The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma

Seweryn Mroczek, Justyna Chlebowska, Tomasz M. Kulinński, Olga Gewartowska, Jakub Gruchota, Dominik Cysewski, Vladyslava Liudkovska, Ewa Borsuk, Dominika Nowis & Andrzej Dziembowski

FAM46C is one of the most frequently mutated genes in multiple myeloma. Here, using a combination of in vitro and in vivo approaches, we demonstrate that FAM46C encodes an active non-canonical poly(A) polymerase which enhances mRNA stability and gene expression. Reintroduction of active FAM46C into multiple myeloma cell lines, but not its catalytically-inactive mutant, leads to broad polyadenylation and stabilization of mRNAs strongly enriched with those encoding endoplasmic reticulum-targeted proteins and induces cell death. Moreover, silencing of FAM46C in multiple myeloma cells expressing WT protein enhance cell proliferation. Finally, using a FAM46C-FLAG knock-in mouse strain, we show that the FAM46C protein is strongly induced during activation of primary splenocytes and that B lymphocytes isolated from newly generated FAM46C KO mice proliferate faster than those isolated from their WT littermates. Concluding, our data clearly indicate that FAM46C works as an onco-suppressor, with the specificity for B-lymphocyte lineage from which multiple myeloma originates.

DOI: 10.1038/s41467-017-00578-5
mass sequencing of cancer genomes has revealed genomic landscapes of potential tumor suppressors and oncogenes. FAM46C, a gene whose physiological function was largely unknown, is one of the most frequently mutated genes in multiple myeloma (MM), following the well-known proto-oncogenes, KRAS, NRAS, and BRAF. Deletions of FAM46C gene 1p12 locus (del(1p)) have been found in ~20% of MM cases and are associated with short progression-free survival and decreased overall survival. Except of chromosomal aberrations, recurrent homozygotic or hemizygotic somatic point mutations have been identified in about 10% of MM cases, depending on studies based on whole-genome- or whole-exome sequencing. To date, more than 70 unique somatic mutations across whole FAM46C gene sequence have been identified, many of which are frameshift or nonsense mutations (https://research.themmrf.org).

Importantly, FAM46C mutations are specific to MM since no other cancer type with statistical significant enriched in FAM46C mutations has been described so far. The high frequency of mutation in the FAM46C gene allowed it to be classified as MM driver-gene, which may function as a tumor suppressor even though it does not contain mutational hotspots. FAM46C mutations are also frequently found in stable human myeloma cell lines. In addition, FAM46C has been identified as a type I interferon-stimulated gene, overexpression of which slightly enhances replication of some viruses. FAM46C belongs to a FAM46 metazoan-specific family of proteins, which has 4 members in humans that are very similar at the protein level with a degree of sequence identity of at least 56.9%. There is currently very little functional data on FAM46 and positional cloning in mouse revealed that mutations in the gene cause anemia. The only publication about this FAM46 family of proteins. Positional cloning in mouse revealed that mutations in the driver-gene, which may function as a tumor suppressor even though it does not contain mutational hotspots.

FAM46C encodes a poly(A) polymerase that enhances gene expression. The FAM46 family of proteins exists only in animals. In vertebrates, all its members have the same architecture. They contain domains that are very distinctly related to the catalytic and associated domains of poly(A) polymerases, and they lack any detectable RNA-binding domains. The putative catalytic residues of FAM46C are preserved, indicating that this protein may indeed be an active poly(A) or poly(U) polymerase (Supplementary Fig. 1).

Results

FAM46C encodes a poly(A) polymerase that enhances gene expression. The FAM46 family of proteins exists only in animals. In vertebrates, all its members have the same architecture. They contain domains that are very distinctly related to the catalytic and associated domains of poly(A) polymerases, and they lack any detectable RNA-binding domains. The putative catalytic residues of FAM46C are preserved, indicating that this protein may indeed be an active poly(A) or poly(U) polymerase (Supplementary Fig. 1).

---

**Fig. 1** FAM46C interacts with RNA and is an active RNA poly(A) polymerase in vitro and in vivo. **a** Recombinant FAM46C WT (lanes 8-13), but not its catalytic mutant FAM46C mut (lanes 2-7), displays poly(A) polymerase activity in vitro. Reaction products (using 32P-labeled (A)15 as substrate) were separated in denaturing PAGE gels and visualized by autoradiography. **b** SDS-PAGE analysis of recombinant FAM46CWT and its catalytic mutant FAM46Cmut. **c** FAM46CWT is an active poly(A) polymerase in vitro and requires Mn2+ ions for its activity. Purified protein was incubated with 32P-labeled (A)15 primer in the presence of ATP and divalent cations as follows: Mg2+ (lanes 4-6), both Mg2+/Mn2+ (lanes 7-9), or Mn2+ (lanes 10-12). Control reactions were carried out without the protein (lanes 1-3). **d** FAM46C interacts with RNA in human cells. Autoradiography of UV cross-linked 32P-labeled RNAs co-purified with FAM46CWT-GFP from HEK293 cells stably expressing the fusion protein (lanes 3-4) and from control empty cells (lanes 1-2). Immunoprecipitated RNA-protein complexes were separated by SDS-PAGE, transferred to nitrocellulose membrane, stained with Ponceau S, and subsequently autoradiographed. The right panel shows the Ponceau S stained blot merged with autoradiogram.

---

**ARTICLE**

**NATURE COMMUNICATIONS** | DOI: 10.1038/s41467-017-00578-5

---

2 NATURE COMMUNICATIONS | 8:619 | DOI: 10.1038/s41467-017-00578-5 | nature.com/naturecommunications
In order to confirm predicted molecular activity, we purified recombinant FAM46C and its mutant version, in which acidic residues of the putative catalytic center, expected to coordinate divalent cations, were replaced by alanines (D90A and D92A; FAM46C<sup>mut</sup>). This was followed by an in vitro polyadenylation assay using an A<sub>15</sub> oligoribonucleotide RNA as a substrate. Wild type, but not catalytic mutant, protein was able to extend the RNA substrate in vitro in the presence of ATP and divalent metal ions (Mn<sup>2+</sup>) (Fig. 1a, b), indicating that FAM46C is a poly(A) polymerase. However, due to the relatively low activity of recombinant FAM46C (likely related to its intrinsic low solubility and aggregation), we validated the result using recombinant FAM46D, which was more soluble; the sequences of the two proteins are highly similar (Supplementary Fig. 1). FAM46D displayed poly(A) polymerase activity and was able to extend A<sub>15</sub> oligoribonucleotide substrate; its activity was significantly higher compared to that of FAM46C (Fig. 1c).

In order to determine the molecular function of FAM46C, we evaluated whether this protein could interact with RNA at the cellular level. Stable HEK293 cells expressing FAM46C<sup>WT</sup>-GFP were UV cross-linked and the tagged proteins were immunoprecipitated using anti-GFP antibodies. Co-purified RNA was visualized by 32P-labeling of RNA-protein covalent crosslinks. Strong enrichment of RNA was observed in cross-linked, FAM46C-positive samples compared to control cells and non-crosslinked samples (Fig. 1d), demonstrating that FAM46C binds to RNA and may act as an RNA poly(A) polymerase in human cells.

Non-canonical poly(A) polymerases may enhance overall gene expression by extending the poly(A) tails of dormant messenger RNAs (like GLD2) in the cytoplasm or trigger rapid exosome-mediated RNA decay by oligoadenylation of target RNA in the nucleus (as in the case of TRF4) [15]. In order to examine intracellular localization of the FAM46C protein, HEK293T, RPMI826 (MM), H929 (MM), SKMM1 (MM), HL60 (promyelocytic leukemia) and Raji (B-cell lymphoma) cell lines were transduced with constructs encoding the FAM46C<sup>WT</sup>-GFP fusion gene. Confocal microscopy revealed simultaneous nuclear and cytoplasmic localization of FAM46C in all tested cell lines, as expected for non-canonical poly(A) polymerases (Fig. 2), demonstrating the possibility of both RNA stabilizing and destabilizing functions.

