF cells from 12 $DIS3^{+/+}$ and 13 $DIS3^{G766R/+}$ littermates were used. Cells passaged 2 times to disintegrate into a single cell suspension and noted. Cells were plated into a 10-cm cell culture dish at a density of 9×10^5 cells per dish. Cells were cultured at standard conditions with medium change after 2 days in culture. Three days after plating, cells were trypsinized, counted using a hemocytometer, and replated 9×10^5 into a 10-cm cell culture dish. The strict passaging regime was repeated for 30 passages or until the immortalization of cells was obtained. Cumulative cell numbers were plotted.

Pristane-induced mouse plasmacytoma (MPC) model

Six- to ten-week old mice were injected 3 times intraperitoneally, in 2-month intervals, with 0.5 ml of pristane oil (2,6,10,12-tetramethylpentadecane) to induce the development of mouse plasmacytoma (MPC), an established murine model of multiple myeloma (Gadó *et al*, 2001; Potter, 1997). The day of the first injection is considered as day 0 of the experiment. Mice were regularly observed for abdominal swelling. The neoplastic nature of the tumors was confirmed by paracentesis and ascitic smear stained with Wright-Giemsa stain. Animals were closely followed-up by the animal caretakers and researchers, with with particular attention to lupus-like syndrome signes such as alveolar hemorrhage or fur loss. Regular inspection by a veterinarian was conducted, according to the standard health and animal welfare procedures of the local animal facility.

Mice for the pristane-induced mouse plasmacytoma experiment were maintained in individually ventilated cages in a room with controlled temperature (22 ± 2°C) and humidity under a 12 h light/12 h dark cycle. To obtain G766R mice on a BALB/C genetic background, $DIS3^{G766R/+}$ mice were backcrossed to BALB/CanNCrlCmd mice for 6 generations.

Tissue collection and primary cell culture, *in vitro* B-cell activation, CSR efficiency assay

Splenic naïve B-cells were prepared using the CD43 negative selection commercial kit (Miltenyi Biotec.; 130-090-862). Spleens were pulled from 2-3 sacrificed sex- and age-matched animals mechanically disintegrated and filtered through a 70 µm cell strainer to prepare a single cell suspension essential for microbead separation. Spleen extracts were depleted from red blood cells using ACK lysis buffer. The magnetic bead, column-based

enrichment of naïve B-cells was performed according to the manufacturer's instructions. Cells were resuspended and cultured in RPMI 1640 containing 15% FBS supplemented with 100 nM 2-mercaptoethanol and penicillin/streptomycin (Sigma; P4083).

Ex vivo activation of B lymphocytes for class switch recombination, proliferation studies and RNA sequencing and was achieved by culturing them in RPMI 1640 containing 15% FBS supplemented with 100 nM 2-mercaptoethanol, 20 µg/ml LPS and 20 ng/ml IL-4. CSR efficiency was assessed in B-cells 3 days after activation. In brief, cells were collected and stained for 30 min with Alexa Fluor 555 conjugated goat antibody raised against mouse IgG1 (Invitrogen; A-21127) at 10 µg/ml. The percentage of positive cells that underwent IgG1 class switch recombination was quantified using flow cytometry on an Attune NxT Cytometer (Thermo) or CytoFLEX Flow Cytometer (Beckman Coulter) and analyzed using FlowJo (v10.7.1) software.

CellTrace Proliferation Assay

CellTrace™ Yellow Proliferation kit was used according to the manufacture's instructions. Briefly: Cells were stained by adding 1 µl of CellTrace™ Yellow (Invitrogen, Cat #C34567) stock solution in DMSO per each ml of 10^6/ml cell suspension in PBS (5 µM working concentration), and subsequently incubated for 40 minutes at room temperature, while being protected from light. Afterward, five times the original staining volume of culture medium, containing FBS, was added to the cells and incubated for 5 minutes. The cells were then pelleted and resuspended in a fresh pre-warmed complete culture medium. 10 minute incubation period was allowed before analysis to enable the acetate hydrolysis of the CellTrace™ reagent. Cells were pipetted in technical triplicates, $3x10^5$ per well, into round bottom 96 well plate in 200 µl growth medium. Cells were activated by supplementing media with 20 µg/ml LPS and 20 ng/ml IL-4 and cultured for 3 days. Naïve B-cells and unstained control cells were parallelly cultured in triplicates. Before cytometric analysis, dead cells were stained with LIVE/DEAD fixable dyes (Invitrogen™) according to manufactures instructions and analized on a CytoFLEX Flow Cytometer (Beckman Coulter). Data were analyzed using FlowJo (v10.7.1) software using the Proliferation functionality.

RNA-seq, Mate-Pair and Paired-End DNA-seq

Total RNA from collected tissues and primary cultures was isolated using a standard TRIzol protocol followed by ribodepletion with a Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat) according to the manufacturer's instructions (Epicentre; RZG1224). Strand-specific libraries were prepared using a TruSeq RNA Sample Preparation Kit (Illumina) and the dUTP method. Three independent replicate sample sets were prepared for

each condition. DNA was purified using a SherlockAX kit optimized for small samples (A&A Biotechnology; 095-25). NEXTERA XT mate-pair and TruSeq DNA libraries were prepared according to manufacturer's instructions (Illumina; FC-131-1024 and FC-121-2001). All of library types were sequenced using Illumina NextSeq or NovaSeq platform in a paired-end mode with different average sequencing depth: MEF RNA-seq: 29M reads, primary B-cell cultures RNA-seq: 23M reads, Mate-pair DNA libraries: 42.7M reads, standard DNA-seq libraries: 281.5M reads (NovaSeq).

Bioinformatic analysis

MMRF CoMMpass study whole-genome DNA sequencing reads were mapped to the GRCh38 human reference genome, using STAR (Dobin *et al*, 2013) processed and filtered using SAMtools (Li & Barrett, 2011). To identify translocations, significant structural variants were called using SVDetect (Zeitouni *et al*, 2010) with default settings. Significant translocations were filtered for those known to be commonly initiated in the IGH locus and genotype. RNA sequencing reads were mapped using STAR split read aligner. RNA sequencing samples of patients genotyped based on exome sequencing as being a DIS3 somatic mutation or a recurrent DIS3 mutation were compared to RNA samples from patients not possessing such a genotype as well as not possessing a germline DIS3 mutation or a mutant DIS3 variant as a minor clone. Reads were counted onto a gene annotation (gencode v34 basic) using FeatureCount (subread package release 2.0.0) (Liao *et al*, 2014). Differential expression RNA-seq analyses were performed as previously using DESeq2 R package (Lovén *et al*, 2012). Whole-exome DNA sequencing reads were mapped to the GRCh38 human reference genome, using STAR split read aligner (Dobin *et al*, 2013), and quantified using SAMtools and BCFtools pipeline (Li & Barrett, 2011). Healthy tissue control whole-exome DNA sequencing data were mapped and variants that were identified in at least one control sample were defined as germline mutations. The Single Base Substitution (SBS) profiles were calculated using SigProfilerMatrixGenerator (Bergstrom *et al*, 2019) and fitted to the COSMIC Mutational Signatures (v3.2) using mmsig (Rustad *et al*, 2021).

To identify translocations in plasmacytoma DNA samples, mate-pair DNA sequencing data were preprocessed with NXtrim(O'Connell *et al*, 2015). Both mate-pair and standard TruSeq DNA libraries were mapped by STAR split read aligner. Statistically significant structural variants were called using SVDetect with default settings. All significant structural variants from all sequencing experiments have been concatenated together according to the genotype and merged using bedtools merge to exclude loci involved in more than one structural variant to be analyzed multiple times. These annotations have

been used for the meta-analysis. For the filtering to identify those involving an extended region, including the immunoglobulin heavy chain locus, custom scripts were used.

DamID-seq data were stripped of sequencing adapters using cutadapt and mapped to the mouse reference genome (GRCm38) using STAR aligner. Uniquely mapping reads were counted onto DpnI restriction fragments in the mm10 genome. Count tables were library size corrected. Counts for the 3 replicates of the DamOnly negative control have been averaged and subtracted from all replicates to eliminate the background signal. The annotation of putative PROMPT 3kb regions upstream and antisense to the known transcription start sites in the gencode release M25. Mean values have been calculated of DIS3 occupancy.

