Minigene-based splicing analysis and ACMG/AMP-based tentative classification of 56 *ATM* variants.

Running title: Functional classification of *ATM* splice-site variants

Elena Bueno-Martínez,[1] Lara Sanoguera-Miralles,[1] Alberto Valenzuela-Palomo,[1] Ada Esteban-Sánchez,[2] Víctor Lorca,[2] Inés Llinares-Burguet,[1] Jamie Allen,[3] Alicia García-Álvarez,[1] Pedro Pérez-Segura,[2] Mercedes Durán,[4] Douglas F. Easton,[3] Peter Devilee,[5] Maaike P. G. Vreeswijk,[5] Miguel de la Hoya,[2] Eladio A. Velasco-Sampedro[1]

MdlH and EAV-S are joint senior authors.

Affiliations:

[1] Splicing and genetic susceptibility to cancer, Unidad de Excelencia Instituto de Biología y Genética Molecular. Consejo Superior de Investigaciones Científicas (CSIC-UVa), 47003 Valladolid, Spain.

[2] Molecular Oncology Laboratory CIBERONC, Hospital Clínico San Carlos, IdISSC (Instituto de Investigación Sanitaria del Hospital Clínico San Carlos), Madrid, Spain.

[3] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, CB1 8RN, Cambridge, UK.

[4] Cancer Genetics, Instituto de Biología y Genética Molecular. Valladolid, Spain.

[5] Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands.

Correspondence to:

Eladio A. Velasco-Sampedro
Grupo de Splicing y Cáncer
Instituto de Biología y Genética Molecular (IBGM)
Consejo Superior de Investigaciones Científicas (CSIC)-UVa
Sanz y Forés 3.
47003 Valladolid.
Spain
Phone: +34 983184829
e-mail: eavelsam@ibgm.uva.es

Data Availability Statement: After publication, all sequencing and fragment analysis data will be made available at Digital.CSIC (http://hdl.handle.net/10261/265669; https://doi.org/10.20350/digitalCSIC/14573).

Funding Statement: PD, MPGV, DFE, MdlH and EAV have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 634935. EAV lab is supported by grants from the Spanish Ministry of Science and Innovation, Plan Nacional de I+D+I 2013-2016, ISCIII (PI17/00227 and PI20/00225) co-funded by FEDER from Regional Development European Funds (European Union) and from the Consejería de Educación, Junta de Castilla y León, ref. CSI242P18 (actuación cofinanciada P.O. FEDER 2014-2020 de Castilla y León). MdlH lab is supported by a grant from the Spanish Ministry of Science and Innovation, Plan Nacional de I+D+I 2013-2016, ISCIII (PI20/00110) co-funded by FEDER from Regional Development European Funds (European Union). EB-M is a postdoctoral researcher funded by the University of Valladolid (POSTDOC-UVA05, 2022-2025). LS-M is supported by a predoctoral fellowship from the AECC-Scientific Foundation, Sede Provincial de Valladolid (2019–2023). AV-P and IL-B are supported by predoctoral fellowships from the Consejería de Educación, Junta de Castilla y León (2018–2022). AE-S is supported through the Operational Program for Youth Employment and Youth Employment Initiative (YEI), called by the Community of Madrid in 2020, and co-financed by the European Social Fund.

Conflicts of Interest: The authors declare no conflict of interest.

Ethics approval: Ethical approval for this study was obtained from the Ethics Committee of the Spanish National Research Council-CSIC (28/05/2018).

The manuscript comprises original unpublished and is not under consideration for publication elsewhere.

ABSTRACT

The ataxia telangiectasia-mutated (ATM) protein is a major coordinator of the DNA damage response pathway. *ATM* loss-of-function variants are associated with two-fold increased breast cancer risk. We aimed at identifying and classifying spliceogenic *ATM* variants detected in subjects of the large-scale sequencing project BRIDGES. A total of 381 variants at the intron-exon boundaries were identified, 128 of which were predicted spliceogenic. After further filtering, we ended up selecting 56 variants for splicing analysis. Four functional minigenes (mgATM) spanning exons 4-9, 11-17, 25-29 and 49-52 were constructed in the splicing plasmid pSAD. Selected variants were genetically engineered into the four constructs and assayed in MCF-7/HeLa cells. Forty-eight variants (85.7%) impaired splicing, 32 of which did not show any trace of the full-length (FL)-transcript. A total of 43 transcripts were identified where the most prevalent event was exon/multi-exon skipping. Twenty-seven transcripts were predicted to truncate the ATM protein. A tentative ACMG/AMP (American College of Medical Genetics and Genomics/Association for Molecular Pathology)-based classification scheme that integrates mgATM data allowed us to classify 29 *ATM* variants as pathogenic/likely pathogenic and 7 variants as likely benign. Interestingly, the likely pathogenic variant c.1898+2T>G generated 13% of the minigene FL-transcript due to the use of a non-canonical GG-5'-splice-site (0.014% of human donor sites). Circumstantial evidences in three *ATM* variants (leakiness uncovered by our mgATM analysis together with clinical data) provides some support for a dosage-sensitive expression model in which variants producing ≥30% of FL-transcripts would be predicted benign, while variants producing ≤13% of FL-transcripts might be pathogenic.

Keywords: Hereditary Breast Cancer; Susceptibility genes; ATM; VUS; Splicing; Aberrant splicing; Splicing assay; Minigenes; Variant Classification.

INTRODUCTION

The *ATM* gene [MIM#607585], located on chromosome 11q22-23, is composed of 62 coding exons and encodes a large Serine/Threonine kinase of 3056 amino acids [1,2]. This protein plays an essential role in cellular homeostasis, being responsible for global orchestration of the cellular response to double-strand breaks. Biallelic germline mutations in *ATM* result in the autosomal recessive A-T (Ataxia Telangiectasia) syndrome, characterized by neurodegeneration, progressive ataxia, immunodeficiency, ocular telangiectasia, regular respiratory infections, gonadal atrophy and infertility [3], as well as increased cancer susceptibility, mostly lymphoid cancer [4,5].

Two recent large-scale studies of breast cancer (BC) patients have estimated that at least eight genes are significantly associated with breast cancer susceptibility, including *ATM*[6–8]. *ATM* is widely tested on commercial gene panels; heterozygous protein-truncating variants have been associated with a BC risk of around 2-fold [6,7], and a pancreatic cancer risk of 6.5-fold [9]. Possible increased risks of melanoma, stomach, and prostate cancers have also been reported [10].

Splicing is an essential and highly regulated RNA processing mechanism that is carried out by the spliceosome, an ensemble of ribonucleoproteins and other splicing factors that identify the cis-acting sequences needed for exon recognition, which include, among others, the basic donor or 5' (5'ss) and the acceptor or 3' splice-sites (3'ss) [11]. Historically, the role of splicing disruptions has been underestimated in genetic diseases, because often only variants in the "canonical" +/- 1,2 positions have been considered as potentially disease-causing. However, most potential spliceogenic variants are classified as variants of uncertain significance (VUS) because splicing outcomes cannot be accurately predicted. In this regard,

RNA assays provide information that might become critical for accurate clinical classification. In principle, RNA analysis of variant carriers can be used to classify variants as likely spliceogenic but, unfortunately, these samples cannot always be collected, and analysis in patient RNA is hampered by the presence of the wild-type allele. Certainly, the latter can be somehow overcome by RNA-seq approaches that detect allele-specific expression and calculate percent splicing index [12]. Alternatively, minigene assays provide a valuable approach to perform functional analysis of variants [13,14].

The BRIDGES project (Breast Cancer Risk after Diagnostic Gene Sequencing; https://bridges-research.eu/) is an international initiative that has sequenced 34 known or suspected BC genes in more than 113,000 women. Previously, we had performed comprehensive studies of BRIDGES' splice-site variants in *RAD51C*, *RAD51D* and *PALB2* by the splicing reporter minigene technology [15–17]. Here, we have selected and functionally analyzed 56 potential spliceogenic variants in *ATM* identified in BRIDGES subjects, using four different *ATM* splicing reporter minigenes. Further, we have integrated minigene data into an ACMG/AMP-based classification scheme that allows us to propose a tentative classification of all 56 tested variants.

MATERIALS AND METHODS

Ethical statement

Ethical approval for this study was obtained from the Ethics Committee of the Spanish National Research Council-CSIC (28/05/2018).

Annotation

All splicing events and predicted protein products were described according to the Human Genome Variation Society (HGVS) guidelines, using the Ensembl reference transcript ID ENST00000278616.8 (Genbank NM_000051.4). For clarity, we also used abbreviated notations using any of the following symbols [18,19]: ▼ (incorporation of intronic sequences not present in the reference transcript), ∆ (deletion of exonic sequences present in the reference transcript), E (exon), p (alternative 3' splice-site, new acceptor site), q (alternative 5' splice-site, new donor site) and a number representing the exact number of nucleotides incorporated or skipped. For example, transcript ▼(E8q5) denotes the use of an alternative donor site 5 nucleotides downstream of exon 8, causing the incorporation of 5-nt into the mature mRNA.

Selection of candidate *ATM* variants

A total of 381 unique variants at the *ATM* exon/intron boundaries were identified in the BRIDGES consortium sequencing data (Dorling et al., 2021). *In silico* splicing predictions were performed in all 381 variants using MaxEntScan (MES) [20] **(Supplementary Table S1)**. We selected likely spliceogenic variants based on: (i) ≥20% decrease of MES scores [21,22], (ii) creation of putative *de novo* sites (MES cut-off ≥ 3.0), or (iii) changes at conserved positions (-3, -2, -1, exon 5'-3'-ends, +3, +4, +5, and +6) of the consensus splice-site, regardless of MES predictions [17]. The

latter included eight variants with scores above the -20% threshold (c.3994-3C>T, c.902G>A, c.3577G>C, c.3746+4A>C, c.4436+4A>G, c.3993+5G>T, c.4109+6T>G and c.7788+6T>G).

Based on these criteria, we ended up selecting 137 likely spliceogenic variants spread all over the gene **(Supplementary Table S1)**. Since cloning all 63 *ATM* exons into minigenes was not feasible, we focused our attention in four exon clusters (4 to 9, 11 to 22, 25 to 29 and 49 to 52) in which a substantial proportion of candidate variants (61%) occur. After discarding candidate variants located in exons 17 to 22 (the dedicated minigene did not perform well, see results), and filtering-out several candidate variants located at the same splice-site positions with similar MES impact (e.g. c.1898+3A>G and c.1898+3A>T), we ended up with a list of 56 variants to be tested in minigenes **(Table 1, Supplementary Table S1)**.

Minigene construction and site-directed mutagenesis

Given that RNA from BRIDGES carriers had not been collected, we envisioned a minigene-based strategy similar to that we adopted in other BC susceptibility genes [15–17]. Minigenes mgATM_ex4-9, mgATM_ex11-17, mgATM_ex17-22, mgATM_ex25-29 and mgATM_ex49-52 were designed to include *ATM* exons 4 to 9, 11 to 17, 17 to 22, 25 to 29 and 49 to 52, respectively, and 200 nucleotides of flanking intronic sequences upstream and downstream from each exon **(Supplementary Figure S1)**. Subsequently, each insert was synthesized (Genewiz, South Plainfield, Waltham, MA, USA) and subcloned into the splicing plasmid pSAD (Patent P201231427-CSIC) **(Figure 1; Supplementary Figure S1, Supplementary Methods)** [23,24]. mgATM_ex4-9 minigene was obtained by inserting exon 9 into mgATM_ex4_8 minigene using HindIII/SalI restrictions

enzymes. The final minigenes were confirmed by sequencing (Macrogen, Madrid, Spain) and functionally checked (i.e. expressing the expected transcripts) in MCF-7 cells. All DNA variants were introduced into the wild type minigenes by site-directed mutagenesis using the QuikChange Lightning kit (Agilent, Santa Clara, CA) **(Supplementary Table S2)**. All mutant constructs were confirmed by sequencing (Macrogen).

Transfection

Approximately $2x10^5$ MCF-7 cells (human breast adenocarcinoma cell line) were grown to 90% confluence in 4-well plates (Nunc, Roskilde, Denmark) in 0.5 mL of medium (MEME, 10% Fetal Bovine Serum, 2 mM glutamine, 1% non-essential amino acids and 1% Penicillin/Streptomycin solution). The reproducibility of the minigene outcomes was tested in MDA-MB-231 (triple-negative breast cancer cell line) cells that were transfected with the wild type and mutant minigenes with variants c.901+2T>C, c.2377-2A>G, c.3746+5G>A and c.7629+2T>G. Cells were transiently transfected with 1 μg of each minigene and 2 μL of Lipofectamine LTX (Life Technologies, Carlsbad, CA). Nonsense mediated decay (NMD) was inhibited by incubating cells with cycloheximide 300 μg/mL (Sigma-Aldrich, St. Louis, MO) for 4 hours. RNA was purified using the Genematrix Universal RNA Purification Kit (EURx, Gdansk, Poland) with on-column DNAse I digestion following the manufacturer's instructions.