To further analyze FAM46C function in cells, we performed an RNA-tethering assay. HEK293 cells were co-transfected with a construct expressing <i>Renilla</i> luciferase (RL) containing five boxB sites in its 3′-UTR, Firefly luciferase (FL) control reporter and FAM46C<sup>WT</sup> harboring the N-terminal λN boxB-binding domain and HA-tag [19]. After 24 h, the levels of reporter protein were analyzed, which revealed that tethering a wild-type FAM46C enhances expression of RL more than 7-fold (Fig. 3a, d). Several additional controls were applied to prove that the catalytic activity of FAM46C was indeed responsible for the enhanced expression of RL reporter: (1) tethering of the catalytic FAM46C mutant had little effect on RL reporter expression (Fig. 3a, d); (2) a RL reporter with a cyclic phosphate at the 3′-end generated by a hammerhead ribozyme, which cannot be polyadenylated, was insensitive to FAM46C tethering (Fig. 3b, d); (3) expression of FAM46C without the λN domains did not enhanced expression of the reporter (Fig. 3c, d). At the RNA level, northern blot analysis revealed that enhanced expression of RL upon FAM46C<sup>WT</sup> tethering correlated with increased steady state levels and slower migration of its mRNA (Fig. 3e, f). This was not observed for the RL reporter with a cyclic phosphate at the 3′ (Fig. 3e, f). Quantifications of control FL reporter mRNA revealed no effect of FAM46C on its expression (Fig. 3g). Furthermore, northern blot analysis of poly(A) enriched-RNA fractions revealed significant increases in steady-state levels and slower migration of RL mRNA in samples where FAM46C<sup>WT</sup> was tethered (Fig. 3h). Next, we confirmed that the observed RL mRNA lengthening is indeed caused by polyadenylation of mRNA 3′-end using an RNase H cleavage assay with oligo(dT)<sub>20</sub> (Fig. 3i). In addition, RNA from cells after FAM46C was tethered to RL mRNA was fractionated based on the lengths of poly(A) tails using oligo(dT) sepharose resin. Northern blot analysis revealed that more RL mRNA molecules were present in fractions with longer poly(A) tails, confirming that it is polyadenylated by FAM46C (Supplementary Fig. 2a). The same effect was observed for FAM46D indicating that they have similar properties (Supplementary Fig. 2b).

In order to determine whether the observed increase in reporter expression was due to mRNA stability, we evaluated the half-life of RL mRNA. Cells were transfected with pRL-5BoxB and pNHAFAM46C<sup>WT</sup> or pNHAFAM46C<sup>mut</sup>. After 24 h, transcription was inhibited for 4, 8, and 12 h by the addition of actinomycin D. Northern blot, followed by densitometry...
FAM46C tethering leads to polyadenylation and enhanced expression of a Renilla luciferase (RL) reporter. Analysis at the protein level (a-d): a FAM46C tethering increases RL reporter protein levels. HEK293 cells were co-transfected with a construct expressing RL containing five boxB sites in its 3′-UTR, Firefly luciferase (FL) control reporter and FAM46CWT harboring the N-terminal HA-boxB-binding domain and HA-tag. Western blot detection of RL in mock-transfected cells (lane 1) and after NHA-FAM46CWT (lane 2) or NHA-FAM46Cmut (lane 3) tethering. Expression of NHA-tagged FAM46C proteins were confirmed using an α-HA antibody. DBC1 served as a loading control. Asterisks indicate cross-hybridization signals. b Expression of RLSboxHSL + HhR reporter with a cyclic phosphate at the 3′ end generated by a hammerhead ribozyme was not enhanced upon FAM46C tethering. The experiment was performed as in a. c Expression of the FAM46C without λN domains did not enhance expression of the reporter. The experiment was performed as in a. d Quantifications of RL protein normalized by the internal FL reporter related to experiments from a–c. RNA analysis (e-i): e northern blot detection of RL mRNA using total RNA from HEK293 cells after tethering of NHA-FAM46CWT or NHA-FAM46Cmut to RLSbox (lanes 1–3) or RLSboxHSL + HhR with a cyclic phosphate at the 3′ end (lanes 4–6). f Quantifications of RL mRNA. g RT-qPCR analysis of the FL control reporter. h Northern blot detection of Renilla luciferase using poly(A) + fraction from HEK293 cells after tethering of NHA-FAM46CWT or NHA-FAM46Cmut (i) Poly (A) tails added to reporter mRNA can be removed by RNase H treatment in the presence of oligo(dT). High-resolution northern blot analysis of RL mRNA from control HEK293 cells (lanes 1–2), after tethering of NHA-FAM46CWT (lanes 3–4) or NHA-FAM46Cmut (lanes 5–6). The data in d, f, g are shown as a mean value ± SD (n = 3).

FAM46C is an MM cell growth suppressor. To study FAM46C in the context of MM pathogenesis, we took advantage of the established MM cell lines, some of which harbor early frameshift mutations in FAM46C: SKMM1 (Homozygous p.1173fsX36) and H929 (hemizygous p.1936fsX15) (Supplementary Table 1). In order to reintroduce wild-type FAM46C into these cell lines, we designed a lentiviral vector allowing for the expression of C-terminally GFP-tagged protein under the control of the spleen focus-forming virus (SFFV) promoter, which is active in hematopoietic cells. Initially, we evaluated transduction efficiency and protein expression levels using control HEK293T cells, which were transduced with lentiviruses carrying FAM46CWT-GFP, FAM46Cmut-GFP, or GFP as control, using different multiplicities of infection (MOIs; from 1 to 4). In all cases, stable cell lines were obtained with >95% efficiency as verified by flow cytometry for GFP-expressing living cells (Supplementary Fig. 3a). No significant growth phenotypes were observed in HEK293T cells. A cytometric analysis of the cell cycle revealed higher number of cells in the G0/G1 phase; however, this observation could be a side effect associated with lentiviral particles, as it was not correlated with the activity of the FAM46C protein.

Next, we transduced SKM1, H929 (both mutated in the endogenous FAM46C locus), and RPMI8226 (with wild-type FAM46C sequence) with lentiviruses carrying FAM46CWT-GFP, FAM46Cmut-GFP, and GFP constructs at MOIs from 1 to 4. Cytometric analysis of cell cultures was used to monitor the effect of FAM46C expression up to 11 days post transduction. Three days after transgene delivery, most of the tested cell lines expressing FAM46CWT had reduced growth rate and increased number of dead cells (Fig. 4a, b); this effect was positively correlated with high FAM46CWT virus titer levels (Supplementary Fig. 4b). Interestingly, in SKMM1 cells, the cytopastic effect of FAM46CWT expression was observed later in comparison to H929 or RPMI8226 cells and was best visible from the ninth day after transduction. Cytostatic effect was also observed for constructs with other tags at the N and C termini of FAM46C.
**Fig. 4** FAM46C control survival and proliferation of multiple myeloma cells. **a–c** Expression of FAM46C WT induces death in multiple myeloma cells. 

**a** An example gating strategy for defining transduction efficiency and cell viability. Forward scatter (FSC) and side scatter (SSC) gate were used to separate debris from intact cells. Viability of RPMI8226 cells overexpressing either FAM46C WT-GFP or FAM46Cmut-GFP was analyzed using propidium iodide (PI) staining on the 11th day post-transduction. GFP expression (GFP+) was evaluated in parallel in PI-negative cells. 

**b** Summary of flow cytometry analyses presented as bar graphs showing GFP expression level and reduced viability of multiple myeloma cell lines (SKMM1, H929, and RPMI8226) overexpressing either FAM46C WT-GFP or FAM46Cmut-GFP. 

**c** Flow cytometry analyses for RPMI8226 in **b** and Raji and HL60 in **e** throughout the time course of GFP, FAM46C WT-GFP, and FAM46Cmut-GFP expression. The data are presented as percentage of cells ± SD (n = 3). P values were calculated using two-way ANOVA tests (*P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001). 

**d** shRNA-mediated silencing of FAM46C enhances proliferation rate of RPMI8226 MM cell line expressing wild-type protein but not SKMM1 harboring FAM46C mutation. Cells were transduced with lentiviral vectors expressing shRNA and empty vector as control. Stable transduced cell lines were stained with CFSE. Cell division was monitored by flow cytometry after 48 and 72 h by levels of CFSE dilution. 

**e** Reverse transcription qPCR analysis of the FAM46C silencing efficiency. Bars represent mean values ± SD (n = 3). 

**P** values were calculated using Student’s t-test (*P < 0.05, (n = 3).
protein (FLAG-FAM46CWT, FAM46CWT-FLAG, GFP-FAM46CWT, and NHA-FAM46CWT; Supplementary Fig. 4). We also tested other MM cell lines (MM1.S, MM1.R) and observed the same effects. Importantly, in contrast to FAM46CWT-GFP, expression of FAM46Cmut-GFP or GFP only, did not cause cell death and stable MM cell lines were established (Fig. 4a, b). This observation indicates that cytotoxic and cytostatic effects of FAM46C overexpression depend on its enzymatic activity. Moreover, overexpression of GFP-tagged nucleotidyltransferases such as canonical (PAPOLA) or non-canonical poly(A) polymerases (POLS or GLD2) in SKMM1 cell lines did not cause strong cytostatic effects (Supplementary Fig. 5), since the toxicity was visible only at early time points post transduction and stable cell lines were derived. This suggests a unique influence of FAM46C on MM cell viability.