ChIP-seq reads were aligned to the mouse reference genome (GRCm38.p6) using bowtie2. All replicates were pooled and used to call peaks with MACS2 (v2.2.7.1) using IGG control as the background. The obtained tracks and peak annotations were utilized in downstream analysis.

Meta-analyses of genomic regions were performed using custom scripts employing BWtool. Structural variants hotspots were tested for nucleotide composition similarities and their relation to AID chromatin occupancy or transcription using custom scripts. The AID DNA recognition motifs tracks were prepared by calculating the density of RGYW, GAGCT, GGGGW, and GGGCT sequence occurrence in the GRCm38.p6 mouse reference genome. Two G-quadruplex prediction annotations were performed by scanning the GRCm38.p6 mouse genome for at least 4 stretches of at least three G separated by up to 7 of any nucleotides for a more stringent annotation or separated by 30 of any nucleotides for the less stringent annotation. From these annotations, strand-specific density tracks were prepared and used for meta-analysis.

The likelihood of G766R mutation pathogenicity, was assessed using the PON-P2 algorithm (Niroula *et al*, 2015), with default settings. This machine learning-based classifier categorizes variants into pathogenic, neutral, and unknown classes by utilizing various information, such as evolutionary conservation of sequences, physical and biochemical properties of amino acids, GO annotations, and, if accessible, functional annotations of variation sites.

Publicly available data from the following publications were used: ChIP-seq AID SRA: SRP003605 (Pavri *et al*, 2010); EXOSC3 KO RNA-seq SRA: SRP042355 (Pefanis *et al*, 2014); GRO-seq B-cells and CH12F3 SRA: SRP193758 (Zhang *et al*, 2019).

Ethical issues

All procedures were approved by the First Local Ethical Committee in Warsaw affiliated at the University of Warsaw, Faculty of Biology (approval numbers WAW/092/2016, WAW/177/2016, WAW/642/2018). Housing in animal facilities was performed in conformity with local and European Commission regulations under the control of veterinarians and with the assistance of trained technical personnel.

Acknowledgments:

We thank Justyna Chlebowska, Jakub Gruchota, Kamila Kłosowska-Kosicka, Dorota Adamska, Michał Kamiński, Radosław Salamon, Katarzyna Prokop, and Monika Kusio-Kobiałka for help with selected experiments, Maria Anna Ciemerych-Litwinienko, Dominika Nowis, Jakub Gołąb, Joanna Kufel, Katarzyna Matylla-Kulińska for attentive readings of the manuscript and all Andrzej Dziembowski lab members for fruitful discussions. We thank the expert support of the Mouse Clinic Institute (Illkirch) in mouse construction and handling. Efforts of the Multiple Myeloma Research Foundation (MMRF) and centers that contributed to the CoMMpass study are also acknowledged.

This work was supported by grant funding from the National Science Centre (NCN) to AD (UMO-2013/10/M/NZ4/00299; UMO-2016/22/A/NZ4/00380) and TK (UMO-2019/32/C/NZ2/00558), and grant funding from Ligue Contre le Cancer (Equipe Labellisée 2020), USIAS 2015, within the framework of ANR-10-IDEX-0002-02, ANR-10-LABX-0030-INRT, and ANR-17-EURE-0023 to BS. This research was also supported by funding from the European Union's Horizon 2020 research and innovation program (grant agreement no. 810425).

Author contributions:

TK: conceptualization, all bioinformatic analysis, and experimental work; OG, MM, JD, KK: experimental work; BS: conceptualization and knock-in mice; AD: conceptualization, supervision, original draft preparation.

Competing interests: All authors declare that they have no competing interests.

References

Affer M, Chesi M, Chen WD, Keats JJ, Demchenko YN, Tamizhmani K, Garbitt VM, Riggs DL, Brents LA, Roschke A V, *et al* (2014) Promiscuous MYC locus rearrangements hijack enhancers but mostly super-enhancers to dysregulate MYC expression in

multiple myeloma. *Leukemia* 28: 1725–1735

Alexandrov LB, Jones PH, Wedge DC, Sale JE, Campbell PJ, Nik-Zainal S & Stratton MR (2015) Clock-like mutational processes in human somatic cells. *Nat Genet* 47: 1402–1407

Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR, *et al* (2011) The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* 144: 353–363

Bergstrom EN, Huang MN, Mahto U, Barnes M, Stratton MR, Rozen SG & Alexandrov LB (2019) SigProfilerMatrixGenerator: A tool for visualizing and exploring patterns of small mutational events. *BMC Genomics* 20: 1–12

Carpenter MA, Rajagurubandara E, Wijesinghe P & Bhagwat AS (2010) Determinants of sequence-specificity within human AID and APOBEC3G. *DNA Repair (Amst)* 9: 579–587

Casellas R, Basu U, Yewdell WT, Chaudhuri J, Robbiani DF & Di Noia JM (2016) Mutations, kataegis and translocations in B cells: Understanding AID promiscuous activity. *Nat Rev Immunol* 16: 164–176

Castaneda O & Baz R (2019) Multiple Myeloma Genomics-A Concise Review. *Narrat Rev Acta Medica Acad* 48: 57–67

Cevher MA, Zhang X, Fernandez S, Kim S, Baquero J, Nilsson P, Lee S, Virtanen A & Kleiman FE (2010) Nuclear deadenylation/polyadenylation factors regulate 3′ processing in response to DNA damage. *EMBO J* 29: 1674–1687

Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet JP, Ahmann GJ, Adli M, *et al* (2011) Initial genome sequencing and analysis of multiple myeloma. *Nat 2011 4717339* 471: 467–472

Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E & Alt FW (2003) Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nat 2003 4226933* 422: 726–730

Chen Z & Wang JH (2019) Signaling control of antibody isotype switching. In *Advances in Immunology* pp 105–164. Academic Press Inc.

Cieśla M, Ngoc PCT, Muthukumar S, Todisco G, Madej M, Fritz H, Dimitriou M, Incarnato D, Hellström-Lindberg E & Bellodi C (2023) m6A-driven SF3B1 translation control steers

splicing to direct genome integrity and leukemogenesis. *Mol Cell* 83: 1165-1179.e11

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) Sequence analysis STAR: ultrafast universal RNA-seq aligner. 29: 15–21

Domingo-Prim J, Endara-Coll M, Bonath F, Jimeno S, Prados-Carvajal R, Friedländer MR, Huertas P & Visa N (2019) EXOSC10 is required for RPA assembly and controlled DNA end resection at DNA double-strand breaks. *Nat Commun*

Dziembowski A, Lorentzen E, Conti E & Séraphin B (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat Struct Mol Biol* 14: 15–22

Fonseca R, Oken MM & Greipp PR (2001) The t(4;14)(p16.3;q32) is strongly associated with chromosome 13 abnormalities in both multiple myeloma and monoclonal gammopathy of undetermined significance. *Blood* 98: 1271–1272 doi:10.1182/blood.V98.4.1271 [PREPRINT]

Gadó K, Silva S, Pálóczi K, Domján G & Falus A (2001) Mouse plasmacytoma: An experimental model of human multiple myeloma. *Haematologica* 86: 227–236

Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, Krysiak K, Harris CC, Koboldt DC, Larson DE, *et al* (2011) Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet 2011 441* 44: 53–57

Gritti I, Basso V, Rinchai D, Corigliano F, Pivetti S, Gaviraghi M, Rosano D, Mazza D, Barozzi S, Roncador M, *et al* (2022) Loss of ribonuclease DIS3 hampers genome integrity in myeloma by disrupting DNA:RNA hybrid metabolism . *EMBO J* 41: 1–22

Hoang PH, Cornish AJ, Sherborne AL, Chubb D, Kimber S, Jackson G, Morgan GJ, Cook G, Kinnersley B, Kaiser M, *et al* (2020) An enhanced genetic model of relapsed IGH-translocated multiple myeloma evolutionary dynamics. *Blood Cancer J* 10: 101