RT-PCR and cDNA amplification

The specific minigene-exon V2 primer RTPSPL3-RV (5'-TGAGGAGTGAATTGGTCGAA-3') was used to carry out a reverse transcription with 400 ng of RNA using the RevertAid First-Strand cDNA Synthesis Kit (Life

Technologies). Two μl of the resultant cDNA were used for amplification of the regions of interest using Platinum Taq DNA polymerase (Life Technologies). For all variants, the amplification was performed using the primers SD6-PSPL3_RTFW (5'-TCACCTGGACAACCTCAAAG-3') and RTpSAD-RV (Patent P201231427, CSIC). Samples were denatured at 94ºC for 2 min, followed by 35 cycles consisting of 94ºC for 30 sec, 59ºC for 30 sec and 72ºC (1 min/kb), and a final extension step at 72ºC for 5 min. RT-PCR products were sequenced by Macrogen. The expected minigene full-length (mgFL) transcripts are the following: mgATM_ex4-9 ($mgFL^{4-9}$: 1231 nt); mgATM_ex11-17 ($mgFL^{11-17}$: 1212 nt); mgATM_ex17-22 ($mgFL^{17-22}$: 999 nt); mgATM_ex25-29 ($mgFL^{25-29}$: 1041 nt); mgATM_ex49-52 ($mgFL^{49-52}$: 880 nt).

In order to assess the relative contribution of each transcript to the overall mgATM expression, semi-quantitative fluorescent RT-PCRs (26 amplification cycles) were performed in triplicate (in the case of c.1898+2T>G, experiments were replicated six times) using Platinum Taq DNA polymerase (Life Technologies) and the primers PSPL3_RTFW and RTpSAD-RV (both FAM-labeled) under standard conditions [24]. FAM-labeled products were run with LIZ-1200 Size Standard at the Macrogen facility (Seoul, Korea) and analyzed using the Peak Scanner software V1.0 (Life Technologies). Only peak heights ≥50 RFU (Relative Fluorescence Units) were considered. The protocol is summarized in the **Supplementary Figure S2**.

ACMG/AMP-based tentative classification of *ATM* genetic variants

We classified 56 *ATM* genetic variants according to a recently proposed ACMG/AMP point system, a Bayesian framework that outperforms the original classification guidelines, and allows for increased flexibility and accuracy in combining different ACMG/AMP criteria and strengths of evidence [25,26]. In this

framework, point-based variant classification categories are defined as follows: Pathogenic **(P)** ≥+10; Likely Pathogenic **(LP)** +6 to +9; Variant of Uncertain Significance **(VUS)** 0 to +5; Likely Benign **(LB)** -1 to -6; and Benign **(B)** ≤-7.

To assign ACMG/AMP scores [27] to individual variants, we based our analysis primarily on recently released (January 19, 2022) *ATM* specifications defined by the ClinGen Hereditary Breast, Ovarian and Pancreatic Cancer Variant Curation Expert Panel (clinicalgenome.org/affiliation/50039/). For some specific variants, we also used *ATM* specifications elaborated by the Spanish ATM Cancer Susceptibility Variant Interpretation Working Group (Feliubadaló et al., 2021). Finally, we have introduced some *ad-hoc* rules, in particular to incorporate mgATM complex read-outs (≥2 transcripts) into the classification scheme as PVS1_O/BP/_O codes of variable strength depending on the actual outcome. As results, we do not intend to provide an ACMG/AMP or ClinGen endorsed final classification of any *ATM* variant ready to be used in the clinical setting, but rather to highlight the complexity of incorporating complex minigene read-outs into an ACMG/AMP-based classification scheme. A comprehensive description of the classification scheme is provided in **Supplementary Methods**, **Supplementary Table S3.1-S3.3**, and **Supplementary Figure S3 A-C**. For comparative purposes only, we performed an alternative classification incorporating predictive splicing codes PVS1/PP3/BP4 rather than experimental splicing codes PVS1_O/BP7_O (see **Supplementary Table S3.4**).

RESULTS

A total of 381 unique variants at the *ATM* exon/intron boundaries were identified in the BRIDGES cohort. After filtering, we selected for minigene analysis up to 56 likely pathogenic variants clustering in a subset of *ATM* exons (exon 4 to 9, 11 to 17, 25 to 29 and 49 to 52) (see Materials and Methods for further details).

ATM **Minigenes**

We constructed five *ATM* minigenes (mgATM_ex4_9, mgATM_ex11_17, mgATM_17-22, mgATM_ex25_29 and mgATM_ex49_52) that we tested in MCF-7 cells. Four minigenes mimic the reference transcript NM_000051.3, producing as main outcomes the expected FL-transcripts: V1-*ATM* exons 4 to 9-V2, 1,231-nt (65.7%); V1-*ATM* exons 11 to 17-V2, 1,212 nt (84.1%); V1-*ATM* exons 25 to 29-V2, 1,041-nt (100%); V1-*ATM* exons 49 to 52-V2, 880-nt (75.8%). Likewise, the alternative isoforms ∆(E7) (34.2%; mgATM_ex4_9), ∆(E11) (15,9%; mgATM_ex11_17), ∆(E52) (24.2%; mgATM_ex49_52) and other uncharacterized transcripts were also detected (**Figure 1 B-D-F-H**, **Table 1**). On the other hand, we discarded for variant testing minigenes mgATM_17-22 and mgATM_ex49-54 (insertion of exons 53-54 into mgATM_ex49-52), because they did not produce clean splicing profiles (**Supplementary Methods**).

Splicing assays of ***ATM*** **variants**

Fifty-six variants were genetically engineered into the four minigenes: 14 in mgATM_ex4_9, 11 in mgATM_ex11-17, 20 in mgATM_ex25_29 and 11 in mgATM_ex49-52. For the purpose of the present analysis, splicing was considered “impaired” if the proportion of the corresponding mgFL-transcript was, at least, 10% lower than in WT construct. After RNA isolation, a semi-quantitative cDNA-

amplification was performed to analyze the impact of each variant. Forty-eight out of 56 (86%) variants disrupted splicing (**Table 1; Figures 2-5**). Twenty variants affected the ±1,2 positions and 28 targeted other splice-site positions: the polypyrimidine tract (3 variants), -3 (1 variant), +3 (4 variants), +4 (4 variants), +5 (5 variants), +6 (4 variants), as well as the first (1 variant), and the two last exonic nt (6 variants).

Up to 32 spliceogenic variants (underlined in **Table 1**) demonstrated strong impact on splicing (i.e. mgFL-transcripts not detected, or representing <5% of the overall signal), including one variant predicted missense [c.7787A>T (p.Glu2596Val)], and three variants predicted synonymous [c.3993G>A (p.Gln1331=), c.7515G>A (p.Lys2505=) and c.7788G>A (p.Glu2596=)] that did not produce any trace of the mgFL-transcript. The remaining 16 spliceogenic variants demonstrated weak to moderate splicing impacts, producing a non-negligible proportion of mgFL-transcripts (13 -71.4% of the overall signal). Curiously, four out of eight non-spliceogenic variants (c.2377-6T>A, c.2467-3A>G, c.2638+3A>G and 3994-3C>T) improved inclusion efficiency of the corresponding exons (i.e., the proportion of mgFL-transcripts were increased relative to their wild type counterpart). Unexpectedly, variant c.1898+2T>G produced mgFL-transcripts (up to 13%; average of 6 replicas) that might be explained by the use of the atypical GG-5'ss (0.01% of human exons) [29] created by this variant (Figure 3C). Finally, to check splicing reproducibility, one variant of each minigene (c.901+2T>C, c.2377-2A>G, c.3746+5G>A and c.7629+2T>G was tested in MDA-MB-231 cells, showing identical outcomes (**Supplementary Figure S4**).

Transcript analysis

Fluorescent-fragment analysis of minigene read-outs allowed us to characterize the mgFL-transcripts produced by the 4 WT minigenes, and up to 43 other transcripts **(Table 1 and Supplementary Table S4)**. The latter includes three alternative splicing isoforms, Δ(E7), Δ(E11) and Δ(E52), produced by the corresponding WT minigenes. Twenty-seven transcripts, including Δ(E7), introduced a premature termination codon (PTCs), while 15, including Δ(E11) and Δ(E52), kept the reading frame **(Table 1, Supplementary Table S4)**. One transcript of 970 nucleotides could not be characterized.

It is important to highlight the distinction between variant-induced transcripts (i.e. transcripts not produced by WT minigenes) and variant-induced splicing events (i.e. splicing events not detected in WT minigenes). For instance, the *ATM* variant c.332-1G>A (targeting exon 5 acceptor site) produces up to three variant-induced transcripts [Δ(E5)+[Δ(E5),Δ(E7)]+Δ(E5p1)], but only two variant-induced splicing events [Δ(E5) and Δ(E5p1)]. The Δ[(E5)(E7)] transcript combines variant-induced exon 5 skipping with exon 7 skipping, a splicing event already observed in WT minigenes (**see Figure 2B, Supplementary Tables S3.2 and S4**).

A significant proportion of variants (N=32) induced two or more splicing events, and/or demonstrated a partial effect on splicing, producing a non-negligible amount of mgFL-transcripts (leaky variants). These complex read-outs represented a challenge for variant interpretation (see latter). Exon (or multi-exon) skipping, observed in 48 variants, was the most frequent variant-induced event. Alternative site-usage was observed in 15 variants. Five leaky variants produced FL-transcripts that harbor missense (r.496G>A, r.902G>A, r.1898G>U, r.3557G>C) or synonymous (r.903U>G) changes (**Figures 2-4, Supplementary Table S3.2**).

ACMG/AMP-based tentative classification of 56 *ATM* variants

Once mgATM data was available, we decided to classify all 56 ATM variants according to ACMG/AMP variant classification guidelines [27], integrating mgATM data as PVS1_O/BP7_O evidence codes. We classified 29 variants as **P/LP** (six of them as pathogenic) and seven non-GT-AG intronic variants as **LB**. Up to 20 variants (36%) were classified as **VUS** (**Table 2 and Supplementary Table S3.1)**.

Overall, 37 of the 56 *ATM* variants here analyzed have been reported previously in ClinVar (last accessed 09/02/202, see **Supplementary Table S3.4**), but only 21 of them by multiple submitters with no conflicts (two-star review status). Focusing our analysis on the subgroup of 21 ClinVar no conflicting-variants, we conclude that our classification scheme (integrating mgATM data) does not reduce the number of **VUSs**, but rather reclassifies variants (7 variants, 33%) in both directions. Specifically, three variants reported in ClinVar as **VUSs** are upgraded to **P/LP**, while three variants reported as **P/LP** are downgraded to **VUSs**, and one variant reported as **LB/B** is upgraded to **VUS**. **Supplementary Table S5** shows a comparative analysis in this subgroup of variants.

To evaluate the contribution of mgATM data to variant classification, we compared our final classification (**Table 2, Supplementary Table S3.1**) with an alternative classification in which we simply replaced PVS1_O/BP7_O evidences with PVS1/PP3/BP4 predictive splicing evidences (**Supplementary Table S3.5**). **Supplementary Table S6** summarizes the comparative analysis. Experimental splicing data has an impact on the final classification of 16 *ATM* variants (29% of the tested variants). Overall, experimental splicing data has a positive effect on variant classification, reducing uncertainty (reducing the number of VUSs from 29 to 20).

DISCUSSION

NGS technology is an efficient screening approach to detect variants associated with cancer risk with high sensitivity, cost effectiveness and speed. Genetic testing has become available for larger groups of patients, allowing more variant carriers to be identified, better management of risk and, in some cases, better treatment [30]. However, NGS also presents some challenges. One of them is the high rate of VUSs, for which the association with cancer risk in unclear, and genetic counselling of carriers is difficult [31]. A recent large scale sequencing study reported that the prevalence of VUSs in 12 BC genes was 18.9% [7]. Classification of VUS may be improved through large-scale splicing or functional assays.

The *ATM* gene is one of the eight "core" genes that displayed a significant association with breast cancer in the two large-scale studies already mentioned [6–8]. *ATM* pathogenic variants are associated with a moderate risk of breast cancer (1.8-2.1) and an overall lifetime female breast cancer risk above 20%. Protein truncating variants in *ATM* accounts for 0.63% of all breast cancer cases [6,7]. We performed a comprehensive evaluation of potential spliceogenic variants to aid their interpretation. Here, we have analyzed *in silico* 381 *ATM* splice-site variants identified in BRIDGES patients and controls, through which 128 candidate variants were predicted to impair splicing.