Finally, in order to determine if FAM46C toxicity is restricted to MM cells, other hematological cell lines, including HL60 (promyelocytic leukemia) and Raji (B-cell lymphoma), were transduced with lentiviruses carrying FAM46CWT-GFP, FAM46Cmut-GFP, and GFP. Unlike in MM cells, expression of FAM46C had no effect on HL60 or Raji cell growth and viability, and these lines were established after cell sorting (Fig. 4c). It was concluded that FAM46C expression specifically induces cell death of MM cells, however, we cannot exclude that reduced toxicity of FAM46C for HL60 and Raji cell is not related to lower expression levels of the transgene.

Furthermore, in order to verify if FAM46C may positively regulate gene expression in MM cells we have conducted tethering assays performed similarly to the one described for HEK293 cells but implemented for SKMM1 and H929 cells. As expected, tethering of FAM46CWT but not FAM46Cmut in MM cell lines enhanced the expression of the reporter mRNA (Supplementary Fig. 6).

Mutations in FAM46C are found only in a fraction of MM cases. In order to check whether the effect of FAM46C on cell survival and proliferation is general to MM cells, we silenced FAM46C expression in RPMI8226, a MM cell line expressing wild-type FAM46C. To this end, we derived stable cell lines transduced with lentiviral vectors expressing three different shRNA targeting FAM46C. Although the level of silencing was moderate (~50%, Fig. 4d), basically all shRNAs led to an increased proliferation rate of RPMI8226 cell line as measured using CFSE cell tracer (Fig. 4e). Such an effect was not observed in the SKMM1 MM cell line harboring FAM46C mutations (Fig. 4e). This result strongly suggests that FAM46C suppresses the growth of all MM cells.
FAM46C KO enhances B lymphocyte proliferation rate. The analysis of MM cell lines revealed that FAM46C expression affects their growth and survival. Our intensive trials to raise or purchase specific antibodies that recognized FAM46C failed and so we were unable to verify endogenous FAM46C levels in MM cells. Thus, to study the expression of FAM46C in the B-cell lineage, from which MM originates, we have generated a FAM46C-KO knock-in mouse using CRISPR/Cas9 methodology. The animals did not display any detectable phenotypes. We have used western blot analysis of IL4 and LPS activated splenocytes and found that FAM46C expression was very strongly enhanced upon activation (Fig. 5a). Furthermore, cell fractionation experiments revealed that it is localized mainly in the cytoplasm (Fig. 5b), which is generally in agreement with our localization studies based on confocal imaging. Such result suggests that FAM46C may act as a non-canonical poly(A) polymerase in the B lymphocyte lineage.

In order to study the potential role for FAM46C in B lymphocytes and more generally in hematopoesis, we have generated FAM46C KO mutant animals, which harbor a 17 bp deletion and a 301 bp insertion using the CRISPR. This mutation causes a frameshift at position C88 and premature FAM46C protein termination, which takes place 26 amino acids after the breakpoint. Although, detailed analysis of FAM46C KO phenotypes are outside of the scope of this study, the mice did not display any major developmental phenotypes. However, we did detect several hematologic abnormalities (Fig. 6a–h). Heterozygous FAM46C KO mice (FAM46C+/−) had significantly lower hemoglobin level than age- and sex-matching FAM46C+/+ controls, suggesting that they suffer from anemia. The red blood cells count was also increased while MCV, MCH and MCHC were all significantly decreased. The microcytic and hypochromic erythrocytes imply that insufficiency of hemoglobin production is the most probable cause of the anemia. The MCV/RBC ratio (Mentzer index) is lower in FAM46C−/− (2.52) than in FAM46C+/− (4.39), suggesting the anemia is a consequence of a block in globin synthesis rather than iron insufficiency. This hypothesis has to be further tested experimentally. These results are generally in agreement with the unpublished work from the Fleming laboratory described in the Meng Tian PhD thesis, but runs contrary to the Mouse Phenotypic Consortium, which has found FAM46C an essential gene. We suppose, that this discrepancy may be due to some linked lethal mutation generated by the consortium.

Importantly, analysis of in culture-activated primary splenic B lymphocytes revealed that FAM46C mutation leads to an
increased proliferation rate, what indicates that FAM46C decreases the growth rate of B cells (Fig. 6i). In conclusion, FAM46C acts as a growth suppressor not only for MM cells but also more generally in B-cell lineage cells.

**FAM46C polyadenylates multiple ER-targeted protein mRNAs.**

In order to identify FAM46C substrates in MM cells, deep sequencing of total RNA from SKMM1 and H929 MM cell lines was performed after 3 days of FAM46CWT-GFP and FAM46Cmut-GFP overexpression. RNA samples were prepared in triplicates or quadruplicates and depleted of rRNA, and strand-specific total RNA libraries were prepared. The samples were sequenced and mapped to an average depth of 15 million uniquely aligning reads. Differential expression analyses revealed significant differences in transcriptomes and responses to FAM46C expression among tested MM cell lines (Fig. 7a, b and Supplementary Data 1). When FAM46C was overexpressed, H929 cells had an increase in steady-state levels of many transcripts that encode proteins involved in the interferon response. This effect was not observed in the SKMM1 cell line. Nevertheless, there was a significant overlap (538 mRNAs) between upregulated transcripts among MM cell lines (Fig. 7c). Notably, functional annotation clustering analysis of upregulated transcripts in SKMM1 and H929 MM cells expressing FAM46CWT (FDR < 1%), (using
**Fig. 8** FAM46C expression results in polyadenylation of selected mRNAs in MM cells. 

**a** Northern blot analysis of SSR4, CD320, and FTL transcripts from H929 (lanes 1–3), SKMM1 (lanes 4–6), Raji (lanes 7–9) and HL60 (lanes 10–12) cells transduced with GFP (lanes 1, 4, 7, 10), FAM46C<sup>WT</sup>-GFP (lanes 2, 5, 8, 11), and FAM46C<sup>mut</sup>-GFP (lanes 3, 6, 9, 12). Asterisks indicate cross-hybridization signals.

**b** The SSR4 transcript is extensively polyadenylated by FAM46C in MM cells. High-resolution northern blot analysis of SSR4 transcripts from SKMM1 (lanes 1–6) and H929 (lanes 7–12) transduced with GFP (lanes 1, 2, 7, 8), FAM46C<sup>WT</sup>-GFP (lanes 3, 4, 9, 10), and FAM46C<sup>mut</sup>-GFP (lanes 5, 6, 11, 12) after RNase H treatment (lanes 1, 3, 5, 7, 9, 11) to remove the poly(A) tail in presence of oligo(dT)<sub>25</sub>. Control reactions were carried out in the presence of oligo(dT) without RNase H (lanes 2, 4, 6, 8, 10, 12). Kinetics of polyadenylation of SSR4, NAPSA, FTL, and CD320 transcripts over the time course of FAM46C expression.

**c** High- and low-resolution northern blot analysis of SSR4, NAPSA, GAPDH, and RN7 transcripts from H929 (lanes 1–7) and SKMM1 cells (lanes 8–14) transduced with FAM46C<sup>WT</sup>-GFP up to 24 h.

**d** High-resolution northern blot analysis of SSR4 and FTL transcripts from H929 (lanes 1–8) and SKMM1 cells (lanes 9–16) transduced with FAM46C<sup>WT</sup>-GFP (lanes 1–4, 9–12) and FAM46C<sup>mut</sup>-GFP (lanes 5–8, 13–16) up to 72 h.

**e** Northern blot analysis of NAPSA and CD320 transcripts from H929 (lanes 1–4) and SKMM1 cells (lanes 5–8) transduced with FAM46C<sup>WT</sup>-GFP up to 72 h.
DAVID) revealed significant enrichment towards transcripts encoding proteins that pass through ER/Golgi and are transported into lysosomes or exported outside the cell (Fig. 7d and Supplementary Fig. 7a, b).