Hwang JK, Alt FW & Yeap L-S (2015) Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. *Microbiol Spectr* 3

Kamp JA, Lemmens BBLG, Romeijn RJ, González-Prieto R, Olsen J V., Vertegaal ACO, van Schendel R & Tijsterman M (2022) THO complex deficiency impairs DNA double-strand break repair via the RNA surveillance kinase SMG-1. *Nucleic Acids Res* 50: 6235–6250

Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, Lawrence MS, Kiezun A, Fernandes SM, Bahl S, *et al* (2015) Whole-genome sequencing reveals activation-

induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun* 6: 8866

Keats JJ, Reiman T, Belch AR & Pilarski LM (2006) Ten years and counting: So what do we know about t(4;14)(p16;q32) multiple myeloma. *Leuk Lymphoma* 47: 2289–2300

Kind J, Pagie L, De Vries SS, Nahidiazar L, Dey SS, Bienko M, Zhan Y, Lajoie B, De Graaf CA, Amendola M, *et al* (2015) Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. *Cell* 163: 134–147

Kuliński TM, Gewartowska O, Mroczek S, Szpila M, Sałas K, Liudkovska V & Dziembowski A (2023) Recurrent Multiple Myeloma DIS3 alleles arise early but are later counter-selected due to toxicity. [Manuscript submitted for publication].

Kumar SK, Rajkumar V, Kyle RA, Van Duin M, Sonneveld P, Mateos MV, Gay F & Anderson KC (2017) Multiple myeloma. *Nat Rev Dis Prim* 3: 1–20

Laffleur B, Batista CR, Zhang W, Lim J, Yang B, Rossille D, Wu L, Estrella J, Rothschild G, Pefanis E, *et al* (2022) B CELLS RNA exosome drives early B cell development via noncoding RNA processing mechanisms. 2738: 1–14

Laffleur B, Lim J, Zhang W, Chen Y, Pefanis E, Bizarro J, Batista CR, Wu L, Economides AN, Wang J, *et al* (2021) Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in B cells. *Nat Genet* 53: 230–242

Li H & Barrett J (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27: 2987–2993

Liao Y, Smyth GK & Shi W (2014) Sequence analysis featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. 30: 923–930

Lionetti M, Barbieri M, Todoerti K, Agnelli L, Fabris S, Tonon G, Segalla S, Cifola I, Pinatel E, Tassone P, *et al* (2015) A compendium of *DIS3* mutations and associated transcriptional signatures in plasma cell dyscrasias. *Oncotarget* 6

Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, Sougnez C, Knoechel B, Gould J, Saksena G, *et al* (2014) Widespread genetic heterogeneity in multiple myeloma: Implications for targeted therapy. *Cancer Cell* 25: 91–101

Lovén J, Orlando DA, Sigova AA, Lin CY, Rahl PB, Burge CB, Levens DL, Lee TI & Young RA (2012) Revisiting global gene expression analysis. *Cell* 151: 476–482

Mah LJ, El-Osta A & Karagiannis TC (2010) γh2AX: A sensitive molecular marker of DNA

damage and repair. *Leukemia* 24: 679–686

Manfrini N, Mancino M, Miluzio A, Oliveto S, Balestra M, Calamita P, Alfieri R, Rossi RL, Sassoe-Pognetto M, Salio C, *et al* (2020) FAM46C and FNDC3A Are Multiple Myeloma Tumor Suppressors That Act in Concert to Impair Clearing of Protein Aggregates and Autophagy. *Cancer Res* 80: 4693–4706

Manier S, Salem KZ, Park J, Landau DA, Getz G & Ghobrial IM (2017) Genomic complexity of multiple myeloma and its clinical implications. *Nat Rev Clin Oncol* 14: 100–113

Marin-Vicente C, Domingo-Prim J, Eberle AB & Visa N (2015) RRP6/EXOSC10 is required for the repair of DNA double-strand breaks by homologous recombination. *J Cell Sci* 128: 1097–1107

Maura F, Rustad EH, Yellapantula V, Łuksza M, Hoyos D, Maclachlan KH, Diamond BT, Greenbaum BD, Morgan G, Lesokhin A, *et al* (2020) Role of AID in the temporal pattern of acquisition of driver mutations in multiple myeloma. *Leukemia* 34: 1476–1480

Milbury KL, Paul B, Lari A, Fowler C, Montpetit B & Stirling PC (2019) Exonuclease domain mutants of yeast DIS3 display genome instability. *Nucleus* 10: 21–32

Mroczek S, Chlebowska J, Kuliński TM, Gewartowska O, Gruchota J, Cysewski D, Liudkovska V, Borsuk E, Nowis D & Dziembowski A (2017) The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma. *Nat Commun* 8

Mukarami H, Goto DB, Toda T, Chen ES, Grewal SI, Martienssen RA & Yanagida M (2007) Ribonuclease Activity of Dis3 Is Required for Mitotic Progression and Provides a Possible Link between Heterochromatin and Kinetochore Function. *PLoS One* 2: e317

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y & Honjo T (2000) Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell* 102: 553–563

Murphy MR & Kleiman FE (2020) Connections between 3′ end processing and DNA damage response: Ten years later. *Wiley Interdiscip Rev RNA* 11: 1–24

Niroula A, Urolagin S & Vihinen M (2015) PON-P2: Prediction method for fast and reliable identification of harmful variants. *PLoS One* 10: 1–17

O'Connell J, Schulz-Trieglaff O, Carlson E, Hims MM, Gormley NA & Cox AJ (2015) NxTrim: optimized trimming of Illumina mate pair reads. *Bioinformatics* 31: 2035–2037

Ohkura H, Adachi Y, Kinoshita N, Niwa O, Toda T & Yanagida M (1988) Cold-sensitive and caffeine-supersensitive mutants of the Schizosaccharomyces pombe dis genes

implicated in sister chromatid separation during mitosis. *EMBO J* 7: 1465–1473

Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, Pellagatti A, Wainscoat JS, Hellstrom-Lindberg E, Gambacorti-Passerini C, *et al* (2011) Somatic SF3B1 Mutation in Myelodysplasia with Ring Sideroblasts . *N Engl J Med* 365: 1384–1395

Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, San-Martin BR, Barreto V, *et al* (2010) Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* 143: 122–133

Pefanis E, Wang J, Rothschild G, Lim J, Chao J, Rabadan R, Economides AN & Basu U (2014) Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* 514: 389–393

Pefanis E, Wang J, Rothschild G, Lim J, Kazadi D, Sun J, Federation A, Chao J, Elliott O, Liu ZP, *et al* (2015) RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* 161: 774–789

Potter M (1997) Experimental plasmacytomagenesis in mice. *Hematol Oncol Clin North Am* 11: 323–347

Qiao Q, Wang L, Meng FL, Hwang JK, Alt FW & Wu H (2017) AID Recognizes Structured DNA for Class Switch Recombination. *Mol Cell* 67: 361-373.e4

Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, Honjo T, Nussenzweig A & Nussenzweig MC (2004) AID Is Required for c-myc/IgH Chromosome Translocations In Vivo. *Cell* 118: 431–438

Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Lagelouse R, Gennery A, *et al* (2000) Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2). *Cell* 102: 565–575

Rustad EH, Nadeu F, Angelopoulos N, Zicchedu B, Bolli N, Puente XS, Campo E, Landgren O & Maura F (2021) Mmsig: a Fitting Approach To Accurately Identify Somatic Mutational Signatures in Hematological Malignancies. *Commun Biol* 4: 1–12

Segalla S, Pivetti S, Todoerti K, Chudzik MA, Giuliani EC, Lazzaro F, Volta V, Lazarevic D, Musco G, Muzi-Falconi M, *et al* (2015) The ribonuclease DIS3 promotes let-7 miRNA maturation by degrading the pluripotency factor LIN28B mRNA. *Nucleic Acids Res* 43:

5182–5193

Shinkura R, Tian M, Smith M, Chua K, Fujiwara Y & Alt FW (2003) The influence of transcriptional orientation on endogenous switch region function. *Nat Immunol* 4: 435–441