NGS based RNA-seq provides high-quality qualitative and quantitative data for the characterization of splicing variants in hereditary cancer genes [32]. However, lacking RNA from carriers, we designed five minigenes covering 27 out of the 62 *ATM* coding exons in which most of the pre-selected variants were located, although, unfortunately, minigenes mgATM_17-22 and mgATM_49-54 did not show the correct transcript profiles and variants at exons 18 to 22, 53 and 54 were excluded from the

analysis. It is conceivable that smaller constructs, such as those with exons 19-20 or 21-22 (both with short introns), may be functional. Indeed, we introduced deletions of exons 17-18 and 17-20 into the wild type mgATM_ex17-22 that generated the corresponding full-length transcripts without any other isoform (**Supplementary Methods**).

Hybrid minigene technology has proven to be efficient for the description of the splicing outcomes of variants in the absence of patient RNA [33,34], as it is the case in the present study. There are many examples verifying the reproducibility of this strategy, including previous minigene studies of our group [19,35]. In this regard, the present study includes 10 *ATM* variants for which previous experimental RNA data in carriers has been published: c.496+5G>A [36]; c.901+3A>T [37]; c.902-1G>T [38]; c.1066-6T>G [39]; c.1898+2T>G [40]; c.1898+3_+4del [37]; c.3993+1G>A [37,41]; c.4110-9C>G [42], c.7630-2A>C [37,38,41,43], and c.7788G>A [44]. **Supplementary Table S7** shows a comparative analysis with mgATM data. Overall, concordance is high and do not affect PVS1_O/BP7_O code strengths (**Supplementary Table S3.2**). The only possible exception are three variants (c.496+5G>A, c.1066-6T>G, c.1898+2T>G) in which mgATM data has uncovered leaky effects not reported by previous RNA studies in carriers. In brief, the present study further supports the notion that hybrid minigenes are very good proxies for splicing assays in carriers.

SpliceAI is a neural network that predicts splicing from a pre-mRNA sequence [45]. Recent evaluations have identified SpliceAI as the best predictor of variants that impact splicing, here termed spliceogenic variants [46–49]. To further evaluate our analysis, we have compared mgATM data with SpliceAI predictions (note that SpliceAI was not used for the initial bioinformatics selection of *ATM* likely spliceogenic variants). The comparative analysis is shown **Supplementary Table**

S7. Taken together, the data further supports the robustness of the mgATM assay and, equally relevant, the accuracy of SpliceAI in predicting the actual outcome of spliceogenic variants. In relation with the latter, it is worth highlighting that:

(i) Four *ATM* variants targeting consensus positions of the splice-sites (c.2377-6T>A, c.2467-3A>G, c.2638+3A>G and c.3994-3C>T) do not disturb splicing (a remarkable finding correctly predicted by SpliceAI).

(ii) Eight *ATM* variants (c.1065+1G>T, c.1065+3A>G, c.3577G>C, c.3993G>A, c.3993+1G>A, c.3993+5G>T, c.7307+1G>A and c.7307+4A>G) contributed to the incorrect recognition of natural donor sites and to the use of cryptic 5'ss. SpliceAI predicted these variant impacts correctly, except for c.1065+3A>G.

The vast majority of human introns (~99%) are of the GT-AG type while the most frequent atypical 5'ss is a GC-donor [29]. It is known that GC donor splice-sites are related to alternative splicing events [50,51]. However, the mechanisms underlying the GC-5'ss recognition are not completely understood yet, so variants disrupting GC donor sites are particularly interesting. We have focused our attention on the normal splicing of GC-exons of breast cancer genes (e.g. *BRCA2* exon 17 or *PALB2* exon 12) as well as the anomalous GC usage induced by variants [17,19,34]. Here, we have analyzed two variants, c.7515G>A and c.7515+6T>C, affecting *ATM* exon 50 GC-5'ss. Given the intrinsic weakness of these 5'ss, it is expected that any sequence change may disrupt splicing. Indeed, both variants impaired exon recognition and provoked exon skipping, especially c.7515G>A (last exon nt), where the mgFL$^{49\text{-}52}$-transcript could not be detected. Taking into account the predicted effect on protein translation (p.Lys2505=), this should be *a priori* re-classified as a spliceogenic variant that produces two likely deleterious transcripts (discussed below). On the other hand, variant c.1898+2T>G induced the use of an extremely

rare GG 5'ss, which functions as a donor site in 0.01% of human introns [29], so that it partially restored the canonical splicing, generating 13% of the expected mgFL$^{11\text{-}17}$-transcript (V1-*ATM* exons 11 to 17-V2; Figure 3B-C).

Our classification schema integrates mgATM data (PVS1_O/BP7_O evidence) and provides an informative classification (P/LP or LB) for 36 variants, but 20 (36%) remained as VUS, including 14 intronic and 2 synonymous variants (the type of variants in which splicing data is expected to be a major contributor to classification). The relative high proportion of VUSs in our study is (partly) explained by the high proportion of mgATM read-outs for which inferring a pathogenic or benign evidence (i.e. deciding the appropriate PVS1_O or BP7_O code strength) is far from obvious, a complexity that we summarize as follows:

(i) For several variants, mgATM read-outs produced two or more altered transcripts with different coding potential and different contribution to the overall expression.
(ii) mgATM analysis identified 24 variants that produced altered transcripts, but also a significant proportion of full-length transcripts ("leaky variants").

To deal with these issues, we have been very conservative (as per ACMG/AMP recommendations), assuming that mgATM read-outs are non-informative (i.e., not adding points to the classification scheme) if both transcripts supporting a pathogenic call and transcripts supporting a benign call represent >10% of the overall expression (see **Supplementary Figure S3A**). Further, if a variant produces only transcripts supporting a pathogenic call (different strength), we have selected the most conservative option for overall PVS1_O code strength, even if representing only 10% of the overall expression (see **Supplementary Figure S3C**).

This conservative approach is reflected in the poor contribution of mgATM read-outs to the final point-based classification: adding only ≥(-1) and ≤(+2) points to

the final classification of 22 variants, including 16 variants for which no points were added (i.e. PVS1_O not applicable and BP7_O not applicable). In brief, many mgATM read-outs are non-informative, reflecting the complexity of integrating mgATM read-outs into an ACMG/AMP-based classification scheme.

In this regard, “leaky variants” are particularly challenging, as we do not know the precise relationship between *ATM* allele-specific expression levels of full-length transcripts and phenotype. It is conceivable to postulate a dosage-sensitive expression model in which some leaky variants producing full-length transcripts above a certain threshold are benign, leaky variants producing full-length transcripts below a certain threshold are pathogenic and leaky variants in between associate with an intermediate phenotype. Yet, as far as we know, there is no clinical and/or functional data in the scientific literature supporting (or addressing) this issue.

Said that, we noticed that evidence of leakiness uncovered by our mgATM analysis in three variants (c.1898+2T>G, c.496+5G>A, and c.1066-6T>G) together with clinical data available in the scientific literature [36,52] provides some support for an *ATM* dosage-sensitive expression model (see **Supplementary Figure S5** for further details). According to this tentative model, leaky variants producing ≥30% of full-length transcripts are predicted benign, leaky variants producing ≤13% are predicted pathogenic, and leaky variants in between might be associated with an intermediate phenotype. At present, this is just a tentative model based on circumstantial evidence. If confirmed by clinical evidences in a sufficient number of leaky variants and/or by functional studies, the dosage-sensitive expression model might be relevant to refine future iterations of the ACMG/AMP specifications for *ATM*.

In summary, here, we have carried out an exhaustive study of *ATM*, in which 56 preselected variants were tested using minigene assays (85.7% spliceogenic).

Once again, minigenes have proven to be a robust and useful tool to assess potential spliceogenic variants. These splicing assays provide key data for the interpretation of variants, so, despite the complexity of the *ATM* gene (63 exons), efforts should be made to test additional variants identified in the clinical setting (minigene approach, or NGS based RNA-seq analysis of patient RNA whenever possible). According to our ACMG/AMP-based tentative classification scheme, 29 variants end-up as pathogenic/likely pathogenic and 7 variants as likely benign. Finally, we provide circumstantial evidence supporting a dosage-sensitive model that might be relevant to classify leaky variants.

Acknowledgments: PD, MPGV, DFE, MdlH and EAV have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 634935. EAV lab is supported by grants from the Spanish Ministry of Science and Innovation, Plan Nacional de I+D+I 2013-2016, ISCIII (PI17/00227 and PI20/00225) co-funded by FEDER from Regional Development European Funds (European Union) and from the Consejería de Educación, Junta de Castilla y León, ref. CSI242P18 (actuación cofinanciada P.O. FEDER 2014-2020 de Castilla y León). MdlH lab is supported by a grant from the Spanish Ministry of Science and Innovation, Plan Nacional de I+D+I 2013-2016, ISCIII (PI20/00110) co-funded by FEDER from Regional Development European Funds (European Union). Programa Estratégico Instituto de Biología y Genética Molecular (IBGM), Escalera de Excelencia, Junta de Castilla y León (Ref. CLU-2019-02). EB-M is a postdoctoral researcher funded by the University of Valladolid (POSTDOC-UVA05, 2022-2025). AV-P and IL-B are supported by predoctoral fellowships from the Consejería de Educación, Junta de Castilla y León. LS-M is supported by a predoctoral fellowship from the AECC-Scientific Foundation, Sede Provincial de Valladolid (2019–2023). AE-S is supported through the Operational Program for Youth Employment and Youth Employment Initiative (YEI), called by the Community of Madrid in 2020, and co-financed by the European Social Fund.

Author Contributions statement

Conceptualization, M.d.l.H. and E.A.V.-S.; Data curation, E.B.-M., L.S.-M., A.V.-P., J.A., D.F.E.; Formal analysis, E.B.-M., L.S.-M., A.V.-P., M.d.l.H. and E.A.V.-S.; Funding acquisition, P.D., D.F.E., M.P.G.V., M.d.l.H. and E.A.V.-S.; Investigation, E.B.-M., L.S.-M., A.V.-P., A.E.-S., V.L., I.L—B., M.D., P.P.-S., M.P.G.V., M.d.l.H. and E.A.V.-S.; Methodology, A.V.-P., L.S.-M., E.B.-M., A.G.-A., I.L—B., M.D., M.d.l.H. and E.A.V.-S.; Supervision, E.A.V.-S.; Writing – original draft, E.B.-M., M.d.l.H. and E.A.V.-S.; Writing – review & editing, E.B.-M., P.D., D.F.E., M.P.G.V., M.d.l.H. and E.A.V.-S. All authors approved the final version of the manuscript.

List of supplementary material online

[Supplementary Figures S6-S7-S8 are embedded within the Supplementary Methods.]

Supplementary Table S3 (S3.1-S3.5). ACMG/AMP-based tentative classification according to a Bayesian point system.

Supplementary Table S4. RNA and protein HGVS descriptions according to the reference sequence NM_000051.3.

Supplementary Table S5. Comparative classification of 21 *ATM* variants.

Supplementary Table S6. Impact of mgATM data on the classification of 56 *ATM* variants.

Supplementary Table S7. Comparative analysis of SpliceAI predictions, mgATM read-outs, and experimental splicing data in carriers.

FIGURE LEGENDS

Figure 1. Structure and functional validation of the WT *ATM* minigenes used in this work. Schematic representation of the *ATM* minigenes with A) exons 4 to 9 (mgATM_ex4-9), B) 11 to 17 (mgATM_ex11-17), C) 25 to 29 (mgATM_ex25-29), D) 49 to 52 (mgATM_ex49-52). Exons are boxed; black arrows locate specific vector RT-PCR primers. Functional assays of the wild type minigene are shown below. RT-PCR products were run by fluorescent capillary electrophoresis, where the full-length and alternative transcripts are shown as blue peaks and the Liz1200 size standard is shown as orange/faint peaks. The x-axis indicates size in bp (electropherograms on the top) and the y-axis represents Relative Fluorescence Units (RFU).

Figure 2. Splicing Functional Assays of selected *ATM* variants in mgATM_ex4-9 minigene. A) Location of tested variants. B) Fluorescent fragment analysis of transcripts generated by the wild type and mutant minigenes. FAM-labelled products (blue peaks) were run with LIZ1200 (orange peaks) as size standard. FL, Full-length transcript. The x-axis indicates size in bp (electropherograms on the top) and the y-axis represents Relative Fluorescence Units (RFU).