In order to take a closer look at changes after wild-type FAM46C expression, we fractionated mRNA based on poly(A) tail length from H929 and SKMM1 MM cells expressing FAM46CWT-GFP, FAM46Cmut-GFP, and GFP for 72 h. RNA-seq libraries were prepared from two RNA fractions with the longest tails (fractions 5 and 6) and one fraction with the shortest poly(A) tail (fraction 1), followed by library preparation, sequencing, and initial analysis (analogous to the one described for RNA-seq).

In order to identify FAM46C substrates, we aimed to determine differences in the enrichment of mRNA species in cells transduced with wild-type FAM46C compared to cells transduced with the catalytically inactive mutant. To this end, we calculated "ratio fractions (5 and 6) relative to the short fraction 1 in FAM46C expression, we fractionated mRNA based on poly(A) tail (fraction 1), followed by library preparation, sequencing, and initial analysis (analogous to the one described for RNA-seq).

In order to identify FAM46C substrates, we aimed to determine differences in the enrichment of mRNA species in cells transduced with wild-type FAM46C compared to cells transduced with the catalytically inactive mutant. To this end, we calculated "polyadenylation ratios" by dividing the levels of individual mRNA in long poly(A) fractions (5 and 6) relative to the short fraction 1 in FAM46CWT cells using relative levels observed in FAM46Cmut cells. Plotting the histograms of polyadenylation ratios revealed a significant number of transcripts with elongated poly(A) tails, which did not fit the normal distribution (Fig. 7e, f and Supplementary Fig. 7c, d). This approach allowed us to identify 1744 mRNAs for SKMM1 and 659 mRNAs for H929 that shift towards longer poly(A) fractions in response to forced FAM46CWT expression; thus, they are presumably polyadenylated by this protein. The top-scoring transcripts were compared and as much as 60% of all H929 statistically significant hits were in common with SKMM1 hits, suggesting strong substrate specificity of FAM46C (Fig. 7g and Supplementary Data 2). Moreover, in the case of SKMM1 cell line the levels of basically all polyadenylated transcripts were increased confirming the role of FAM46C-mediated polyadenylation in enhancing mRNA stability. The correlation between the polyadenylation and mRNA levels was also visible for H929 MM cells. The effect was, however, less pronounced, most probably due to the interferon response observed in H929—but not in SKMM1—which is known to have a strong influence on transcriptome hemostasis.

Importantly, clustered GO-term analysis of transcripts polyadenylated by FAM46C using DAVID (https://david.ncifcrf.gov/) revealed that, as in the case of standard RNA-seq analysis, many FAM46C substrates encode proteins that pass through ER/Golgi and are transported into lysosomes or exported outside the cell (Fig. 7h)21, 22. Moreover, despite filtering for highly-expressed transcripts in an analysis that usually creates a bias towards longer mRNAs, we found that short mRNAs with short 3’-UTRs are more prone to being substrates of FAM46C (Fig. 7i). Sequence analysis of 3’-UTRs of FAM46C-polyadenylated mRNA revealed that, although not a single specific-sequence motif could be identified, there was a striking enrichment of highly statistically significant GpU- and CpU-rich motifs compared to 3’-UTRs of all expressed genes (Supplementary Fig. 7e). These motifs or secondary structures may be recognized by FAM46C directly or indirectly through other recruiting factors.

SSR4 (signal sequence receptor subunit delta), which is involved in the translocation of proteins across the endoplasmic reticulum membrane, was identified as a top hit FAM46C substrate in both MM cell lines. SSR4, as well as other
ER-targeted proteins, such as CD320, NAPSA, and FTL, were selected for further validation analysis based on the fact that ER stress is considered to be one of the most important processes in maintaining and targeting MM cells and augmentation of ER overload by proteasome inhibitors has emerged as an important therapeutic strategy for MM treatment.

Northern blot analysis confirmed that SSR4, CD320, and FTL transcripts migrate much slower, as their poly(A) tails are extended in H929 and SKMM1 MM cells expressing FAM46C \(^{\text{WT}}\)-GFP (Fig. 8a); this effect was not observed when GFP or FAM46C \(^{\text{mut}}\)-GFP were expressed. Moreover, poly(A) tails were not extended in other cell lines derived from hematological malignancies, such as in Raji and HL60 transduced with FAM46C. Furthermore, SSR4 mRNA is polyadenylated at the 3’ end as revealed by analysis using RNase H cleavage assay with oligo(dT)25 followed by northern blot (Fig. 8b); similar molecular phenotypes were determined for the translocon-associated protein SSR2 (Signal Sequence Receptor, Beta) (Supplementary Fig. 8a, b).

Additionally, poly(A) tail length dynamics of selected mRNAs during FAM46C expression were also evaluated in a time course. Cells were transduced with FAM46C \(^{\text{WT}}\)-GFP and FAM46C \(^{\text{mut}}\)-GFP, collected at indicated time points, and analyzed using northern blot. The first extended molecules were detected after 16 h of FAM46C \(^{\text{WT}}\)-GFP expression in both cell lines (Fig. 8c–e). In order to determine if the effect of SSR4 mRNA polyadenylation was specific to the FAM46C protein, we constructed lentiviral vectors allowing for expression of other RNA poly(A) polymerases, including GLD2-GFP, PAPOLA-GFP, and POLS-GFP, and transduced H929 and SKMM1 cells, and confirmed that this phenotype is caused by expression of FAM46C gene (Supplementary Fig. 8c). It was concluded that FAM46C polyadenylates a broad spectrum of substrates in MM cell lines, inducing strong enrichment in those encoding ER-targeted proteins.

FAM46C is a monomer and co-purifies with BCCIP\(^{\beta}\) and PABPC1. Cytoplasmic non-canonical poly(A) polymerases usually obtain specificity through associated RNA-binding proteins\(^{25, 24}\). In order to identify the proteins interacting with FAM46C, co-immunoprecipitation (Co-IP) experiments were performed using SKMM1 and H929 MM cells expressing FAM46C \(^{\text{WT}}\)-GFP. High-resolution mass spectrometry (MS) followed by label-free quantification (LFQ) revealed that BCCIP\(^{\beta}\) is the only protein enriched in both MM cell lines (Fig. 9a, Supplementary Data 3). Analyses performed in HEK293 cells also allowed for the identification of cytoplasmic poly(A)-binding protein PABPC1 as a potential FAM46C interactor (Fig. 9b, Supplementary Data 3). PABPC1 was an interesting target due to the poly(A) polymerase enzymatic activity of FAM46C. Interactions between FAM46C and BCCIP\(^{\beta}\), PABPC1 were confirmed by Co-IPs followed by western blot analyses of HEK293, SKMM1, and H929 MM stable cell lines expressing FAM46C \(^{\text{WT}}\)-GFP. Interaction of FAM46C with BCCIP\(^{\beta}\) was very stable under all experimental conditions tested; however, interaction with PABPC1 was identified only at physiological salt concentration (Fig. 9b). Despite the fact that HEK293 cells express both FAM46A and FAM46B proteins, none of these co-purified with FAM46C, suggesting that they do not form heteromers.

PABPC1 and BCCIP\(^{\beta}\) proteins were selected for functional evaluation as potential FAM46C interactors. First, it was determined whether these proteins had any impact on mRNA after tethering to an RL reporter construct. Similarly to previous experiments, cells were transiently transfected with pRL-5BoxB, pNHA-BCCIP\(^{\beta}\), and pNHA-PABPC1 plasmids and, after 24 h, RL mRNA and protein levels were analyzed. BCCIP\(^{\beta}\) did not affect mRNA; however, PABPC1 tethering led to increased RL mRNA and protein levels (Supplementary Fig. 9a–c). Next, we examined FAM46C poly(A) polymerase activity in cells with reduced BCCIP and PABPC1 levels after siRNA treatment. HEK293 cells were treated with siRNA against BCCIP or PABPC1 mRNAs and, after 3 days, reseeded and co-transfected with pRL-5BoxB, pNHA-FAM46C, and pNHA-FAM46C\(^{\text{mut}}\). BCCIP\(^{\beta}\) and PABPC1 protein levels after siRNA treatment were determined by western blot analysis (Supplementary Fig. 9d). After 24 h, steady state levels of reporter mRNA and protein were analyzed (Supplementary Fig. 9e). In all cases, FAM46C tethering led to increases in steady state RL mRNA and protein levels (Supplementary Fig. 9e). In conclusion, FAM46C does not form stable macromolecular assemblies and interacting proteins have no direct effect on its poly(A) polymerase activity. Thus, the mechanism through which FAM46C gains specificity towards particular mRNAs remains unknown.