Skaar JR, Ferris AL, Wu X, Saraf A, Khanna KK, Florens L, Washburn MP, Hughes SH & Pagano M (2015) The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res* 25: 288–305

Snee MJ, Wilson WC, Zhu Y, Chen SY, Wilson BA, Kseib C, O'Neal J, Mahajan N, Tomasson MH, Arur S, *et al* (2016) Collaborative control of cell cycle progression by the RNA exonuclease Dis3 and ras is conserved across species. *Genetics* 203: 749–762

Szczepińska T, Kalisiak K, Tomecki R, Labno A, Borowski LS, Kulinski TM, Adamska D, Kosinska J & Dziembowski A (2015) DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts. *Genome Res* 25: 1622–1633

Takizawa M, Á HT, Li Z, Dubois W, Lim S, Callen E, Franco S, Mosaico M, Feigenbaum L, Alt FW, *et al* (2008) AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J Exp Med* 205: 1949–1957

Taylor J & Lee SC (2019) Mutations in spliceosome genes and therapeutic opportunities in myeloid malignancies. *Genes Chromosom Cancer* 58: 889–902 doi:10.1002/gcc.22784 [PREPRINT]

Todoerti K, Ronchetti D, Favasuli V, Maura F, Morabito F, Bolli N, Taiana E & Neri A (2021) DIS3 mutations in multiple myeloma impact the transcriptional signature and clinical outcome. *Haematologica* 106

Tomecki R, Drazkowska K, Kucinski I, Stodus K, Szczesny RJ, Gruchota J, Owczarek EP, Kalisiak K & Dziembowski A (2014) Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target. *Nucleic Acids Res* 42: 1270–1290

Tomecki R, Kristiansen MS, Lykke-Andersen S, Chlebowski A, Larsen KM, Szczesny RJ, Drazkowska K, Pastula A, Andersen JS, Stepien PP, *et al* (2010) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J* 29: 2342–2357

Vogel MJ, Peric-Hupkes D & van Steensel B (2007) Detection of in vivo protein - DNA interactions using DamID in mammalian cells. *Nat Protoc* 2: 1467–1478

Walker BA, Mavrommatis K, Wardell CP, Cody Ashby T, Bauer M, Davies FE, Rosenthal A, Wang H, Qu P, Hoering A, *et al* (2018) Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. *Blood* 132: 587–597

Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, Werner L, Sivachenko A, DeLuca DS, Zhang L, *et al* (2011) SF3B1 and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia . *N Engl J Med* 365: 2497–2506

Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M, *et al* (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. *Nat 2011 4787367* 478: 64–69

Yu K, Chedin F, Hsieh C-LL, Wilson TE & Lieber MR (2003) R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol 2003 45* 4: 442–451

Zarrin AA, Alt FW, Chaudhuri J, Stokes N, Kaushal D, Du Pasquier L & Tian M (2004) An evolutionarily conserved target motif for immunoglobulin class-switch recombination. *Nat Immunol* 5: 1275–1281

Zeitouni B, Boeva V, Janoueix-Lerosey I, Loeillet S, Legoix-né P, Nicolas A, Delattre O & Barillot E (2010) SVDetect: A tool to identify genomic structural variations from paired-end and mate-pair sequencing data. *Bioinformatics* 26: 1895–1896

Zhang X, Zhang Y, Ba Z, Kyritsis N, Casellas R & Alt FW (2019) Fundamental roles of chromatin loop extrusion in antibody class switching. *Nature*

Figure legends

Figure 1. Generalized growth phenotype of Dis3$^{G766R/+}$ knock-in mice. (A) Body weight male WT and heterozygous littermate Dis3$^{G766R/+}$ mice aged 30- to 50-weeks. (B) Cumulative population doubling of WT and Dis3$^{G766R/+}$ MEFs cultured according to the 3T3 protocol. (C) MA plot showing comparisons of the transcriptome of Dis3$^{G766R/+}$ MEFs with WT MEFs. (D) Representative examples of CellTrace™ Yellow proliferation assay.

Figure 2. Predisposition to plasmacytoma development in Dis3$^{G766R/+}$ knock-in mice. (A) Kaplan-Meier survival plot of male WT and Dis3$^{G766R/+}$ aged mice. (B) Incidence of WT and Dis3$^{G766R/+}$ mice developing diffuse alveolar hemorrhage together with a generalized autoimmune phenotype in response to a peritoneal injection of pristane. (C) Cumulative incidence plot showing the frequency of plasmacytoma in Dis3$^{G766R/+}$ mice and WT mice after the pristane injection. Note that the Dis3^{G766R} mouse line was constructed on a C57BL/6N genetic background, naturally resistant to plasmacytoma development. (D-F) MA plots comparing the transcriptome of naïve (D), activated for 2 days (E) and for 5 days (F) Dis3$^{G766R/+}$ and WT B cells

Figure 3. The Dis3^{G766R} allele has a mutagenic effect on B cells. (A) Circos plot illustrates translocations that involve the IGH locus detected in plasmacytoma cells isolated from Dis3$^{G766R/+}$ mice (red) and WT mice (green). (B) Fraction of chimeric reads in plasmacytoma samples isolated from Dis3$^{G766R/+}$ and WT mice identified using standard whole-genome DNA sequencing (top) and the sequencing of mate-pare libraries (bottom). (C-G) Meta-analyses of genomic regions engaged in translocations with features specific for sites of AID activity: AID ChIP-seq signal (C), GC-rich regions (D), AID-specific DNA recognition sequence motifs (E), predicted G-quadruplex sequences (F), and native transcription measured by GRO-seq (G), relative to random genomic distribution. (H) UCSC genome browser snapshot of a fragment of the IGH locus that shows the clustering of translocation sites with regions that coincide with the enrichment of AID-specific recognition sequence motifs, G-quadruplexes, active bidirectional transcription, and AID occupancy.

Figure 4. Mutational signatures reveal increases in AID and DNA damage repair lesions on MM driver genes in DIS3 mutant patients. (A) Circos graphs depicting the MM driver translocations of the IGH locus in WT DIS3 CoMMpass patients *vs*. patients with somatic non-recurrent and recurrent DIS3 MM variants. (B) Barplot of the enrichment of IGH MM-specific translocations in DIS3 mutant patients of different genotypes compared with WT DIS3 patients. (C-G) Contribution of COSMIC mutational signatures in all somatic single-nucleotide polymorphisms (SNPs) acquired by CoMMpass patients with WT DIS3, a mutant DIS3, or a recurrent MM DIS3 variant in whole exome sequencing (WES): exom-wide (C), in

all somatic DIS3 mutations (D), PROMPTs in genes known to be AID off-targets (E), in genes known to be AID off-targets (F) and in MM driver genes (G).

Figure 5. Mechanism of DIS3-dependent translocations. (A) Quantification of intensities of DamID libraries obtained from activated CH12F3-2A cells that express WT and DIS3^{G766R}. Error bars represent the standard deviation. (B) DIS3 DamID-seq in activated CH12F3-2A cells shows signal enrichment compared with the rest of the genome over regions that translocate in plasmacytomas and represent known DIS3 substrates, such us PROMPTs (Mood's median test). (C) GRO-seq shows the enrichment of native transcription at the same genomic sites. (D) Representative confocal microscopy images demonstrating DIS3 and AID interaction in CH12F3-2A cells detected by *in situ* PLA (interaction site in green, merged with DAPI in blue). Only the negative control giving the highest background among is shown. (E) Quantification the number of DIS3 and AID interaction sites (PLA spots) in CH12F3-2A. Analysis was restricted to cells with at least one PLA spot identified (PLA-positive cells). Statistics were performed with the Wilcoxon test. The black horizontal lines represent means. (F) Representative images demonstrating γ-H2AX staining in naïve and day 3 activated B-cells (green, merged with DAPI in blue). (G) Quantification of number of γ-H2AX immunofluorescent spots per sigle cell. Statistics were performed with the Wilcoxon test. Scale bars, 10 µm.