Figure 3. Splicing Functional Assays of selected *ATM* variants in mgATM_ex11-17 minigene. A) Location of tested variants. B) Fluorescent fragment analysis of transcripts generated by the wild type and mutant minigenes. FAM-labelled products (blue peaks) were run with LIZ1200 (orange peaks) as size standard. FL, Full-length transcript. The x-axis indicates size in bp (electropherograms on the top) and the y-axis represents Relative Fluorescence Units (RFU). C) Consensus sequence of exon-intron boundaries of 101 non-canonical human GG-splice junctions [29] (top panel) vs the sequence of the atypical GG-splicing donor used in 13% of transcripts induced by variant c.1898+2T>G (medium panel) and the consensus sequence of canonical GT-donors (bottom panel). The size of each letter represents the nucleotide frequency at each position. Pictograms were obtained using WebLogo (https://weblogo.berkeley.edu/logo.cgi).

Figure 4. Splicing Functional Assays of selected *ATM* variants in mgATM_ex25-29 minigene. A) Location of tested variants. B) Fluorescent fragment analysis of transcripts generated by the wild type and mutant minigenes. FAM-labelled products (blue peaks) were run with LIZ1200 (orange peaks) as size standard. FL, Full-length transcript. The x-axis indicates size in bp (electropherograms on the top) and the y-axis represents Relative Fluorescence Units (RFU).

Figure 5. Splicing Functional Assays of selected *ATM* variants in mgATM_ex49-52 minigene. A) Location of tested variants. B) Fluorescent fragment analysis of transcripts generated by the wild type and mutant minigenes. FAM-labelled products (blue peaks) were run with LIZ1200 (orange peaks) as size standard. FL, Full-length

transcript. The x-axis indicates size in bp (electropherograms on the top) and the y-axis represents Relative Fluorescence Units (RFU).

REFERENCES

*** References 53 – 86 are cited only in the Supplementary Material.**

1 Savitsky K, Bar-Shira A, Gilad S, *et al.* A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* 1995; **268**: 1749-1753

2 Paull TT. Mechanisms of ATM Activation. *Annu Rev Biochem* 2015; **84**: 711-738

3 Choy KR, Watters DJ. Neurodegeneration in ataxia-telangiectasia: Multiple roles of ATM kinase in cellular homeostasis. *Dev Dyn* 2018; **247**: 33-46

4 Reiman A, Srinivasan V, Barone G, *et al.* Lymphoid tumours and breast cancer in ataxia telangiectasia; Substantial protective effect of residual ATM kinase activity against childhood tumours. *Br J Cancer* 2011; **105**: 586-591

5 Suarez F, Mahlaoui N, Canioni D, *et al.* Incidence, presentation, and prognosis of malignancies in ataxia-telangiectasia: A report from the french national registry of primary immune deficiencies. *J Clin Oncol* 2015; **33**: 202-208

6 Dorling L, Carvalho S, Allen J, *et al.* Breast Cancer Risk Genes — Association Analysis in More than 113,000 Women. *N Engl J Med* 2021; **384**: 428-439

7 Hu C, Hart SN, Gnanaolivu R, *et al.* A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med* 2021; **384**: 440-451

8 Narod SA. Which Genes for Hereditary Breast Cancer? *N Engl J Med* 2021; **384**: 471-473

9 Hsu FC, Roberts NJ, Childs E, *et al.* Risk of Pancreatic Cancer among Individuals with Pathogenic Variants in the ATM Gene. *JAMA Oncol* 2021; **7**: 1664-1668

10 Lesueur F, Easton DF, Renault AL, *et al.* First international workshop of the ATM and cancer risk group (4-5 December 2019). *Fam Cancer* June 2021

11 Wang G-SS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* 2007; **8**: 749-761

12 Schafer S, Miao K, Benson CC, *et al.* Alternative Splicing Signatures in RNA-seq Data: Percent Spliced in (PSI). *Curr Protoc Hum Genet* 2015; **87**: 11.16.1-11.16.14

13 Cooper T a. Use of minigene systems to dissect alternative splicing elements. *Methods* 2005; **37**: 331-340

14 Rhine CL, Neil C, Glidden DT, *et al.* Future directions for high-throughput splicing assays in precision medicine. *Hum Mutat* 2019; **40**: 1225-1234

15 Sanoguera-Miralles L, Valenzuela-Palomo A, Bueno-Martínez E, *et al.* Comprehensive Functional Characterization and Clinical Interpretation of 20 Splice-Site Variants of the RAD51C Gene. *Cancers (Basel)* 2020; **12**: 3771

16 Bueno-Martínez E, Sanoguera-Miralles L, Valenzuela-Palomo A, *et al.* RAD51D Aberrant Splicing in Breast Cancer: Identification of Splicing Regulatory Elements and Minigene-Based Evaluation of 53 DNA Variants. *Cancers (Basel)* 2021; **13**: 2845

17 Valenzuela-Palomo A, Bueno-Martínez E, Sanoguera-Miralles L, *et al.* Splicing predictions, minigene analyses, and ACMG-AMP clinical classification of 42 germline PALB2 splice-site variants. *J Pathol* 2022; **256**: 321-334

18 Lopez-Perolio I, Leman R, Behar R, *et al.* Alternative splicing and ACMG-AMP-2015-based classification of PALB2 genetic variants: an ENIGMA report. *J Med Genet* 2019; **56**: 453-460

19 Fraile-Bethencourt E, Valenzuela-Palomo A, Díez-Gómez B, *et al.* Mis-splicing in breast cancer: identification of pathogenic BRCA2 variants by systematic minigene assays. *J Pathol* 2019; **248**: 409-420

20 Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004; **11**: 377-394

21 Houdayer C, Caux-Moncoutier V, Krieger S, *et al.* Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Hum Mutat* 2012; **33**: 1228-1238

22 Moles-Fernández A, Duran-Lozano L, Montalban G, *et al.* Computational Tools for Splicing Defect Prediction in Breast/Ovarian Cancer Genes: How Efficient Are They at Predicting RNA Alterations? *Front Genet* 2018; **9**: 366

23 de Garibay GR, Acedo A, García-Casado Z, *et al.* Capillary Electrophoresis Analysis of Conventional Splicing Assays: IARC Analytical and Clinical Classification of 31 BRCA2 Genetic Variants. *Hum Mutat* 2014; **35**: 53-57

24 Acedo A, Hernández-Moro C, Curiel-García Á, *et al.* Functional classification of BRCA2 DNA variants by splicing assays in a large minigene with 9 exons. *Hum Mutat* 2015; **36**: 210-221

25 Tavtigian S V., Greenblatt MS, Harrison SM, *et al.* Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genet Med* 2018; **20**: 1054-1060

26 Tavtigian S V, Harrison SM, Boucher KM, *et al.* Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. *Hum Mutat* 2020; **41**: 1734-1737

27 Richards S, Aziz N, Bale S, *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015; **17**: 405-424

28 Feliubadaló L, Moles-Fernández A, Santamariña-Pena M, *et al.* A Collaborative Effort to Define Classification Criteria for ATM Variants in Hereditary Cancer Patients . *Clin Chem* 2020; **67**

29 Parada GE, Munita R, Cerda CA, *et al.* A comprehensive survey of non-canonical splice sites in the human transcriptome. *Nucleic Acids Res* 2014; **42**: 10564-10578

30 Laduca H, Stuenkel AJ, Dolinsky JS, *et al.* Utilization of multigene panels in hereditary cancer predisposition testing: Analysis of more than 2,000 patients. *Genet Med* 2014; **16**: 830-837

31 van Marcke C, Collard A, Vikkula M, *et al.* Prevalence of pathogenic variants and variants of unknown significance in patients at high risk of breast cancer: A systematic review and meta-analysis of gene-panel data. *Crit Rev Oncol Hematol* 2018; **132**: 138-144

32 Landrith T, Li B, Cass AA, *et al.* Splicing profile by capture RNA-seq identifies pathogenic germline variants in tumor suppressor genes. *npj Precis Oncol* 2020; **4**

33 Baralle D, Lucassen A, Buratti E. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep* 2009; **10**: 810-816

34 Fraile-Bethencourt E, Díez-Gómez B, Velásquez-Zapata V, *et al.* Functional classification of DNA variants by hybrid minigenes: Identification of 30 spliceogenic variants of BRCA2 exons 17 and 18. *PLoS Genet* 2017; **13**: e1006691

35 Montalban G, Fraile-Bethencourt E, López-Perolio I, *et al.* Characterization of spliceogenic variants located in regions linked to high levels of alternative splicing:

BRCA2 c.7976+5G > T as a case study. *Hum Mutat* 2018; **39**: 1155-1160

36 Dörk T, Bendix-Waltes R, Wegner RD, *et al.* Slow Progression of Ataxia-Telangiectasia with Double Missense and in Frame Splice Mutations. *Am J Med Genet* 2004; **126 A**: 272-277

37 Laake K, Jansen L, Hahnemann JM, *et al.* Characterization of ATM mutations in 41 Nordic families with ataxia telangiectasia. *Hum Mutat* 2000; **16**: 232-246

38 Teraoka SN, Telatar M, Becker-Catania S, *et al.* Splicing defects the ataxia-telangiectasia gene, ATM: Underlying mutations and consequences. *Am J Hum Genet* 1999; **64**: 1617-1631

39 Soukupova J, Dundr P, Kliebl Z, *et al.* Contribution of mutations in ATM to breast cancer development in the Czech population. *Oncol Rep* 2008; **19**: 1505-1510

40 Stankovic T, Kidd AMJ, Sutcliffe A, *et al.* ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: Expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. *Am J Hum Genet* 1998; **62**: 334-345

41 Cavaciuti E, Laugé A, Janin N, *et al.* Cancer risk according to type and location of ATM mutation in ataxia-telangiectasia families. *Genes Chromosom Cancer* 2005; **42**: 1-9

42 Demuth I, Dutrannoy V, Marques W, *et al.* New mutations in the ATM gene and clinical data of 25 at patients. *Neurogenetics* 2011; **12**: 273-282

43 Sandoval N, Platzer M, Rosenthal A, *et al.* Characterization of ATM gene mutations in 66 ataxia telangiectasia families. *Hum Mol Genet* 1999; **8**: 69-79

44 Broeks A, de Klein A, Floore AN, *et al.* ATM germline mutations in classical ataxia-telangiectasia patients in the Dutch population. *Hum Mutat* 1998; **12**: 330-337

45 Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, *et al.* Predicting Splicing from Primary Sequence with Deep Learning. *Cell* 2019; **176**: 535-548.e24

46 Ha C, Kim J-W, Jang J-H. Performance Evaluation of SpliceAI for the Prediction of Splicing of NF1 Variants. *Genes (Basel)* 2021; **12**: 1308

47 Moles-Fernández A, Domènech-Vivó J, Tenés A, *et al.* Role of splicing regulatory elements and in silico tools usage in the identification of deep intronic splicing variants in hereditary breast/ovarian cancer genes. *Cancers (Basel)* 2021; **13**

48 Riepe T V., Khan M, Roosing S, *et al.* Benchmarking deep learning splice prediction tools using functional splice assays. *Hum Mutat* 2021; **42**: 799-810

49 Wai HA, Lord J, Lyon M, *et al.* Blood RNA analysis can increase clinical diagnostic rate and resolve variants of uncertain significance. *Genet Med* 2020; **22**: 1005-1014

50 Thanaraj TA, Clark F. Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions. *Nucleic Acids Res* 2001; **29**: 2581-2593

51 Churbanov A, Winters-Hilt S, Koonin E V., *et al.* Accumulation of GC donor splice signals in mammals. *Biol Direct* 2008; **3**: 30

52 Ding H, Mao C, Li SM, *et al.* Lack of association between ATM C.1066-6T > G mutation and breast cancer risk: A meta-analysis of 8,831 cases and 4,957 controls. *Breast Cancer Res Treat* 2011; **125**: 473-477

53 Brnich SE, Abou Tayoun AN, Couch FJ, *et al.* Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med* 2020; **12**: 3

54 Abou Tayoun AN, Pesaran T, DiStefano MT, *et al.* Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat* 2018; **39**: 1517-1524

55 Lee JH, Paull TT. Cellular functions of the protein kinase ATM and their relevance to human disease. *Nat Rev Mol Cell Biol* 2021; **22**: 796-814

56 Jansma M, Linke-Winnebeck C, Eustermann S, *et al.* Near-Complete Structure and Model of Tel1ATM from Chaetomium thermophilum Reveals a Robust Autoinhibited ATP State. *Structure* 2020; **28**: 83-95.e5

57 Stakyte K, Rotheneder M, Lammens K, *et al.* Molecular basis of human ATM kinase inhibition. *Nat Struct Mol Biol* 2021; **28**: 789-798

58 Baretic D, Pollard HK, Fisher DI, *et al.* Structures of closed and open conformations of dimeric human ATM. *Sci Adv* 2017; **3**: e1700933