Discussion

Template-independent elongation of the 3’-end of mRNA by nucleotidyltransferases, known as poly(A) polymerases, plays a crucial role in RNA processing that affects mRNA stability and translational efficiency. Here we provide strong in vitro and in culture evidence demonstrating that the FAM46C gene encodes a novel poly(A) polymerase that acts as a B-cell lineage growth suppressor. Reintroduction of FAM46C into MM cell lines with dysfunctional endogenous gene leads to cell death while its silencing in cells expressing WT proteins enhances proliferation rate. Furthermore, activated primary B lymphocytes isolated from FAM46C KO animals divide faster than the wild type, suggesting that FAM46C acts more generally in the B lymphocyte lineage. Hematologic phenotypes of FAM46C KO animals, in which red blood cell and platelet counts are increased further suggests that this enzyme controls proliferation of other blood cells or common progenitor. To summarize, we are strongly confident that MM cells must benefit from mutations in FAM46C, however we cannot exclude that FAM46C suppresses the growth of other cell types than ones derived from the B-cell lineage. Recent studies from the Xin Lab suggest that FAM46C can also be involved in regulation of proliferation of hepatocellular carcinoma cells\(^{25, 26}\).

Our data based on SKMM1 and H929 cell lines strongly indicate that FAM46C has a relatively broad spectrum of substrates with significant enrichment towards ER/Golgi-targeted proteins, which may explain the toxicity to MM cells since ER and Golgi apparatus homeostasis is essential for survival. Although we did not determine the specificity of the FAM46C protein for particular substrates, we found that the toxicity of FAM46C to MM cells is unique, since overexpression of other poly(A) polymerases, such as GLD2 and canonical PAP, was not harmful to MM cells. Similarly to the overexpression of the wild-type form of FAM46C in HL60, there was no significant effect on the survival of Raji and 293 T cell lines. These results suggest that the toxicity of FAM46C to MM cell lines is related to the expression of ER/Golgi-targeted proteins, which may result in ER overload and induction of an unfolded protein stress response in these cells\(^{27, 28}\). However, how exactly the numerous genetic differences between cancer cell lines makes them sensitive to FAM46C expression remains to be established.

Importantly, this work provides the first experimental evidence directly linking RNA 3’ end polyadenylation process with cell proliferation, notably in the B-cell lineage implying the possible role of FAM46C mutations as driver mutations in MM. Interestingly, a well-established regulator of polyadenylation in gametogenesis, cytoplasmic polyadenylation element binding (CPEB), is also implicated in senescence and was suggested to be a tumor
suppressor29–31. However, the enzyme responsible for CPEB-dependent polyadenylation in somatic cells remains unknown, as knock-out of the poly(A) polymerase GLD2, which cooperates with CPEB in Xenopus oocytes32, 33, has no visible phenotype in mice34.

PABPC1 and BCCIP proteins are novel FAM46C interactors that had not been reported in high throughput screens. The cytoplasmic-nuclear shuttling protein PABPC1 is a FAM46C interactor and is directly related to its physiological function, since it binds to nascent poly(A) tracts, affecting initiation of protein translation and both RNA processing and stability; however, the additional factors determining FAM46C substrate specificity require further investigation. Furthermore, FAM46C displays very weak distributive activity in vitro, while it seems to be more active in vivo, thus additional factors responsible for its optimal activity probably remain to be identified.

Our identification of FAM46 proteins as active poly(A) polymerases suggests that the human genome encodes at least 11 putative ncPAPs and PUPs (4 FAM46 family members and 7 other previously annotated genes; Table 1). FAM46C, as well as all other members in this family, may positively regulate gene expression in various human tissues and organs, since they are differentially expressed during development and several tissues express multiple FAM46 proteins at the same time35. Further research is necessary in order to determine the physiological roles of this novel family of poly(A) polymerases. Data obtained thus far strongly suggest that gene expression regulation by cytoplasmic polyadenylation may play a role in various other cellular processes. Our initial analysis of FAM46C KO animals suggests several hematologic abnormalities, like anemia, what is in agreement with previously reported data. However, how exactly FAM46C affects erythropoiesis remains to be established.

Recent studies have shown that loss-of-function mutation within FAM46A is associated with bone abnormalities in mice and in autosomal recessive retinitis pigmentosa in humans; however, its molecular function remains unexplained36, 12, 36, 37. FAM46A is expressed in the ameloblast nuclei of tooth germs and may be important for the formation of enamel in teeth38.

Table 1 List of human non-canonical PAPs and PUPs

| Protein  | Architecture | Substrate preference | Function                                      |
|----------|--------------|----------------------|-----------------------------------------------|
| hGld2 (PAPD4) |              | A                     | Cytoplasmic polyadenylation                    |
| U6 TUTase (Star-PAP; TUT1) |              | U or A                | U6 snTUTase, phosphatidylinositol dependent nuclear polyadenylation |
| MtPAP (PAPD1) |              | A                     | Mitochondrial polyadenylation                  |
| GLD4 (PAPD5, hTRF4-2) |              | A                     | Nuclear RNA surveillance, p53 mRNA polyadenylation, orthologue of yeast TRF4 |
| POLS (PAPD7, hTRF4) | Unknown      | Unknown, orthologue of yeast TRF4 |
| ZCCHC6 (TUT7) | U            | Cytoplasmic RNA decay |
| ZCCHC11 (TUT4) | U            | Cytoplasmic RNA decay |
| FAM46A | Unknown      | Mutations in mice cause skeletal dysplasia |
| FAM46B | Unknown      | Unknown, unknown |
| FAM46C | A            | Cytoplasmic polyadenylation, tumor suppressor |
| FAM46D | A            | Unknown, unknown |

Orange box, PAP catalytic domain; green box, PAP-associated domain; blue box, C2H2/RRM/CCHC RNA-binding domains; grey box, inactive PAP catalytic domain.
The FAM46A sequence polymorphisms SNP (rs11040) and VNTR, within its coding sequence in exon 2 (15bp VNTR repeats may vary from two to seven repeats), were also associated with a susceptibility to tuberculosis. FAM46A expression has been reported to increase over two-fold in macrophages after stimulation by heat-attenuated Mycobacterium tuberculosis (H37Rv). The VNTR polymorphism in the FAM46A gene has also been associated with susceptibility to large-joint osteoarthritis (hip and knee) and non-small cell lung cancer. FAM46D expression is strongly upregulated in transgenic mice (MALTT) that exhibit autism spectrum disorder-like behaviors (ASD). Since FAM46 proteins are active poly(A) polymerases, these data strongly suggest that additional phenotypes may be related to the regulation of posttranscriptional gene expression.

In conclusion, this study identified that the human FAM46C gene encodes a novel poly(A) polymerase and showed FAM46C potential to regulate gene expression posttranscriptionally resulting in an significant impact on several cell types in the hematopoietic lineages in mice. Thus, demonstrating importance of RNA poly(A) tail homeostasis for cancer cell survival which could be further pursued as a potential therapeutic target for MM treatment.

**Methods**

**Cell lines.** MM cell line SKMM1 was kindly provided by W. Michael Kuehl, M.D., Genetics Branch Head, Molecular Pathogenesis of Myeloma Section National Cancer Institute), H929 (ATCC; CRL-9068), MM1.S (ATCC; CRL-2974), MM1.R (ATCC; CRL-2975), RPMI8226 (ATCC; CCL-155), and Raji Burkitt’s lymphoma (ATCC; TIB-83) were cultured in RPMI 1640 (ATCC), supplemented with 10% FBS (Invitrogen) and penicillin/streptomycin (Sigma; P4868). HL60 cells were cultured in IMDM (Invitrogen) with 20% fetoal serum (FBS) and penicillin/streptomycin. HEK293T (ATCC; CRL-3216), Flp-In-293 Cell Line (Invitrogen; R78007), HeLa (ATCC; CCL-2), and U2OS (ATCC; HTB-9) were cultured in Dulbecco’s modified Eagle’s medium (DMEM; Invitrogen) supplemented with 10% FBS (Gibco) and penicillin/streptomycin at 37 °C in 5% CO2. All cell lines were tested for mycoplasma contaminations upon arrival and routinely during their culture.

Induction of stable cell lines was obtained by addition of tetracycline (Gibco) to a final concentration of 100 ng/ml. siRNAs and plasmid transfections were performed with Lipofectamine 2000 and Opti-MEM media (Invitrogen) according to the manufacturer’s instructions. For transfection inhibition, actinomycin D (Sigma; A9415) was added to a final concentration of 5 μg/ml at the times indicated.