Figure 6. Immunoglobulin class switch recombination in B and MM cells with DIS3 mutations. (A) No significant changes in frequencies of switching to the IgG1 class are observed in WT and Dis3$^{G766R/+}$ B cells 3 days after activation with lipopolysaccharide and interleukin-4 (p = 1, odds ratio = 0.9). (B-C) Correlation of transcriptome changes in the DIS3$^{G766R/+}$ and DIS3$^{+/+}$ during *in vitro* activation for 3 (B) and 5 (C) days. (D) No significant changes in frequencies of class switching in WT DIS3 patients or patients with MM-specific DIS3 mutations, as determined by the analysis of class expression in RNA-seq experiments.

Figure 7. Model of the involvement of MM-specific DIS3 variants in MM oncogenesis through enhanced AID recruitment to ssDNA. The nuclear exosome that contains WT DIS3 protein (left) recruits AID to transcribed target sequences known RNA exosome properties. The higher affinity of the nuclear RNA exosome complex that contains an MM variant of DIS3 (right) leads to higher occupancy or stalling over certain loci and effectively enhances AID activity on the template strand, thereby causing the promiscuous generation of DNA double-strand breaks and resulting in unrestrained genome rearrangements.

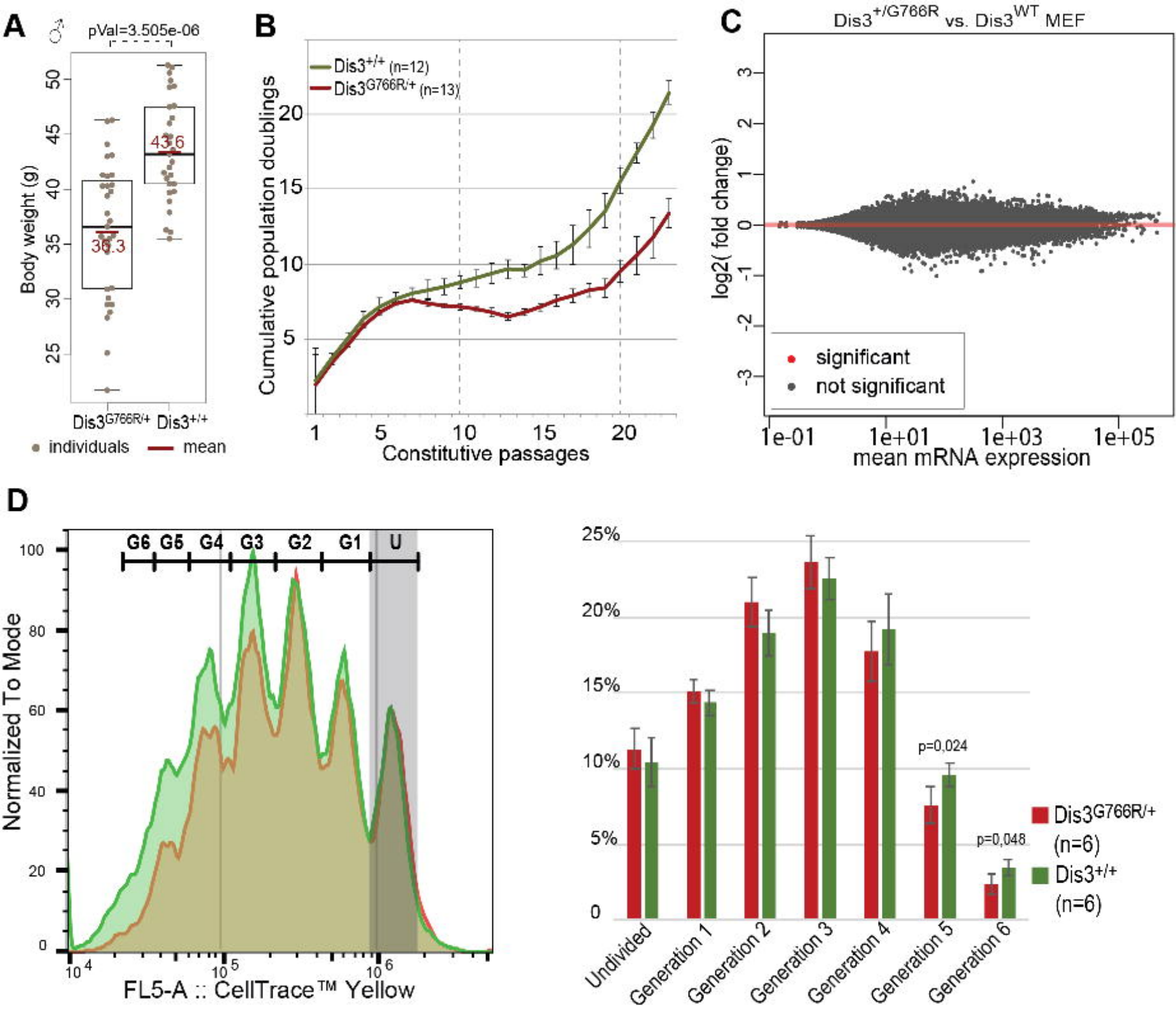
A
♂
pVal=3.505e-06
Body weight (g)
43.6
36.3
Dis3G766R/+
Dis3+/+
individuals
mean
B
Dis3+/+ (n=12)
Dis3G766R/+ (n=13)
Cumulative population doublings
Constitutive passages
C
Dis3+/G766R vs. Dis3WT MEF
log2(fold change)
significant
not significant
mean mRNA expression
D
G6
G5
G4
G3
G2
G1
U
Normalized To Mode
FL5-A :: CellTrace™ Yellow
p=0,024
p=0,048
Dis3G766R/+ (n=6)
Dis3+/+ (n=6)
Undivided
Generation 1
Generation 2
Generation 3
Generation 4
Generation 5
Generation 6

Figure 1.