59 Xiao J, Liu M, Qi Y, *et al.* Structural insights into the activation of ATM kinase. *Cell Res* 2019; **29**: 683-685

60 Xin J, Xu Z, Wang X, *et al.* Structural basis of allosteric regulation of Tel1/ATM kinase. *Cell Res* 2019; **29**: 655-665

61 Yates LA, Williams RM, Hailemariam S, *et al.* Cryo-EM Structure of Nucleotide-Bound Tel1ATM Unravels the Molecular Basis of Inhibition and Structural Rationale for Disease-Associated Mutations. *Structure* 2020; **28**: 96-104.e3

62 Kozlov S V., Graham ME, Peng C, *et al.* Involvement of novel autophosphorylation sites in ATM activation. *EMBO J* 2006; **25**: 3504-3514

63 Cavalieri S, Funaro A, Porcedda P, *et al.* ATM mutations in Italian families with ataxia telangiectasia include two distinct large genomic deletions. *Hum Mutat* 2006; **27**: 1061

64 Sun X, Becker-Catania SG, Chun HH, *et al.* Early diagnosis of ataxia-telangiectasia using radiosensitivity testing. *J Pediatr* 2002; **140**: 724-731

65 Buzin CH, Gatti RA, Nguyen VQ, *et al.* Comprehensive scanning of the ATM gene with DOVAM-S. *Hum Mutat* 2003; **21**: 123-131

66 Soukupova J, Pohlreich P, Seemanova E. Characterisation of ATM mutations in Slavic ataxia telangiectasia patients. *NeuroMolecular Med* 2011; **13**: 204-211

67 Pietrucha BM, Heropolitańska-Pliszka E, Wakulińska A, *et al.* Ataxia-telangiectasia with hyper-IgM and wilms tumor: Fatal reaction to irradiation. *J Pediatr Hematol Oncol* 2010; **32**: e28-e30

68 Allinen M, Launonen V, Laake K, *et al.* ATM mutations in Finnish breast cancer patients [3]. *J Med Genet* 2002; **39**: 192-196

69 Jacquemin V, Rieunier G, Jacob S, *et al.* Underexpression and abnormal localization of ATM products in ataxia telangiectasia patients bearing ATM missense mutations. *Eur J Hum Genet* 2012; **20**: 305-312

70 Pylkäs K, Tommiska J, Syrjäkoski K, *et al.* Evaluation of the role of Finnish ataxia-telangiectasia mutations in hereditary predisposition to breast cancer. *Carcinogenesis* 2007; **28**: 1040-1045

71 Lee JH, Mand MR, Kao CH, *et al.* ATM directs DNA damage responses and proteostasis via genetically separable pathways. *Sci Signal* 2018; **11**

72 Tian Y, Pesaran T, Chamberlin A, *et al.* REVEL and BayesDel outperform other in silico meta-predictors for clinical variant classification. *Sci Rep* 2019; **9**

73 Verhagen MMM, Abdo WF, Willemsen MAAP, *et al.* Clinical spectrum of ataxia-

telangiectasia in adulthood. *Neurology* 2009; **73**: 430-437

74 Tavtigian S V., Oefner PJ, Babikyan D, *et al.* Rare, Evolutionarily Unlikely Missense Substitutions in ATM Confer Increased Risk of Breast Cancer. *Am J Hum Genet* 2009; **85**: 427-446

75 Mitui M, Bernatowska E, Pietrucha B, *et al.* ATM gene founder haplotypes and associated mutations in Polish families with ataxia-telangiectasia. *Ann Hum Genet* 2005; **69**: 657-664

76 Cavalieri S, Funaro A, Pappi P, *et al.* Large genomic mutations within the ATM gene detected by MLPA, including a duplication of 41 kb from exon 4 to 20. *Ann Hum Genet* 2008; **72**: 10-18

77 Davis MY, Keene CD, Swanson PD, *et al.* Novel mutations in ataxia telangiectasia and AOA2 associated with prolonged survival. *J Neurol Sci* 2013; **335**: 134-138

78 Verhagen MMM, Last JI, Hogervorst FBL, *et al.* Presence of ATM protein and residual kinase activity correlates with the phenotype in ataxia-telangiectasia: A genotype-phenotype study. *Hum Mutat* 2012; **33**: 561-571

79 Mitui M, Campbell C, Coutinho G, *et al.* Independent mutational events are rare in the ATM gene: Haplotype prescreening enhances mutation detection rate. *Hum Mutat* 2003; **22**: 43-50

80 Birrell GW, Kneebone K, Nefedov M, *et al.* ATM mutations, haplotype analysis, and immunological status of Russian patients with ataxia telangiectasia. *Hum Mutat* 2005; **25**: 593

81 Nespoli L, Verri A, Tajè S, *et al.* A Precocious Cerebellar Ataxia and Frequent Fever Episodes in a 16-Month-Old Infant Revealing Ataxia-Telangiectasia Syndrome. *Case Reports Immunol* 2013; **2013**: 1-5

82 Telatar M, Teraoka S, Wang Z, *et al.* Ataxia-telangiectasia: Identification and detection of founder-effect mutations in the ATM gene in ethnic populations. *Am J Hum Genet* 1998; **62**: 86-97

83 Aygün FD, Nepesov S, Çokuğraş H, *et al.* Bladder Wall Telangiectasia in a Patient with Ataxia-Telangiectasia and How to Manage? *Case Rep Pediatr* 2015; **2015**: 1-3

84 Ghosh R, Harrison SM, Rehm HL, *et al.* Updated recommendation for the benign stand-alone ACMG/AMP criterion. *Hum Mutat* 2018; **39**: 1525-1530

85 Whiffin N, Roberts AM, Minikel E, *et al.* Using High-Resolution Variant Frequencies Empowers Clinical Genome Interpretation and Enables Investigation of Genetic Architecture. *Am J Hum Genet* 2019; **104**: 187-190

86 Biesecker LG, Harrison SM, ClinGen Sequence Variant Interpretation Working Group. The ACMG/AMP reputable source criteria for the interpretation of sequence variants. *Genet Med* 2018; **20**: 1687-1688

Table 1. Bioinformatics analysis and splicing outcomes of *ATM* variants.

Variant (HGVS)	MaxEntScan	Minigene FL-Transcript	PTC-transcripts	In-frame transcripts	Unknown transcripts
mgATM_ex4-9					
Wild type		(65.7%±0.7)	Δ(E7) (34.2%±0.7)		
c.332-5A>G	[-] 3'ss (7.2→1.6)	(40.0%±8.4)	Δ(E7) (36.7%±5.7) ▼(E5p4) Δ(E7) (16.9%±2.2) Δ[(E5)(E7)] (6.4%±0.3)		
c.332-1G>A	[-] 3'ss (7.2→1.6)	-	Δ[(E5)(E7)] (36.6%±2.0) Δ(E5p1) (19.8%±1.2)	Δ(E5) (43.6%±0.8)	
c.496G>A (p.Glu166Lys)	[↓] 5'ss (7.8→4.6)	(71.8%±1.0)	Δ(E7) (28.2%±1.0)		
c.496+5G>A	[↓] 5'ss (7.8→4.9)	(25.2%±0.3)	Δ(E7) (5.7%±0.1) Δ[(E5)(E7)] (33.0%±0.5)	Δ(E5) (36.1%±0.4)	
c.901G>T (p.Gly301Cys)	[-] 5'ss (7.1→ -3.6)	-	Δ(E7) (91.0%±0.2)		970-nt (9.0%±0.2)
c.901+2T>C	[-] 5'ss (7.1→ -0.7)	-	Δ(E7) (100.0%±0)		
c.901+3A>T	[-] 5'ss (7.08→1.88)	-	Δ(E7) (90.5%±1.1)		970-nt (9.5%±1.1)
c.902-1G>T	[-] 3'ss (7.6→ -1.0)	-	Δ(E8) (37.4%±0.2) Δ(E7_8) (62.6%±0.2)		
c.902G>A (p.Gly301Asp)	[↓] 3'ss (7.6→6.9)	(64.8%±0.9)	Δ(E7) (22.6%±0.5) Δ(E7_8) (12.6%±1.0)		
c.903T>G (p.Gly301=)	[↓] 3'ss (7.6→6)	(61.5%±0.3)	Δ(E7) (21.3%±0.3) Δ(E7_8) (17.2%±0.1)		
c.1065+1G>T	[-] 5'ss (8.7→0.2)	-	▼(E8q5) (54.8%±0.6) [Δ(E7)▼(E8q5)] (45.2%±0.6)		
c.1065+3A>G	[-] 5'ss (8.7→5.0)	(25.3%±0.6)	Δ(E7) (5.7%±0.1) [Δ(E7)▼(E8q5)] (8.9%±0.5) Δ(E7_8) (28.8%±0.8) Δ(E8) (6.1%±0.4) ▼(E8q5) (25.2%±1.3)		
c.1066-6T>G	[↓] 3'ss (10.8→8.3)	(26.7%±1.2)	Δ(E9) (44.8%±0.9) Δ[(E7)(E9)] (14.9%±0.4) Δ(E7) (3.3%±0.1)	Δ(E7_9) (10.3%±0.2)	
c.1235+4_1235+5del	[-] 5'ss (4.1→ -1.3)	(45.1%±0.4)	Δ(E9) (19.4%±0.2) Δ[(E7)(E9)] (11.3%±0.1) Δ(E7) (12.8%±0.2)	Δ(E7_9) (11.4%±0.2)	
mgATM _ex11-17					
Wild type		(84.1%±0.6)		Δ(E11) (15.9%±0.6)	
c.1898G>T (p.Cys633Phe)	[↓] 5'ss (8.4→4.3)	(62.6%±2.9)		Δ(E12) (37.4%±2.9)	
c.1898+2T>G	[-] 5'ss (8.4→0.7)	(13.0%±1.3)[3]		Δ(E12) (87.0%±1.3)	
c.1898+3A>T	[↓] 5'ss (8.4→4.9)	(100.0%±0.0)			
c.1898+3_1898+4del	[-] 5'ss (8.4→-5.1)	-		Δ(E12) (100.0%±0.0)	
c.2251-1G>C	[-] 3'ss (6.6→ -1.4) [+] 5'ss (4.5)	-	Δ(E15p19) (57.4%±0.4)	Δ(E15) (11.7%±0.7) Δ[(E11)(E15)] (18.2±1.1) Δ(E15_16) (12.7%±0.8)	

Variant (HGVS)	MaxEntScan	Minigene FL-Transcript	PTC-transcripts	In-frame transcripts	Unknown transcripts
c.2376+1G>A	[-] 5'ss (10.6→2.4)	-		Δ(E15) (81.3%±0.6) Δ[(E11)(E15)] (18.7%±0.6)	
c.2376+3A>T	[↓] 5'ss (10.6→6.1)	-		Δ(E15) (80.0%±0.0) Δ[(E11)(E15)] (20.0%±0.0)	
c.2377-6T>A	[↓] 5'ss (8.1→6.1)	(83.3%±0.2)		Δ(E11) (16.7%±0.2)	
c.2377-2A>G	[-] 3'ss (8.1→0.1)	-		Δ(E16) (45.2%±0.4) Δ(E16p3) (38.7%±0.6) Δ(E11)Δ(E16) (16.1%±0.2)	
c.2467-3A>G	[-] 3'ss (8.8→0.9)	(100.0%±0.0)			
c.2638+3A>G	[↓] 5'ss (6.6→3.8)	(100.0%±0.0)			
mgATM _ex25_29					
Wild type		(100.0%±0.0)			
c.3577-1G>A	[-] 3'ss (8.5→ -0.2) [+] 3'ss (3.6)	-	Δ(E25) (100%±0.0)		
c.3577G>C (p.Val1931Leu)	[↓] 3'ss (8.5→7.4)	(37.3%±0.5)	Δ(E25) (47.1%±0.8)	Δ(E25p159) (7.9%±0.1) Δ(E25_26) (7.7%±1.0)	
c.3746+1G>A	[-] 5'ss (9.7→1.5)	-	Δ(E25) (100%±0.0)		
c.3746+4A>C	[↓] 5'ss (9.7→8.2)	(71.4%±1)	Δ(E25) (28.6%±1)		
c.3746+5G>A	[↓] 5'ss (9.7→6.6)	-	Δ(E25) (100%±0.0)		
c.3993G>A (p.Gln1331=)	[-] 5'ss (10.0→5.0)	-	Δ(E26) (27.1%±0.5)	Δ(E26q120) (72.9%±0.5)	
c.3993+1G>A	[-] 5'ss (10.0→1.8)	-	Δ(E26) (32.7%±0.7)	Δ(E26q120) (67.3%±0.7)	
c.3993+5G>T	[↓] 5'ss (10.0→8.4)	(76.2%±0.8)		Δ(E26q120) (23.8%±0.8)	
c.3994-3C>T	[↓] 3'ss (11.4→9.9)	(100%±0.0)			
c.3994-2A>G	[-] 3'ss (11.4→3.4)	-	Δ(E27) (100%±0.0)		
c.4109+1G>T	[-] 5'ss (8.3→ -0.2)	-	Δ(E27) (83.4%±0.4) Δ(E27q1) (16.6%±0.4)		
c.4109+3A>G	[↓] 5'ss (8.3→4.5)	(45.7%±0.9)	Δ(E27) (54.3%±0.9)		
c.4109+5G>A	[-] 5'ss (8.3→3.9)	-	Δ(E27) (100.0%±0.0)		
c.4109+6T>G	[↓] 5'ss (8.3→7.0)	(68.5%±0.8)	Δ(E27) (31.5%±0.8)		
c.4110-9C>G	[-] 3'ss (5.6→ -0.1) [+] 3'ss (5.6)	-	Δ(E28) (15.2%±2.5) ▼(E28p8) (84.8%±2.5)		
c.4110-2A>C	[-] 3'ss (5.6→ -2.5)	-	Δ(E28) (84.2%±0.1) Δ(E28p53) (15.8%±0.1)		
c.4236+1G>A	[-] 5'ss (7.5→ -0.7)	-	Δ(E28) (100%±0.0)		
c.4236+5G>A	[-] 5'ss (7.5→ -0.6)	-	Δ(E28) (100%±0.0)		
c.4236+6T>C	[-] 5'ss (7.5→3.7)	-	Δ(E28) (100%±0.0)		
c.4436+4A>G	[↓] 5'ss (8.9→7.2)	(72.8%±3.1)	Δ(E29) (21.1%±0.3)	Δ(E28_29) (6.1%±3.3)	
mgATM _ex49-52					
Wild type		(75.8%±0.04)		Δ(E52) (24.2% ±0.04)	
c.7307+1G>A	[-] 5'ss (8.6→0.4)	-	Δ(E49q38) (62.5%±1.0) [Δ(E49q38) Δ(E52)] (17.7%±0.6) Δ(E49) (8.7%±0.1) [Δ(E49) Δ(E52)] (8.1%±0.2)		
c.7307+4A>G	[↓] 5'ss (8.6→6.5)	-	Δ(E49q38) (86.5%±0.2) [Δ(E49q38) Δ(E52)] (13.5%±0.2)		