**MM cell line genotyping.** The exon 2 of FAM46C gene was amplified using hFAM46C_ex2_1F and hFAM46C_ex2_1R primers, which were designed with Oligo software (Premier BioSoft, USA) and Primer-BLAST. PCR products were subsequently cloned and then thoroughly analyzed with Mutation Surveyor 4.0 software (SoftGenetics) against the presence of mutations and other differences compared to GenBank reference sequences NC_000013 and NC_000001.11. The Mutation Surveyor allowed for detection of minor allelic down to 3%, basing analysis of both strands. Sequencing results are summarized in Supplementary Table 1.

**General molecular biology techniques.** General molecular biology techniques were conducted as described in Molecular Cloning. Low-molecular weight RNA samples were separated on 4–6% acrylamide gels containing 7M urea in 0.5x TBE buffer and transferred to a Hybond N+ membrane by electrophoresis in 0.5x TBE buffer. High-molecular weight RNA samples were separated on 1.2% agarose gels containing formaldehyde in 1x TBE buffer and transferred to membranes by capillary elution using 8x SSC buffer. After transfer membranes, were soaked with 0.03% methylene blue in 0.3 M NaAc pH 5.3 for 10' at room temperature, scanned and then destined with water. RNA was immobilized on membranes by 254 nm UV light using a UVPL-CL1000 crosslinker. Radioactive probes were prepared with a DEC-Aprime II DNA Labeling Kit (Invitrogen). Northern blots were cut out in PerfectHy Plus Hybridization Buffer (Sigma), scanned with Fuji Typhoon FLA 7000 (GE Healthcare Life Sciences) and analyzed with Multi Gauge software Ver. 2.0 (FUJI FILM). All western blotting experiments were performed in accordance with the technical recommendations of the antibodies’ suppliers. All antibodies used in this study are listed in Supplementary Table 2. Unprocessed scans of selected major blots and gels are shown in the Supplementary Fig. 10.

**Cloning and construct preparation.** The primers used for cloning are listed in the Supplementary Data 4. All plasmid-encoding constructs for tethering assays and miRNA were cloned using classical restriction enzyme digestion and ligation approaches; all other constructs were prepared with SLIC19. The pGneo-NHA vector (Invitrogen) containing a BLE8 vector (Invitrogen) were kindly provided by Prof. Witold Filipowicz (FMI). pNHA-GENe constructs were prepared using classical restriction ligation methods. All pcDNA-GENE-TEV-eGFP constructs were prepared by PCR amplification from human cDNA using the appropriate primers listed in Supplementary Data 4. PCR products were subcloned into a modified version of the pcDNA5/FRT/TO vector (Invitrogen) using the SLIC method. Plasmid pcDNAFAM46CmuntGFp was obtained by site-directed mutagenesis using the pcDNANAFAM46CmuntE and pcDNA-NAFAM46CmuntR primers with the pcDNAFAM46CmuntGFp plasmid served as a template; final constructs were verified by sequencing. The HA-tag was inserted into PK-K-NoTag made in PCR reaction with primers N_Met-HA_for and N_Met-HA_rev using Acc65I and BamHI cut PKK-NoTag and then assembled with SLIC approach giving pKK-HA. pKK-NoTag is a modified version of pcDNA5FRTTO (All additional information will be described in forthcoming publication). pKK-HA-FAM46C construct was cloned using classical restriction ligation protocol into pkHKKA BstXI and Nhel restriction sites. To obtain pCI-NHAFAM46DWT and pCI-NHAFAM46Dmrt PCR product with primers NHAFAM46D and NHAFAM46DR and pKK-FAM46DWT or pKK-FAM46Dmrt plasmids, respectively and cloned into EcoRI and NotI sites in pCI-HA vector. All SFFV (pLVX) lentiviral constructs were prepared in modified HIVSFFVRFP plasmids, which was kindly provided by Dr Verhoeyen Els (ENS de Lyon). All GENE-TEV-eGFP inserts were obtained from a PCR reaction with the appropriate plasmid pcDNA-GENE-TEV-eGFP and universal SFFV_F and SFFV_R primers. PCR products were cloned into the plasmid HIVSFFVRFp and digested by BamHI and Xhol using the SLIC protocol. To make HIVSFFV-NHA-FAM46C construct, PCR product with primers 15_125F and 15_125_R and 15_131F and 15_131_R cloned into HIVSFFV vector using SLIC approach. To obtain HIVSFFVFLAG-FAM46C construct, PCR product with primers 15_125F1 and HivN_R using pcDNAFAM46CGFP plasmid as template, then reamplified with HivNFpF and HivNFpR primers was cloned into HIVSFFV vector using SLIC. To make HIVSFFV-RL and HIVSFFV-FL constructs, PCR products were cut using SLIC approach giving pKK-HA. pKK-NoTag is a modified version of pcDNA5FRTTO vector cut by BamHI and XhoI and then assembled with SLIC approach giving pKK-HA. pKK-NoTag using Acc65I and BamHI cut pKK-NoTag and then assembled with SLIC approach giving pKK-HA. PKK-NoTag is a modified version of pcDNA5FRTTO (All additional information will be described in forthcoming publication).

**Tethering assays.** Tethering assays were performed as previously described. In brief, a day before transfection, 0.75 ml of HEK293 cells were seeded on 6-well plates to achieve about 70–80% confluence on the day of transfection. Next, cells were co-transfected with 100 ng of constructs expressing reporter Renilla luciferase (RL-SBox8 or RL), 100 ng of control firefly luciferase (FL, pGL3 plasmid) and 2 μg of plasmid encoding tethered NHA-protein using 5 μl of Lipofectamine 2000 and OPTI-MEM media (Invitrogen) according to manufacturer’s instructions. All transfections were repeated at least three times. Twenty-four hours after transfection, cells were collected, lysed and protein levels were analyzed by western blot. If transfected cells were used solely for the analyses of RNA they were co-transfected only with 0.1 μg pRL-SBox plasmid carrying Renilla luciferase (RL, pGL3 plasmid) and 2 μg of plasmid encoding tethered NHA-protein using 5 μl of Lipofectamine 2000 and OPTI-MEM media (Invitrogen) according to manufacturer’s instructions. All transfections were repeated at least three times. Twenty-four hours after transfection, cells were collected, lysed and protein levels were analyzed by western blot. For special purposes, such as higher yield of RNA, cell cultures and transfections were scaled-up.

**Tethering assay in MM cells.** SKMM1 and H929 cells were co-transduced with lentiviral particles carrying FL and RL or RLSBox and FL at MOI 1. Expression of reporter genes in obtained stable cell lines was verified by western blot then cells were transduced with lentiviruses allowing expression of wild-type and mutated NHA-tagged FAM46C protein at MOI 1. Twenty-four hours after transfection, cells were collected and protein levels was determined by western blot.
Poly(A) fractionation was performed as previously described with the following RNA isolation and poly(A) fractionation procedures. For silencing of BCCIP and PABPC1, three different siRNAs were tested; they are listed in Supplementary Table 3. For further details, see the manufacturer’s instructions.

**RNA isolation and poly(A) fractionation.** RNA isolation was performed with TRIzol reagent (Invitrogen, 15596) according to the manufacturer’s instructions. Poly(A) fractionation was performed as previously described with the following modifications: 1. In brief, 80 µg of total RNA was mixed with 400 µl GTC buffer (4 M guanidine thiocyanate, 25 mM sodium citrate, pH 7.1, 1% β-mercaptoethanol) and 750 pmol of 5′-biotinylated TEG-oligo(dT)25 (Future Synthesis). 850 µl of “dilution buffer” (6 × SSC, 10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.25% SDS, 1% β-mercaptoethanol) was added, the mixture was incubated at 70 °C for 5 min, and then it was centrifuged at 12,000 rpm for 10 min at RT. The resultant supernatant was mixed with 100 µl of M280 beads (Invitrogen; 60210) washed 3 times with 0.5 × SSC buffer. After binding, samples were washed 3 times with 0.5 × SSC buffer for 10 min. RNA fractions were eluted with decreasing concentrations of SSC buffer (from 0.2 × SSC to 0.025 × SSC) and the final elution was performed with water. RNA was precipitated and subjected to further analysis.