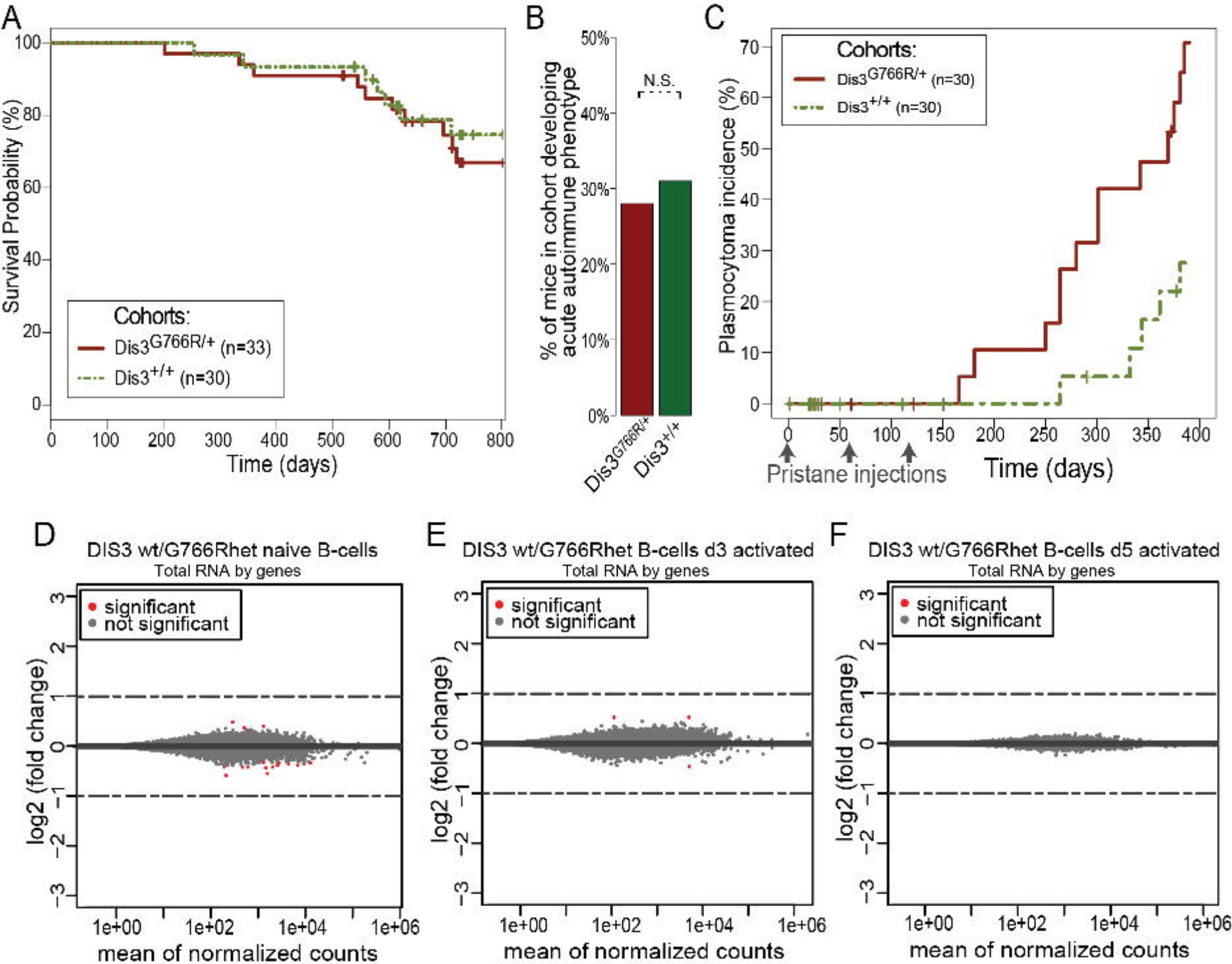
A
Survival Probability (%)
Time (days)
Cohorts:
Dis3G766R/+ (n=33)
Dis3+/+ (n=30)
B
% of mice in cohort developing acute autoimmune phenotype
N.S.
Dis3G766R/+
Dis3+/+
C
Plasmocytoma incidence (%)
Cohorts:
Dis3G766R/+ (n=30)
Dis3+/+ (n=30)
Pristane injections
Time (days)
D
DIS3 wt/G766Rhet naive B-cells
Total RNA by genes
significant
not significant
log2 (fold change)
mean of normalized counts
E
DIS3 wt/G766Rhet B-cells d3 activated
Total RNA by genes
significant
not significant
log2 (fold change)
mean of normalized counts
F
DIS3 wt/G766Rhet B-cells d5 activated
Total RNA by genes
significant
not significant
log2 (fold change)
mean of normalized counts

Figure 2.

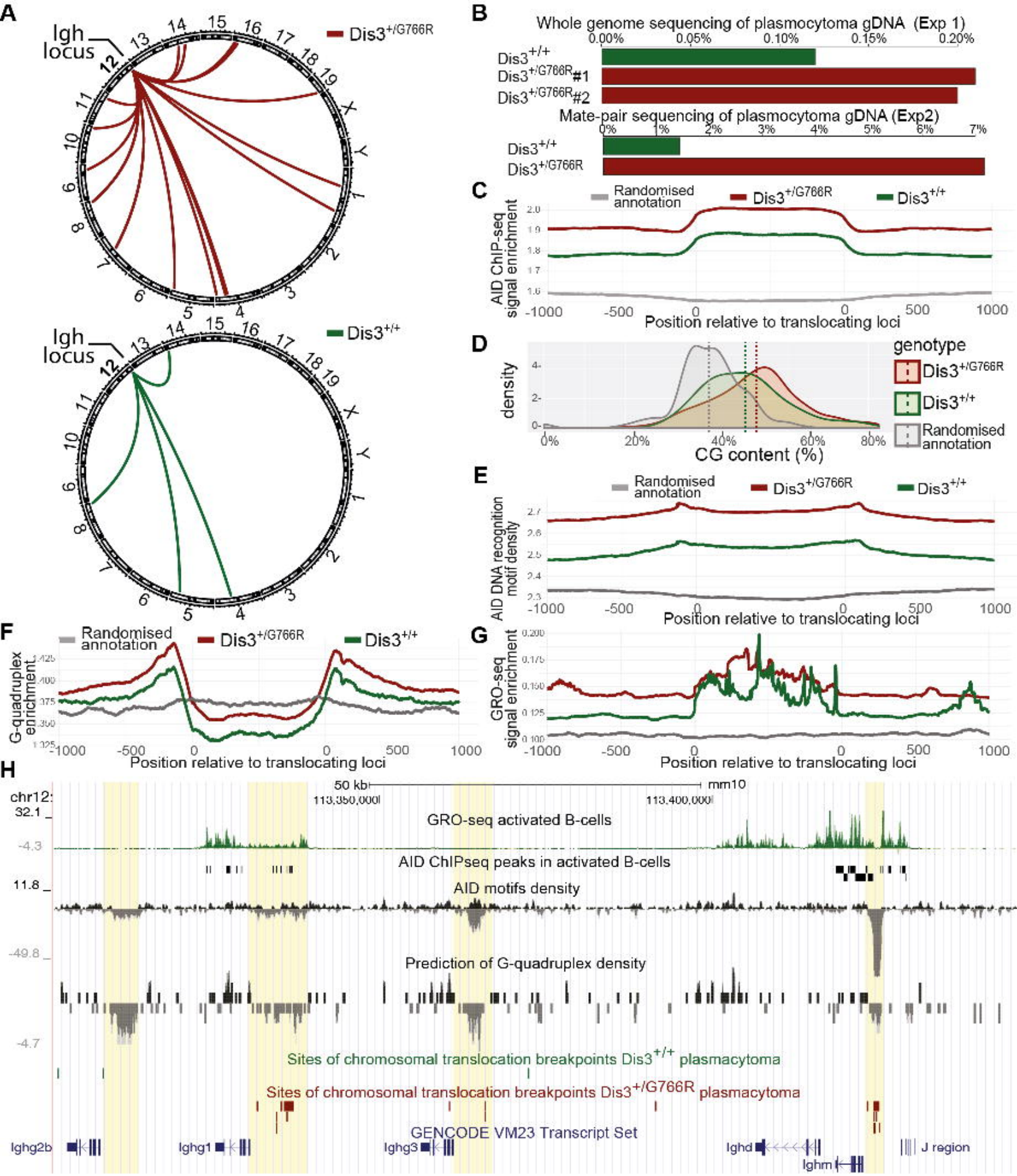
A
Igh locus
Dis3+/G766R
Dis3+/+
B
Whole genome sequencing of plasmocytoma gDNA (Exp 1)
Dis3+/+
Dis3+/G766R#1
Dis3+/G766R#2
Mate-pair sequencing of plasmocytoma gDNA (Exp2)
Dis3+/+
Dis3+/G766R
C
Randomised annotation
AID ChIP-seq signal enrichment
Position relative to translocating loci
D
density
genotype
CG content (%)
E
AID DNA recognition motif density
F
G-quadruplex enrichment
G
GRO-seq signal enrichment
H
chr12:
50 kb
mm10
113,350,000
113,400,000
GRO-seq activated B-cells
AID ChIPseq peaks in activated B-cells
AID motifs density
Prediction of G-quadruplex density
Sites of chromosomal translocation breakpoints Dis3+/+ plasmacytoma
Sites of chromosomal translocation breakpoints Dis3+/G766R plasmacytoma
GENCODE VM23 Transcript Set
Ighg2b
Ighg1
Ighg3
Ighd
Ighm
J region

Figure 3.