Variant (HGVS)	MaxEntScan	Minigene FL-Transcript	PTC-transcripts	In-frame transcripts	Unknown transcripts
c.7515G>A (p.Lys2505=)	[↓] 5'ss (3.2→2.0)	-	Δ(E50) (65.6%±1.1) [Δ(E50) Δ(E52)] (34.4%±1.1)		
c.7515+6T>C	[↓] 5'ss (3.2→2.5)	(48.4%±2.1)	Δ(E50) (34.3%±0.8) [Δ(E50) Δ(E52)] (17.3%±1.3)		
c.7629+2T>G	[-] 5'ss (8.6→1.0)	-		Δ(E51) (100%±0)	
c.7630-3C>T	[↓] 3'ss (7.0→5.5)	(22.1%±0.2)		Δ(E52) (77.9%±0.2)	
c.7630-2A>C	[-] 3'ss (7.0→ -1.1) [+] 3'ss (5.4)	-	Δ(E52p11) (33.6%±1.3)	Δ(E52) (66.4%±1.3)	
c.7787A>T (p.Glu2596Val)	[↓] 5'ss (7,6→3,1)	-		Δ(E52) (100% ±0)	
c.7788G>A (p.Glu2596=)	[-] 5'ss (7.6→0,9)	-		Δ(E52) (100% ±0)	
c.7788+1G>C	[-] 5'ss (7.6→0.9)	-		Δ(E52) (100% ±0)	
c.7788+6T>G	[-] 5'ss (7.6→ -0.6)	(50.9%±4.2)		Δ(E52) (49.1%±4.2)	

[1] Variants without any trace of the full-length transcripts are underlined. [2] [-] site disruption; [+] New site; [↓] Decrease of the splice-site score. [3] Use of a non-canonical GG 5'ss [29].

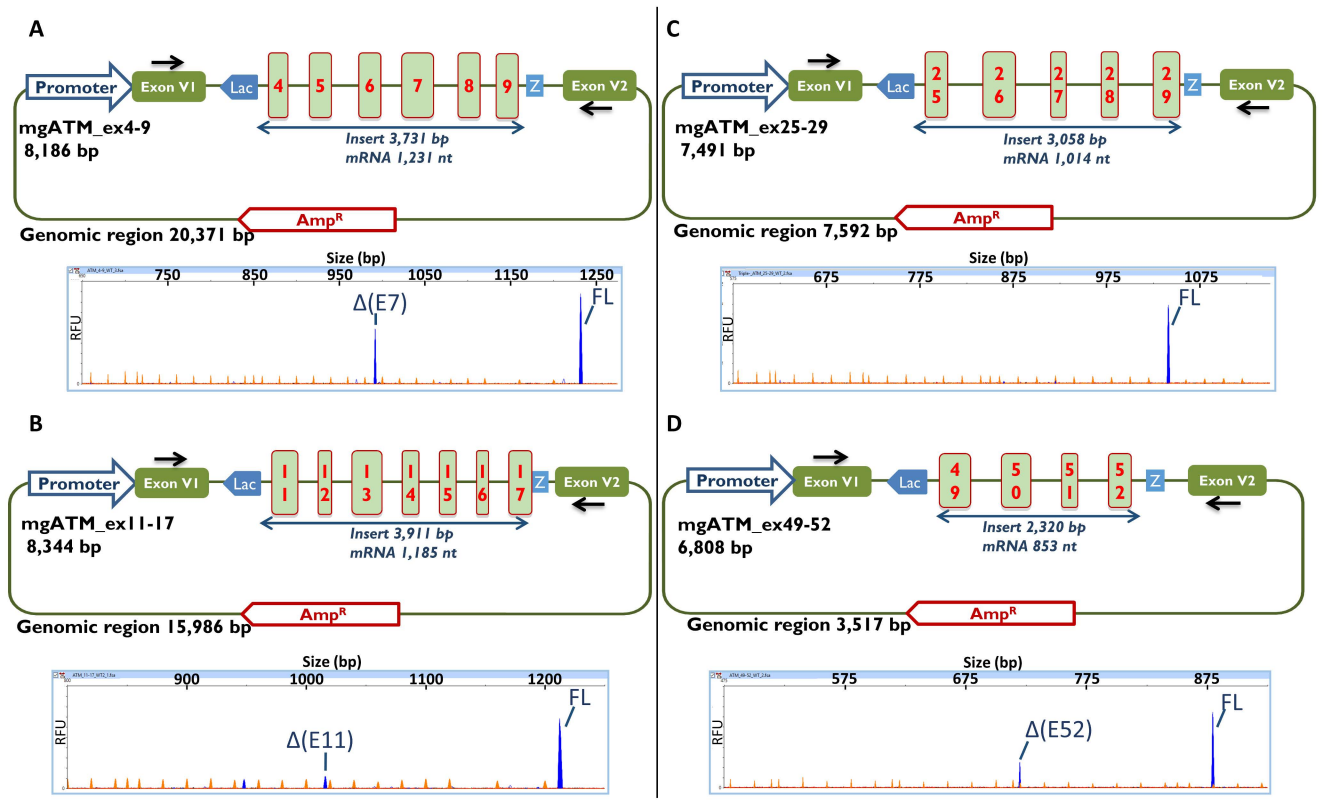
Table 2. ACMG/AMP-based tentative classification of 56 *ATM* variants.

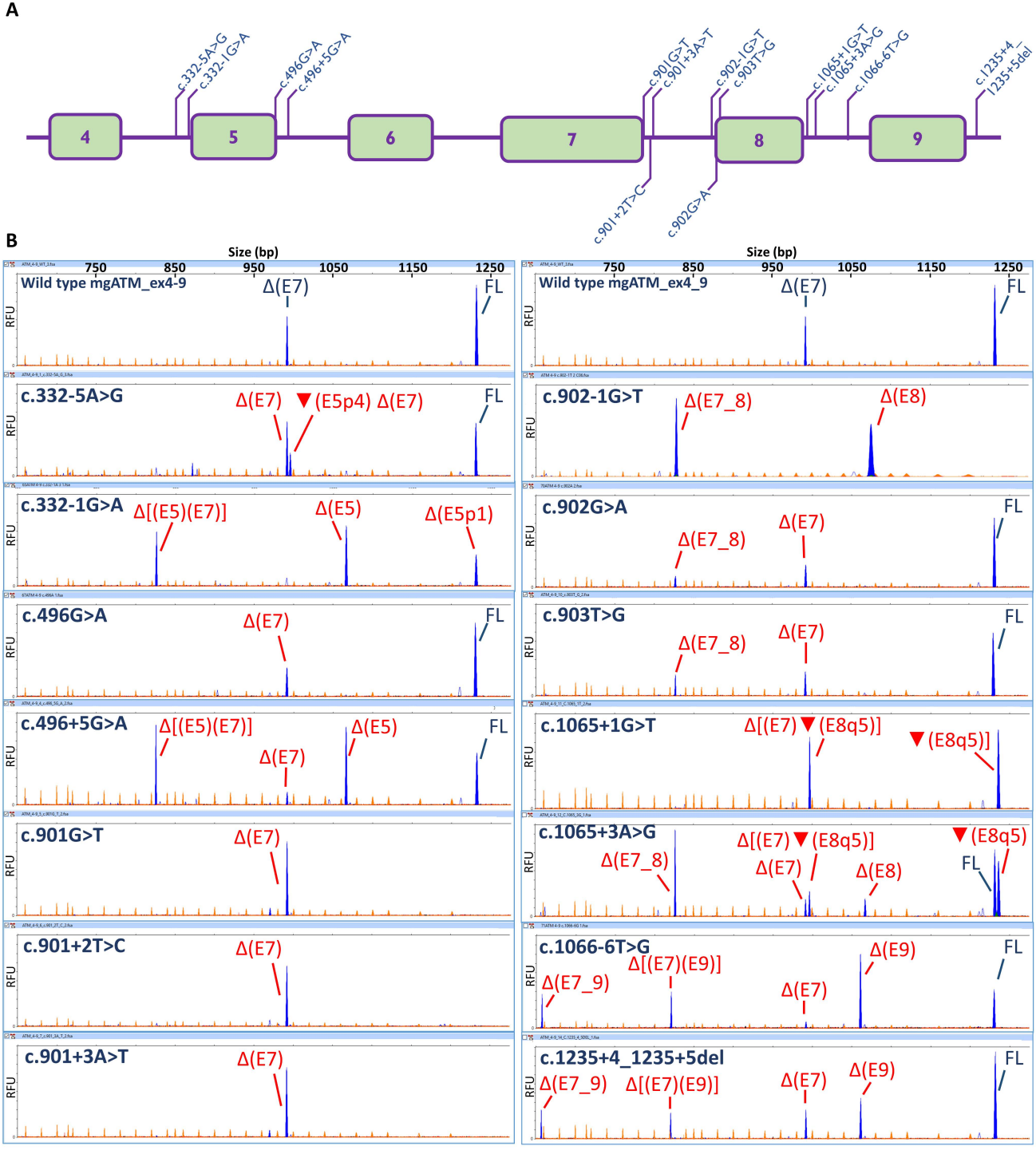
c.HGVS [1]	p.HGVS[1]	ClinVar[2]	Point-based classification[3]		PVS1_O[4]	PM2[5]	PM3[6]	BS7_O[4]	BS1[7]
c.332-5A>G		not reported	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.332-1G>A		LP (2)	**Pathogenic**	**+11 (+8+1+2)**	PVS1_O**(+8)**	PM2_P **(+1)**	PM3**(+2)**	(-)	(-)
c.496G>A	p.(Glu166Lys)	VUS (3)	Uncertain	0 (+1-1)	(-)	PM2_P **(+1)**	(-)	BP7_O**(-1)**	(-)
c.496+5G>A		LP(7)/P(1)	Uncertain	+5 (+1+4)	(-)	PM2_P **(+1)**	PM3_S**(+4)**	(-)	(-)
c.901G>T	p.(Gly301Cys)	not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.901+2T>C		LP (1)	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.901+3A>T		VUS (2)	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.902-1G>T		P (5)	**Pathogenic**	**+11 (+8+1+2)**	PVS1_O**(+8)**	PM2_P **(+1)**	PM3**(+2)**	(-)	(-)
c.902G>A	p.(Gly301Asp)	VUS (17)	Uncertain	(+1)	PVS1_O_P**(+1)**	(-)	(-)	(-)	(-)
c.903T>G	p.(Gly301=)	LB(3)	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.1065+1G>T		LP(7)/P(1)	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.1065+3A>G		VUS (1)	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.1066-6T>G		B(6)/LB(6)/VUS(5)	**Likely Benign**	**(-4)**	(-)	(-)	(-)	(-)	BS1**(-4)**
c.1235+4_1235+5 del		not reported	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.1898G>T	p.(Cys633Phe)	not reported	Uncertain	+2 (+1+1)	PVS1_O_P**(+1)**	PM2_P **(+1)**	(-)	(-)	(-)
c.1898+2T>G		LP(2)/P(7)	**Likely Pathogenic**	**+9 (+1+8)**	(-)	PM2_P **(+1)**	PM3_VS**(+8)**	(-)	(-)
c.1898+3A>T		not reported	**Likely Benign**	**-3 (+1-4)**	(-)	PM2_P **(+1)**	(-)	BP7_O_S**(-4)**	(-)
c.1898+3_1898+4 del		not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.2251-1G>C		LP(1)/P(1)	Uncertain	+5 (+4+1)	PVS1_O_S**(+4)**	PM2_P **(+1)**	(-)	(-)	(-)
c.2376+1G>A		LP(4)	**Likely Pathogenic**	+7 (+4+1+2)	PVS1_O_S**(+4)**	PM2_P **(+1)**	PM3**(+2)**	(-)	(-)
c.2376+3A>T		VUS(1)	Uncertain	+5 (+4+1)	PVS1_O_S**(+4)**	PM2_P **(+1)**	(-)	(-)	(-)
c.2377-6T>A		LB(1)/VUS(3)	**Likely Benign**	**-3 (+1-4)**	(-)	PM2_P **(+1)**	(-)	BP7_O_S**(-4)**	(-)