**RNA-seq libraries.** 1 µg of total RNA was mixed with 100 ng of oligo(dT)25 (Invitrogen) for 1 h at 95 °C, and then annealed to a second 10 µl of 5X RX buffer (100 mM, 0.5 M KCl, 50 mM MgCl2, 50 mM DTT, and 25% sucrose) was added to 5 U of RNaseH (NEB). Reactions were carried out for 45 min at 37 °C and RNA was recovered by extraction with phenol/chloroform, precipitated, and analyzed using northern blots.

**Mass spectrometry analysis.** MS analysis was performed essentially as described. Briefly, precipitated proteins were dissolved in 100 µl of 100 mM ammonium bicarbonate buffer, reduced in 100 mM DTT for 30 min at 55 °C, alkylated in 55 mM iodoacetamide for 40 min at RT in the dark, and digested overnight with 10 ng/ml trypsin (Promega) at 37 °C. Finally, trypsin digestion was completed by adding an additional 0.1% formic acid in water to 35% solution (acetonitrile and 0.1% formic acid). The measurement of each sample was preceded by three washing runs to avoid cross-contamination; the final MS washing run was searched for the presence of cross-contamination between samples. If the protein of interest was identified in the washing run and in the next measured sample at the same or similar intensity, then the sample was regarded as contaminated and thus suggested levels of infection (MOI) was calculated by dividing the vector titer for the number of cells transduced. Efficiency of transduction was determined by flow cytometry for GFP-expressing cells using an Accuri C6 benchtop cytometer (BD Biosciences). Additional staining with 2.5 µg/ml propidium iodide (Sigma-Aldrich) in PBS was performed to distinguish dead cells. If needed, to select GOI-expressing cells, staining for GFP-positive cells using the Aria III cell sorter (BD Biosciences) was performed.

**RNA-seq data analysis.** RiBo-depleted total RNA isolated from mutant and wild type H929 cell line quadruplicates were used to prepare strand-specific libraries (cUTP RNA); RiBo-depleted total RNA isolated from mutant and wild type SKMM1 cell line triplicates were also used to prepare non-strand-specific TRUseq RNA libraries. These libraries were subsequently sequenced using an Illumina HiSeq sequencing platform to the average number of 18.5 million reads per sample in 100-nt pair-end mode. Poly(A) tail length fractionation libraries were sequenced to an average depth of in 75-nt pair-end fashion.

**Mass spectrometry analysis.** MS analysis was performed essentially as described. Briefly, precipitated proteins were dissolved in 100 µl of 100 mM ammonium bicarbonate buffer, reduced in 100 mM DTT for 30 min at 55 °C, alkylated in 55 mM iodoacetamide for 40 min at RT in the dark, and digested overnight with 10 ng/ml trypsin (Promega) at 37 °C. Finally, trypsin digestion was completed by adding a final concentration of 0.1%. MS analysis was performed by LC-MS in the Laboratory of Mass Spectrometry (IBB PAS, Warsaw) using a nanoAcquity UPLC system (Waters) coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). Peptides were separated by a 180 min linear gradient of 95% solution A (0.1% formic acid in water) to 35% solution A (acetonitrile and 0.1% formic acid). The measurement of each sample was preceded by three washing runs to avoid cross-contamination; the final MS washing run was searched for the presence of cross-contamination between samples. If the protein of interest was identified in the washing run and in the next measured sample at the same or similar intensity, then the sample was regarded as contaminated and thus suggested levels of infection (MOI) was calculated by dividing the vector titer for the number of cells transduced. Efficiency of transduction was determined by flow cytometry for GFP-expressing cells using an Accuri C6 benchtop cytometer (BD Biosciences). Additional staining with 2.5 µg/ml propidium iodide (Sigma-Aldrich) in PBS was performed to distinguish dead cells. If needed, to select GOI-expressing cells, staining for GFP-positive cells using the Aria III cell sorter (BD Biosciences) was performed.
Co-immunoprecipitation. Cell collected from two 145 mm plates at a confluency of 80–90% were freshly frozen, thawed on ice, and incubated for 30 min at 4 °C with gentle rotation in 3 ml lysis buffer (20 mM Tris, 150 mM NaCl, 0.5% Triton X-100, 1 mM DTT, 1 mM PMSF, 0.02 μM pepstatin A, 0.02 μg/ml chymostatin, 0.006 mM leupeptin, and 20 μM benzamidine hydrochloride supplemented with proteases and phosphatase inhibitors; Invitrogen). Next, lysates were sonicated for 30 min in Bioruptor Plus (Diagenode) (30s pulses with 30s intervals), after which they were cleared by centrifugation at 13,000 rpm and 4 °C for 15 min. Immunoprecipitations were performed using house-in made affinity slurry CNBr-activated SepFast MAG 4HF [Biotoolomics] coupled with house-in purified anti-GFP antibodies. After 2h of incubation, beads were washed six times with LB buffer (100 mM Tris-HEC pH 7.4, 75 mM NaCl, and 1 mM DTT). Cleavage with TEV protease was conducted for 2h at RT. Then, proteins were precipitated with PRM reagents (0.05 mM pyrogallol red, 0.16 mM sodium hydroxide, 1 mM sodium oxalate, 0.5 mM succinic acid, and pH 2.5 (Siga-Aldrich)) and then analyzed by MS.

Polyadenylation assay. 150 ng of purified recombinant protein was mixed with 32P-labeled RNA substrate (A20), in the presence of 1 mM ATP (NEB), 2 μM dNtase (Fermentas), 0.5 mM MgCl2, or MnCl2 (both) in PAP buffer (25 mM Tris-HCl pH 7.0, 50 mM KCl, 0.02 mM EDTA, 0.2 mM DTT, 100 μg/ml BSA [Invitrogen], and 10% glycerol). Reactions were carried out at 37 °C and stopped at indicated time points by addition of equal volume of RNA loading dye (98% deionized formamide, 25 mM EDTA pH 8.0, 0.025% (w/v) xylene cyanol, and 0.0025% (w/v) bromophenol blue). Reaction products were separated in 8 M urea/15% PAGE in 0.5× TBE.

RNA-protein UV-crosslinking and immunoprecipitation. FAM46C WT-GFP crosslinking to RNA followed by immunoprecipitation was performed as described35. Briefly, five 145 mm plates of stable HEK293 Flp-In TReX cell line expressing FAM46C WT-GFP at 90% confluence were cross-linked with 254 nm UV light (2500 μJ/cm2; UVP Crosslinker). Protein cross-linked to RNA were then immunoprecipitated with anti-GFP resin, followed by RNA labeling with 32P, SDS-PAGE complexes separation, transferred to the nitrocellulose membrane, staining with Pronase S, and autoradiographed.

Mice generation. Basing on UCSC Mice GRCn38/mm10 Assembly, sgRNA and SpCas9_rev primers and subsequently transcribed with T7 RNA polymerase. In silico predicted target sequences were selected and designed with Laemmli sample buffer, separated in SDS-PAGE gels and analyzed using GraphPad Prism software.

RNA synthesis. Streptococcus pyogenes Cas9 cDS was amplified from PX58 (pSpcas9BB-2A-GFP, Addgene Plasmid #48138) using T7_Spcas9_for and SpCas9_rev primers and subsequently transcribed with T7 RNA polymerase.

Collection of zygotcs. Ethical approval for the procedures on animals was obtained from I Local Ethical Commission for Experiments on Animals in Warsaw (decisions no. 527/2013 and 176/2016). Embryos used in all experiments were isolated from F1(C57BL/6×CBA) mouse females, which were induced to superovulation by injection of 10 IU of PMSG (Pregnant Mare Serum Gonadotropin; Sigma) and 10 IU of hCG (Human Chorionic Gonadotropin; Chorulon, Intervet, Netherlands) 48 h later. Females were mated with males of the same strain immediately after hCG injection. Zygotes were collected from mice ovaries 10 h post hCG injection with a syringe attached to a 26G needle. Zygotes were then cultured in drops of M16 medium and stained with Ponceau S, and autoradiographed.

Bone marrow analyses. To determine bone marrow cellularity, the cells were counted on a Bürker hemocytometer and dead cells were excluded using trypsin blue dye. To evaluate percentage of plasma cells bone marrow cells were blocked in 2% BSA for 15 min. Next, the cells were labeled with anti-CD45.2 Horizon V500 (BD 562129) and anti-CD138 PE (BD 553714) monoclonal antibodies for 30 min at room temperature by 1:25 dilution, then washed with PBS and analyzed using LSR Fortessa and BD FACSdiva software version 8.0.1. The plasma cells were identified using CD138 from CD45.2-positive population. Discriminating positive vs. negative signals was ensured with fluorescence minus one (FMO) control.