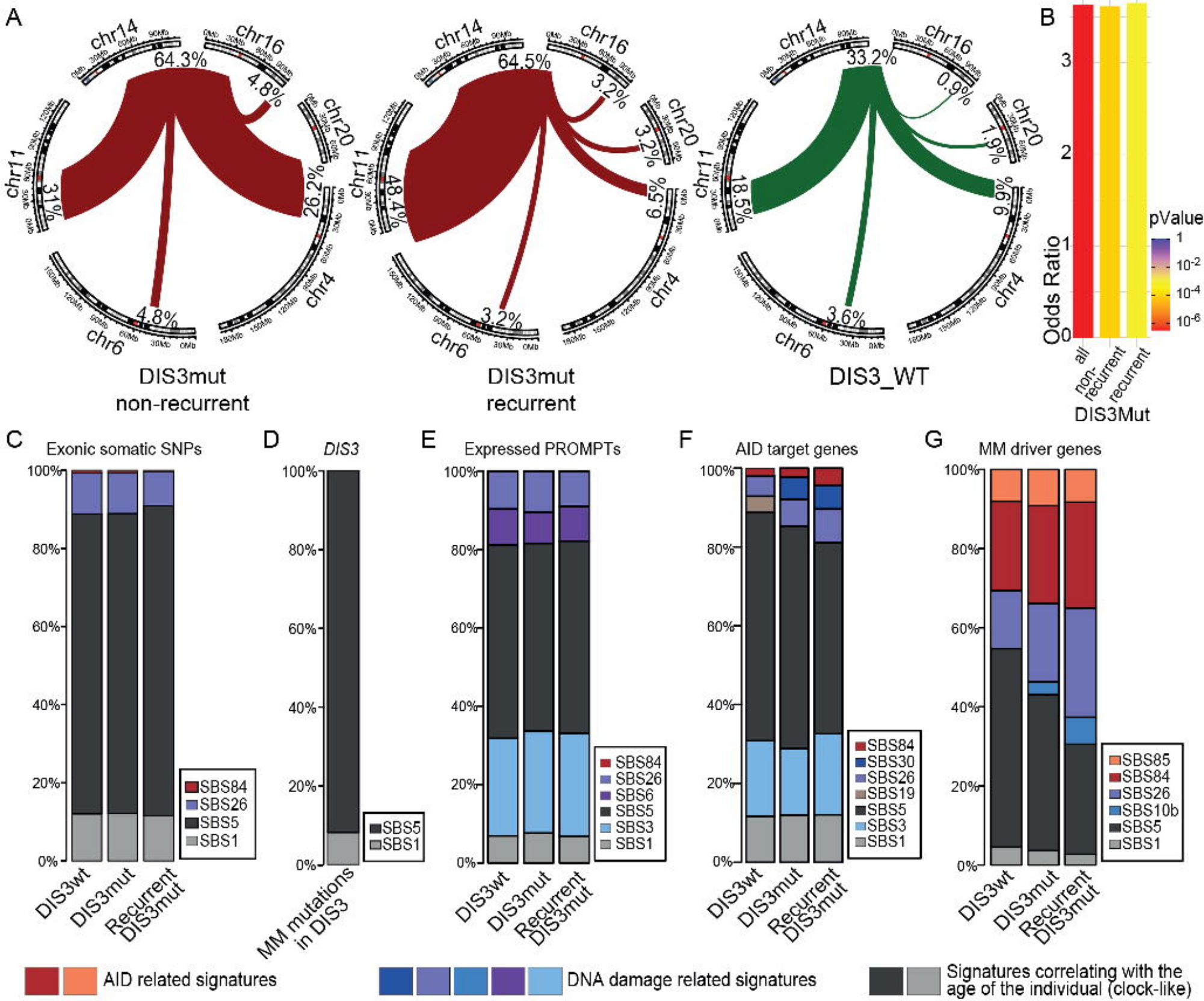
A
chr14
chr16
chr20
chr4
chr6
chr11
64.3%
4.8%
26.2%
31%
4.8%
DIS3mut
non-recurrent
64.5%
3.2%
3.2%
6.5%
48.4%
3.2%
DIS3mut
recurrent
33.2%
0.9%
1.9%
9.9%
18.5%
3.6%
DIS3_WT
B
Odds Ratio
pValue
all
non-recurrent
recurrent
DIS3Mut
C
Exonic somatic SNPs
D
DIS3
E
Expressed PROMPTs
F
AID target genes
G
MM driver genes
DIS3wt
DIS3mut
Recurrent DIS3mut
MM mutations in DIS3
SBS84
SBS26
SBS5
SBS1
SBS6
SBS3
SBS30
SBS19
SBS85
SBS10b
AID related signatures
DNA damage related signatures
Signatures correlating with the age of the individual (clock-like)

Figure 4.

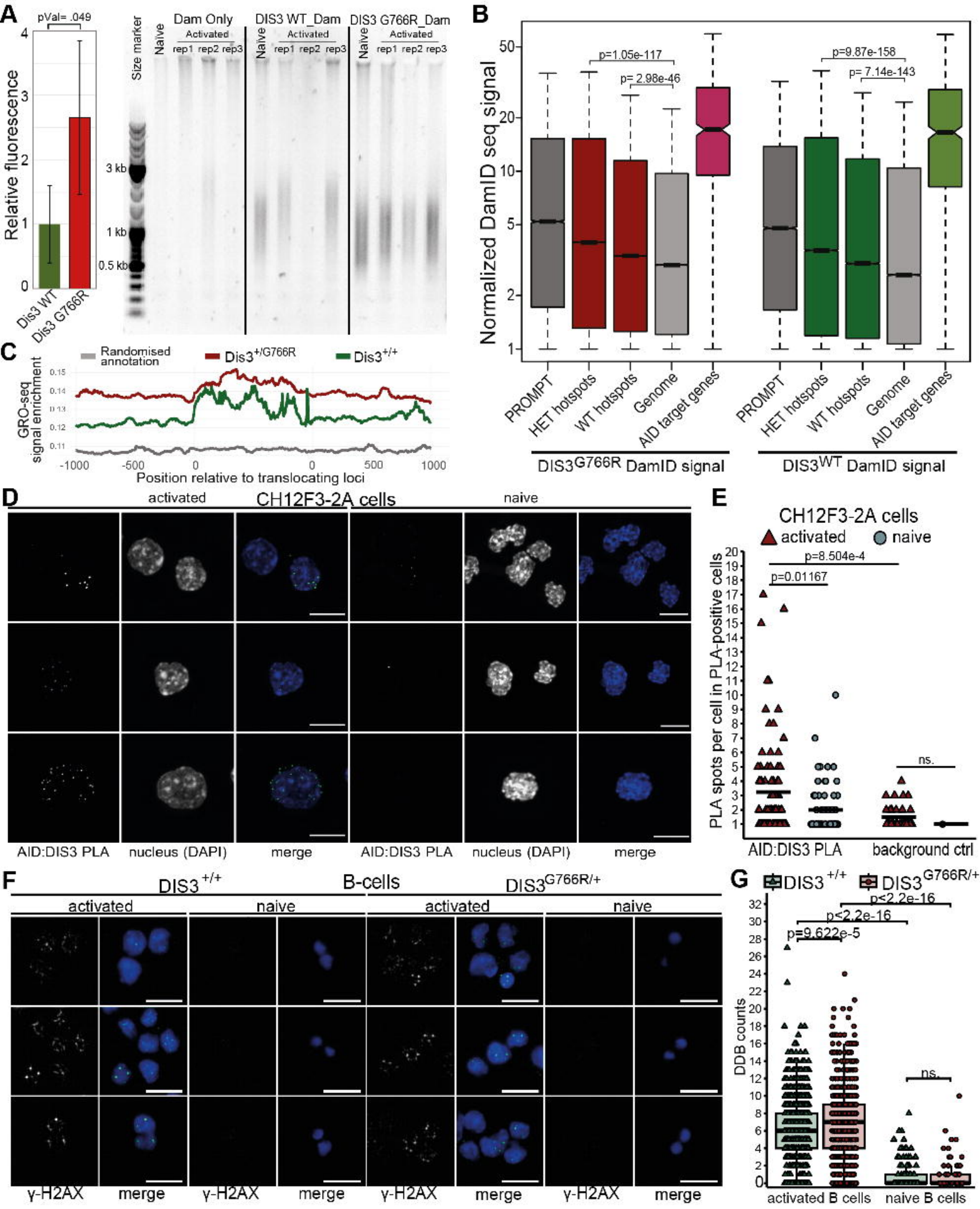
A
pVal= .049
Relative fluorescence
Dis3 WT
Dis3 G766R
Size marker
Dam Only
DIS3 WT_Dam
DIS3 G766R_Dam
Naive
Activated
rep1 rep2 rep3
3 kb
1 kb
0.5 kb
B
Normalized DamID seq signal
p=1.05e-117
p= 2.98e-46
p=9.87e-158
p= 7.14e-143
PROMPT
HET hotspots
WT hotspots
Genome
AID target genes
DIS3G766R DamID signal
DIS3WT DamID signal
C
Randomised annotation
Dis3+/G766R
Dis3+/+
GRO-seq signal enrichment
Position relative to translocating loci
D
activated
CH12F3-2A cells
naive
AID:DIS3 PLA
nucleus (DAPI)
merge
E
CH12F3-2A cells
activated
naive
PLA spots per cell in PLA-positive cells
p=8.504e-4
p=0.01167
ns.
AID:DIS3 PLA
background ctrl
F
DIS3 +/+
B-cells
DIS3G766R/+
activated
naive
γ-H2AX
merge
G
DIS3 +/+
DIS3G766R/+
p<2.2e-16
p<2.2e-16
p=9.622e-5
ns.
DDB counts
activated B cells
naive B cells

Figure 5.