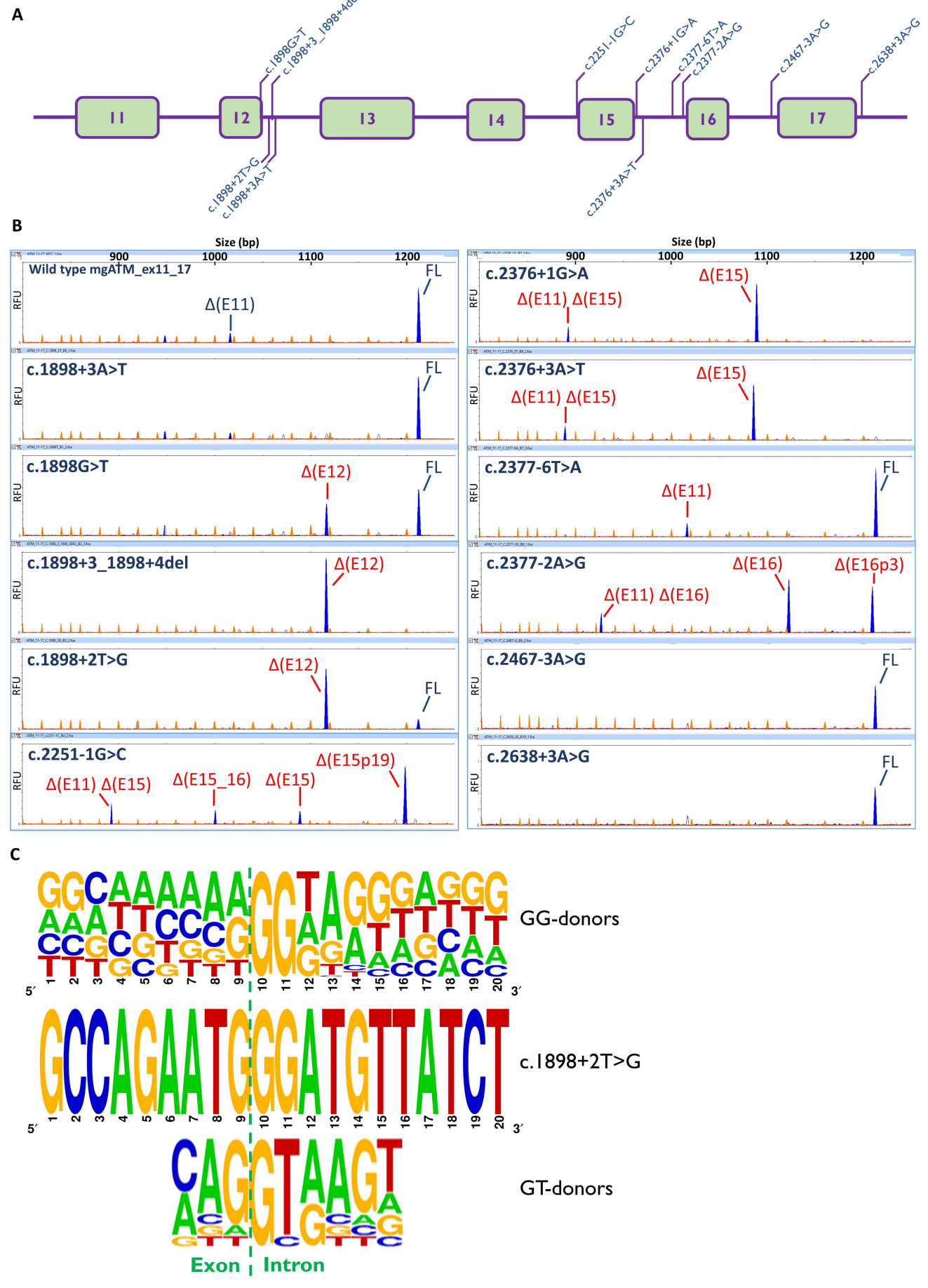
c.2377-2A>G		LP(4)	Uncertain	+2 (+1+1)	PVS1_O_P**(+1)**	PM2_P **(+1)**	(-)	(-)	(-)
c.2467-3A>G		not reported	**Likely Benign**	**-3 (+1-4)**	(-)	PM2_P **(+1)**	(-)	BP7_O_S**(-4)**	(-)
c.2638+3A>G		LB(2)/VUS(2)	**Likely Benign**	**-3 (+1-4)**	(-)	PM2_P **(+1)**	(-)	BP7_O_S**(-4)**	(-)
c.3577-1G>A		not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.3577G>C	p.(Val1193Leu)	VUS(8)	Uncertain		(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.3746+1G>A		not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.3746+4A>C		LB(1)/VUS(2)	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.3746+5G>A		VUS(3)	**Likely Pathogenic**	**+9(+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.3993G>A	p.(Gln1331=)	VUS(1)	Uncertain	+3(+2+1)	PVS1_O_M**(+2)**	PM2_P **(+1)**	(-)	(-)	(-)
c.3993+1G>A		LP(1)/P(4)	**Likely Pathogenic**	**+6(+2+4)**	PVS1_O_M**(+2)**	(-)	PM3_S**(+4)**	(-)	(-)
c.3993+5G>T		B(9)/LB(3)	**Likely Benign**	**(-4)**	(-)	(-)	(-)	(-)	BS1**(-4)**
c.3994-3C>T		not reported	**Likely Benign**	**-3 (+1-4)**	(-)	PM2_P **(+1)**	(-)	BP7_O_S**(-4)**	(-)
c.3994-2A>G		LP(3)/P(1)	**Likely Pathogenic**	**(+8)**	PVS1_O**(+8)**	(-)	(-)	(-)	(-)
c.4109+1G>T		LP(2)	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.4109+3A>G		VUS(1)	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.4109+5G>A		VUS(1)	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.4109+6T>G		LB(4)/VUS(1)	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.4110-9C>G		VUS(4)/P(1)	**Pathogenic**	**+11(+8+1+2)**	PVS1_O**(+8)**	PM2_P **(+1)**	PM3**(+2)**	(-)	(-)
c.4110-2A>C		not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.4236+1G>A		not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.4236+5G>A		VUS(1)	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.4236+6T>C		not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.4436+4A>G		not reported	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.7307+1G>A		LP (1)	**Pathogenic**	**+11(+8+1+2)**	PVS1_O**(+8)**	PM2_P **(+1)**	PM3**(+2)**	(-)	(-)
c.7307+4A>G		VUS(3)	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)

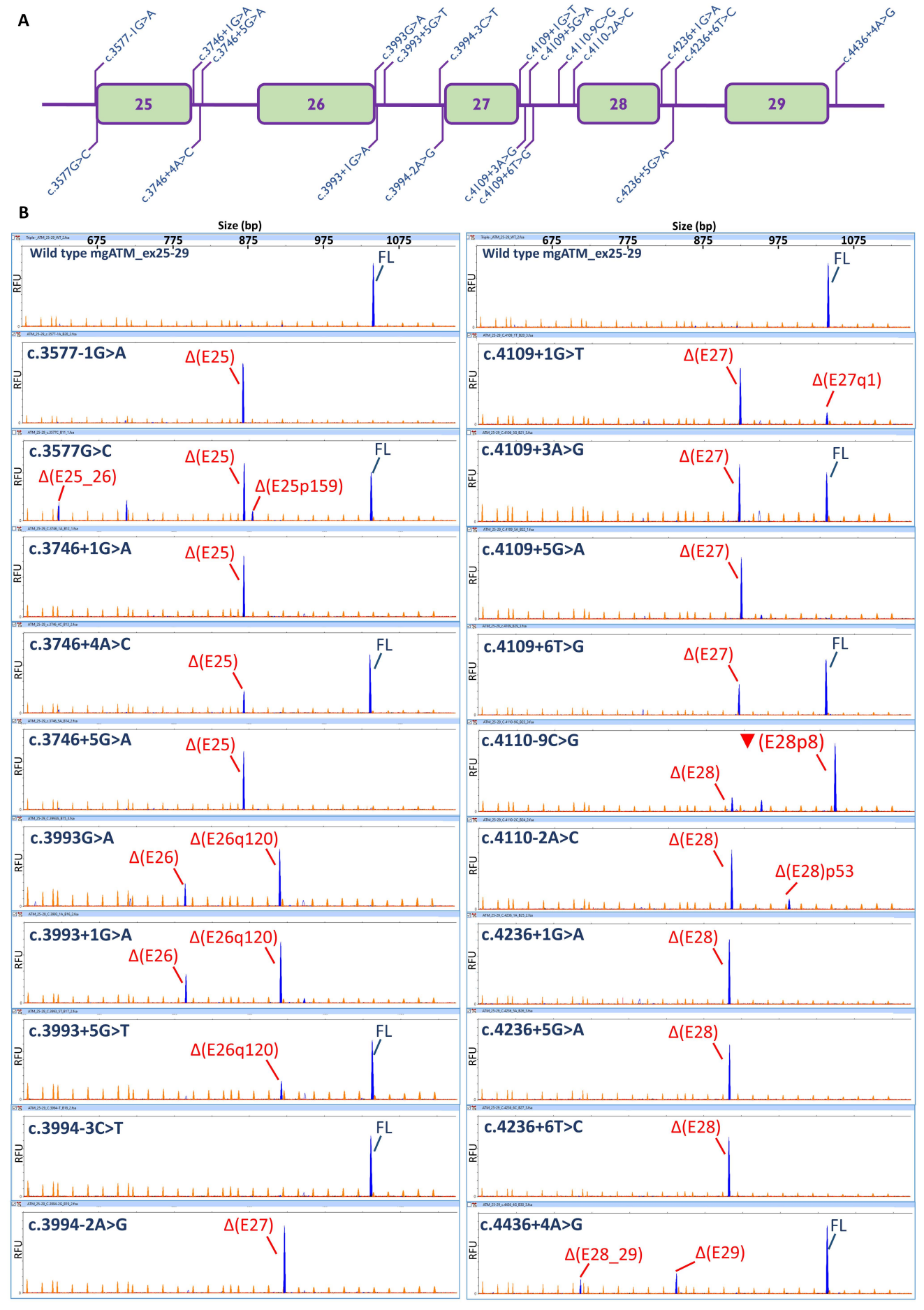
c.7515G>A	p.(Lys2505=)	not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.7515+6T>C		VUS(1)	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.7629+2T>G		not reported	**Likely Pathogenic**	**+9 (+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.7630-3C>T		LB(2)/VUS(2)	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)
c.7630-2A>C		LP(1)/P(15)	**Pathogenic**	**+16 (+8+8)**	PVS1_O**(+8)**	(-)	PM3_VS**(+8)**	(-)	(-)
c.7787A>T	p.(Glu2596Val)	not reported	**Likely Pathogenic**	**+9(+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.7788G>A	p.(Glu2596=)	LP(2)/P(5)	**Pathogenic**	**+13(+8+1+4)**	PVS1_O**(+8)**	PM2_P **(+1)**	PM3_S**(+4)**	(-)	(-)
c.7788+1G>C		not reported	**Likely Pathogenic**	**+9(+8+1)**	PVS1_O**(+8)**	PM2_P **(+1)**	(-)	(-)	(-)
c.7788+6T>G		not reported	Uncertain	(+1)	(-)	PM2_P **(+1)**	(-)	(-)	(-)