Splenocytes fractionation. Cell fractionation was performed using a REAP protocol36. Briefly, cells isolated from spleen of WT and FLAG-tagged knock-in mice were washed in ice-cold PBS then resuspended in ice-cold 0.1% NP40-PBS (Sigma) and centrifuged at 1000 × g for 10 min. Supernatant was collected and centrifuged at 10,000 × g for 10 min at 4 °C. pellet was then washed with PBS and incubated in 0.1% NP40-0.5% PBS containing 500 U of viscosac (AA & Biotechology) incubated at 37 °C for 30 min and designated as “nuclear fraction”. Collected fractions were supplemented with Laemml sample buffer, separated in SDS-PAGE gels and analyzed with western blot approach.

Data availability. The RNA-seq data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE83772) under accession number GSE83772. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD006790. All scripts used in bioinformatics analyses are available for researchers upon request. All other remaining data are available within the Article and Supplementary Files or available from the authors upon request.

Received: 13 June 2016 Accepted: 11 July 2017
Published online: 20 September 2017

References
1. Barbieri, M. et al. Compendium of FAM46C gene mutations in plasma cell dyscrasias. Br. J. Haematol. 174, 642–645 (2015).
2. Raab, M. S., Podar, K., Breitkreutz, I., Richardson, P. G. & Anderson, K. C. Multiple myeloma. Lancet 374, 324–339 (2009).
3. Boyd, K. D. et al. Mapping of chromosome 1p deletions in myeloma identifies FAM46C at 1p12 and CDKN2C at 1p32.3 as being genes in regions associated with adverse survival. Clin. Cancer Res. 17, 7776–7784 (2011).
4. Prideaux, S. M., Conway O’Brien, E. & Chevassut, T. J. The genetic architecture of multiple myeloma. Adv. Hematol. 2014, 864058 (2014).
5. Chapman, M. A. et al. Initial genome sequencing and analysis of multiple myeloma. Nature 471, 467–472 (2011).
6. Bianchi, G. & Gloebrich, I. M. Biological and clinical implications of clonal heterogeneity and clonal evolution in multiple myeloma. Curr. Cancer Ther. Rev. 10, 70–79 (2014).
7. Manier, S. et al. Genomic complexity of multiple myeloma and its clinical implications. Nat. Rev. Clin. Oncol. 14, 100–113 (2017).
8. Lohr, J. G. et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell. 25, 91–101 (2014).
9. Lawrence, M. S. et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505, 495–501 (2014).
10. Forbes, S. A. et al. COSMIC: exploring the world’s knowledge of somatic mutations in human cancer. Nucleic Acids Res. 43, D805–D811 (2015).
11. Bolli, N. et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat. Commun. 5, 2997 (2014).
12. Barragan, I. et al. Genetic analysis of FM46A in Spanish families with autosomal recessive retinitis pigmentosa: characterisation of novel VNTRs. Ann. Hum. Genet. 72, 26–34 (2008).
13. Schoggins, J. W. et al. A diverse range of gene products are effectors of the type I interferon antiviral response. Nature 472, 481–485 (2011).
14. Tian M. The molecular cloning and characterization of Fam46c RNA stability factor. Harvard University; PhD dissertation. (2010).
15. Kuchta, K., Knizewski, L., Wyrwicz, L. S., Rychlewski, L. & Ginalski, K. Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. Nucleic Acids Res. 37, 7701–7714 (2009).
16. Ginalska, K., von Grothhuss, L., Grishin, N. V. & Rychlewski, L. Detecting distant homology with Meta-BASIC. Nucleic Acids Res. 32, W576–W581 (2004).
17. Aravind, L. & Koonin, E. V. DNA polymerase beta-like nucleotidyltransferase superfamliy: identification of three new families, classification and evolutionary history. Nucleic Acids Res. 27, 1609–1618 (1999).
18. Schmidt, M. J. & Norbury, C. J. Polyadenylation and beyond: emerging roles for nucleotidyl transferase. Nucleic Acids Res. 37, 1–14 (2009).
19. Chekulaeva, M. et al. miRNA repression involves GW182-mediated translational repression. Nature 494, 313–316 (2013).
20. Frecha, C. et al. Measles virus glycoprotein-pseudotyped lentiviral vector-based vectorization of dendritic cells elicits potent protective immunity. J. Virol. 87, 1697–1706 (2013).
21. Zhang, Q. Y., Yue, X. Q., Jiang, Y. P., Han, T. & Xin, H. L. FAM46C is critical for the cell invasion and migration in human glioma cells. PLoS ONE 10, e0122651 (2015).
22. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative discovery. Mol. Syst. Biol. 4, 2008.
23. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identity. Nat. Biotechnol. 29, 1277–1282 (2011).
24. Simossa, V. A. & Heringa, J. PRALINE: a multiple sequence alignment tool that integrates homology-extended and secondary structure information. Nucleic Acids Res. 33, W289–W294 (2005).
25. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).
26. Brun, D. M. & Rychlewski, L. J. D. GBP regulation of human cellular senescence. Energy metabolism, and pentas RNA transcription. Genes Dev. 22, 3499–3506 (2008).
27. Nakanishi, T. et al. Identification and characterization of C6orf37, a novel candidate human retinal gene is expressed in developing tooth buds. Arch. Oral. Biol. 54, 1002–1007 (2009).
28. Etokebe, G. E. et al. Family-with-sequence-similarity-46, member A (Fam46a) gene is expressed in developing tooth buds. Arch. Oral. Biol. 54, 1002–1007 (2009).
29. Etokebe, G. E., Bulat-Kardum, L., Munthe, L. A., Balen, S. & Dembic, Z. Association of variable number of tandem repeats in the coding region of the FM46A gene, FM46A rs131040 SNP and BAG6 rs3117582 SNP with susceptibility to tuberculosis. PLoS ONE 9, e91385 (2014).
30. Etokebe, G. E. et al. Susceptibility to large-joint osteoarthritis (hip and knee) is associated with BAG6 rs3117582 SNP and the VNTR polymorphism in the second exon of the FM46A gene on chromosome 6. J. Orthop. Res. 33, 56–62 (2015).
31. Fernandez-Miranda, G. & Mendez, R. The CPEB-family of proteins, translational control in senescence and cancer. Oncogene 641–646 (2012).
32. Kim, J. H. & Richter, J. D. Opposing polymerase-deadenylase activities regulate energy supply factors in mammalian cells. Nucleic Acids Res. 35, e132 (2007).
33. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).
34. Zhu, W. et al. WSEQC: quality control of RNA-seq experiments. Bioinformatics 28, 2184–2185 (2012).
35. Li, H. et al. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
36. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).
37. Mahony, S. & Benos, P. V. STAMP: a web tool for exploring DNA-binding motif similarities. Nucleic Acids Res. 35, W253–W258 (2007).
38. Heringa, J. Two strategies for sequence comparison: profile-preprocessed and structure-oriented multiple alignment. Comput. Chem. 23, 341–364 (1999).
39. Simossa, V. A. & Heringa, J. PRALINE: a multiple sequence alignment tool that integrates homology-extended and secondary structure information. Nucleic Acids Res. 33, W289–W294 (2005).
40. Kalsi, S. et al. A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. Nucleic Acids Res. 45, 2068–2080 (2016).
41. Szczepinska, T. et al. DSS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts. Genome Res. 25, 1622–1633 (2015).
42. Fulton, B. P. & Whittingham, D. G. Activation of mammalian oocytes by intracellular injection of calcium. Nature 273, 149–151 (1978).
43. Suzuki, K., Bose, P., Leong-Quong, R. Y., Fujita, D. J. & Riabowol, K. REAP: a two minute cell fractionation method. BMC. Res. Notes 3, 294 (2010).
Author contributions
A.D.: Developed and directed the studies, S.M.: Performed the vast majority of the biochemical and cell lines experiments, J.C.: with help of S.M. transduced MM cells, J.C. performed flow cytometry analyses, T.M.K.: Performed all bioinformatics analyses and proliferation assays, O.G.: Analyzed mice phenotypes and participated in localization studies, D.C.: Performed all high resolution mass spectrometry analyses, V.L.: Made preliminary RNA-seq libraries, J.G.: Designed and prepared CRISPR reagents, genotyped MM cells and mice, and E.B. generated transgenic animals. A.D. and S.M.: Wrote the manuscript with minor contributions from all authors. D.N.: Supervised J.C.’s work and discussed results.

Additional information
Supplementary Information accompanies this paper at doi:10.1038/s41467-017-00578-5.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017