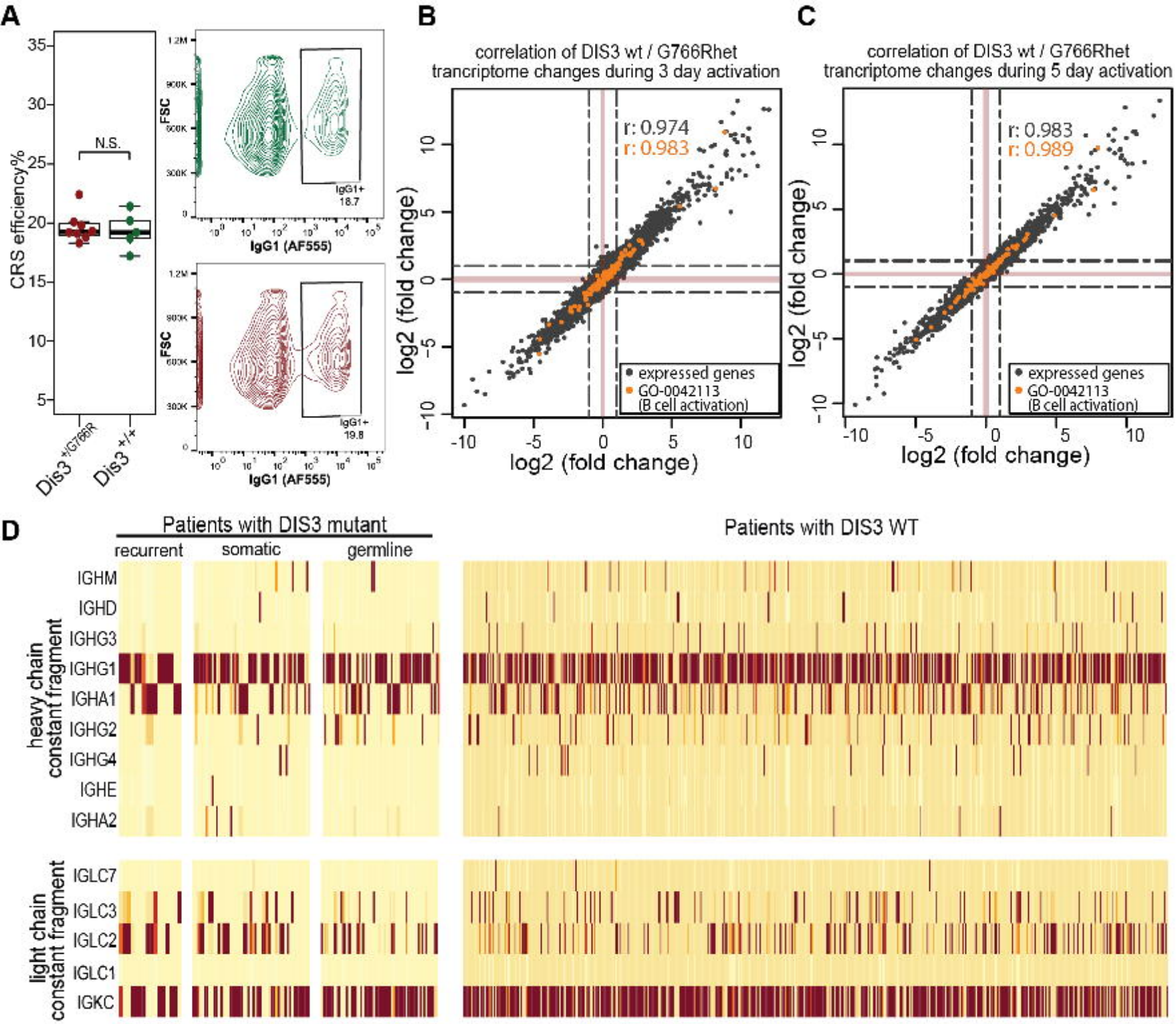
A
CRS efficiency%
N.S.
Dis3 +/G766R
Dis3 +/+
FSC
IgG1 (AF555)
IgG1+
18.7
IgG1+
19.8
B
correlation of DIS3 wt / G766Rhet
trancriptome changes during 3 day activation
r: 0.974
r: 0.983
log2 (fold change)
expressed genes
GO-0042113
(B cell activation)
C
correlation of DIS3 wt / G766Rhet
trancriptome changes during 5 day activation
r: 0.983
r: 0.989
log2 (fold change)
expressed genes
GO-0042113
(B cell activation)
D
Patients with DIS3 mutant
recurrent
somatic
germline
Patients with DIS3 WT
heavy chain constant fragment
IGHM
IGHD
IGHG3
IGHG1
IGHA1
IGHG2
IGHG4
IGHE
IGHA2
light chain constant fragment
IGLC7
IGLC3
IGLC2
IGLC1
IGKC

Figure 6.

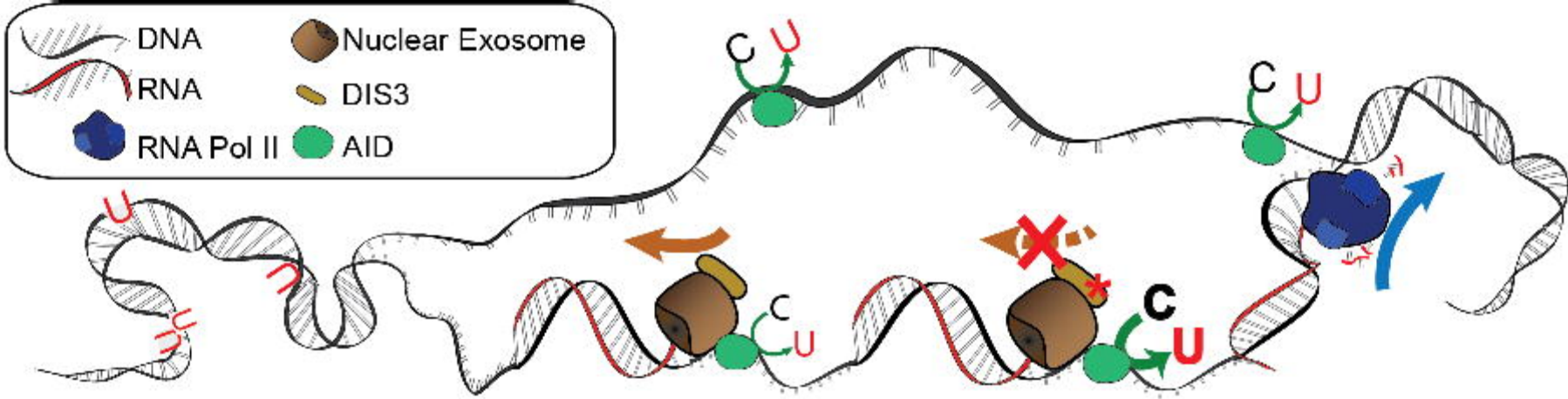

Figure 7.