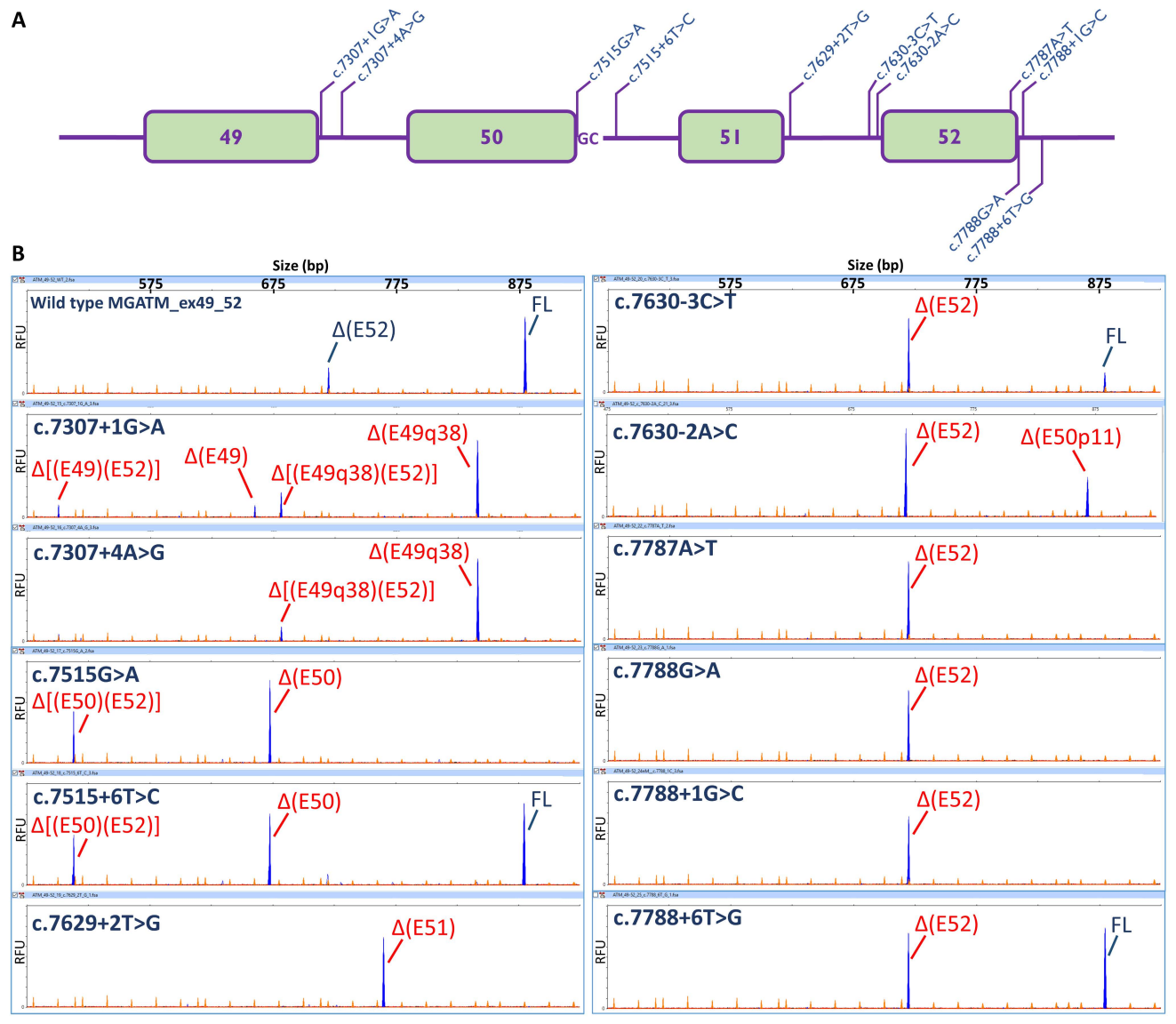
The table shows 56 ATM variants identified in the BRIDGES cohort, its current ClinVar clinical classification, and the ACMG/AMP-based tentative classification that we have performed by combining existing pathogenic and benign evidences with the mgATM data produced in the present study. For each individual *ATM* variant, we have evaluated all applicable evidences, but the table shows only evidences contributing to the final classification. [1]NM_000051.3. [2]ClinVar last accessed 09/02/2022. LB (Likely Benign), VUS (variant of uncertain significance), LP (Likely Pathogenic), P (Pathogenic). In brackets, N submitters supporting each classification. (-) not reported. [3] We used an ACMG/AMP point system Bayesian framework to combine all pathogenic and benign evidences. [4]We have integrated mgATM splicing functional data in the classification scheme as a pathogenic evidence code PVS1_O (or as a benign evidence code BP7_O) of variable strength, as per ClinGen *ATM* expert panel ACMG-AMP specifications. Of note, we have used some ACMG-AMP evidences (not shown in the present table) to decide the PVS1_O or BP7_O code strength applicable to complex mgATM read-outs). [5]Rarity evidence PM2 downgraded to supporting strength, as *per* ClinGen *ATM* expert panel ACMG-AMP specifications. [6]We have assigned the recessive disorders evidence PM3 to variants identified in *trans* with a pathogenic variant in A-T patients (as reported in the scientific literature). Code strength as *per* ClinGen recommendations. BS1 evidence (allele frequency greater than expected for disease) applied as per ClinGen *ATM* expert panel ACMG-AMP specifications. See Supplementary Table S3 and Supplementary Methods for further details.


A
Promoter
Exon V1
Lac
4
5
6
7
8
9
Z
Exon V2
mgATM_ex4-9
8,186 bp
Insert 3,731 bp
mRNA 1,231 nt
AmpR
Genomic region 20,371 bp
Size (bp)
750
850
950
1050
1150
1250
RFU
Δ(E7)
FL
B
Promoter
Exon V1
Lac
11
12
13
14
15
16
17
Z
Exon V2
mgATM_ex11-17
8,344 bp
Insert 3,911 bp
mRNA 1,185 nt
AmpR
Genomic region 15,986 bp
Size (bp)
900
1000
1100
1200
RFU
Δ(E11)
FL
C
Promoter
Exon V1
Lac
25
26
27
28
29
Z
Exon V2
mgATM_ex25-29
7,491 bp
Insert 3,058 bp
mRNA 1,014 nt
AmpR
Genomic region 7,592 bp
Size (bp)
675
775
875
975
1075
RFU
FL
D
Promoter
Exon V1
Lac
49
50
51
52
Z
Exon V2
mgATM_ex49-52
6,808 bp
Insert 2,320 bp
mRNA 853 nt
AmpR
Genomic region 3,517 bp
Size (bp)
575
675
775
875
RFU
Δ(E52)
FL


A
c.332-5A>G
c.332-1G>A
c.496G>A
c.496+5G>A
c.901G>T
c.901+3A>T
c.902-1G>T
c.903T>G
c.1065+1G>T
c.1065+3A>G
c.1066-6T>G
c.1235+4_1235+5del
c.901+2T>C
c.902G>A
4
5
6
7
8
9
B
Size (bp)
750
850
950
1050
1150
1250
RFU
Wild type mgATM_ex4-9
Δ(E7)
FL
c.332-5A>G
Δ(E7)
▼(E5p4) Δ(E7)
c.332-1G>A
Δ[(E5)(E7)]
Δ(E5)
Δ(E5p1)
c.496G>A
c.496+5G>A
Δ[(E5)(E7)]
c.901G>T
c.901+2T>C
c.901+3A>T
Wild type mgATM_ex4_9
c.902-1G>T
Δ(E7_8)
Δ(E8)
c.902G>A
c.903T>G
c.1065+1G>T
Δ[(E7)▼(E8q5)]
▼(E8q5)]
c.1065+3A>G
▼(E8q5)
c.1066-6T>G
Δ(E9)
Δ(E7_9)
Δ[(E7)(E9)]
c.1235+4_1235+5del


A
I1
I2
I3
I4
I5
I6
I7
c.1898G>T
c.1898+3_1898+4del
c.2251-1G>C
c.2376+1G>A
c.2377-6T>A
c.2377-2A>G
c.2467-3A>G
c.2638+3A>G
c.1898+2T>G
c.1898+3A>T
c.2376+3A>T
B
Size (bp)
900
1000
1100
1200
RFU
Wild type mgATM_ex11_17
FL
Δ(E11)
c.1898+3A>T
c.1898G>T
Δ(E12)
c.1898+3_1898+4del
c.1898+2T>G
c.2251-1G>C
Δ(E11) Δ(E15)
Δ(E15_16)
Δ(E15)
Δ(E15p19)
c.2376+1G>A
Δ(E15)
Δ(E11) Δ(E15)
c.2376+3A>T
c.2377-6T>A
c.2377-2A>G
Δ(E16)
Δ(E16p3)
Δ(E11) Δ(E16)
c.2467-3A>G
c.2638+3A>G
C
GG-donors
c.1898+2T>G
GT-donors
5′
3′
Exon
Intron

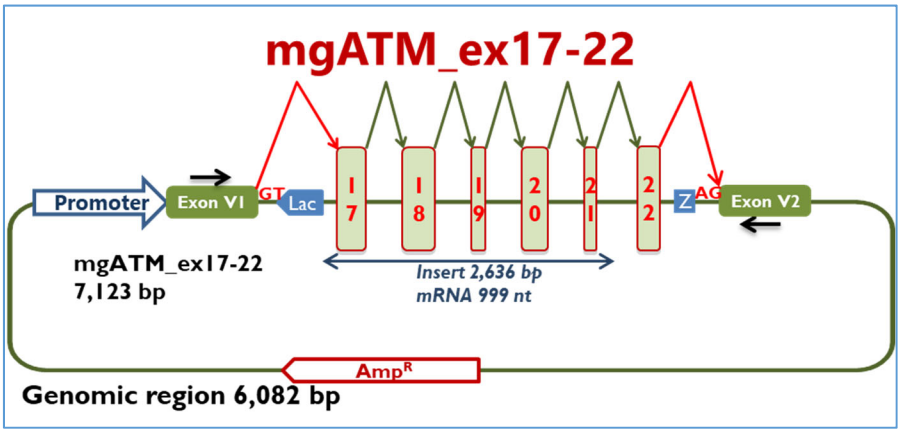

A
c.3577-1G>A
c.3746+1G>A
c.3746+5G>A
c.3993+5G>T
c.3993G>A
c.3994-3C>T
c.4109+1G>T
c.4109+5G>A
c.4110-9C>G
c.4110-2A>C
c.4236+1G>A
c.4236+6T>C
c.4436+4A>G
25
26
27
28
29
c.3577G>C
c.3746+4A>C
c.3993+1G>A
c.3994-2A>G
c.4109+3A>G
c.4109+6T>G
c.4236+5G>A
B
Size (bp)
675
775
875
975
1075
RFU
Wild type mgATM_ex25-29
FL
c.3577-1G>A
Δ(E25)
c.3577G>C
Δ(E25_26)
Δ(E25p159)
c.3746+1G>A
c.3746+4A>C
c.3746+5G>A
c.3993G>A
Δ(E26)
Δ(E26q120)
c.3993+1G>A
c.3993+5G>T
c.3994-3C>T
c.3994-2A>G
Δ(E27)
c.4109+1G>T
Δ(E27q1)
c.4109+3A>G
c.4109+5G>A
c.4109+6T>G
c.4110-9C>G
▼(E28p8)
Δ(E28)
c.4110-2A>C
Δ(E28)p53
c.4236+1G>A
c.4236+5G>A
c.4236+6T>C
c.4436+4A>G
Δ(E28_29)
Δ(E29)


A
c.7307+1G>A
c.7307+4A>G
c.7515G>A
c.7515+6T>C
c.7629+2T>G
c.7630-3C>T
c.7630-2A>C
c.7787A>T
c.7788+1G>C
49
50
GC
51
52
c.7788G>A
c.7788+6T>G
B
Size (bp)
575
675
775
875
RFU
Wild type MGATM_ex49_52
Δ(E52)
FL
c.7307+1G>A
Δ[(E49)(E52)]
Δ(E49)
Δ[(E49q38)(E52)]
Δ(E49q38)
c.7307+4A>G
c.7515G>A
Δ[(E50)(E52)]
Δ(E50)
c.7515+6T>C
c.7629+2T>G
Δ(E51)
c.7630-3C>T
c.7630-2A>C
Δ(E50p11)
c.7787A>T
c.7788G>A
c.7788+1G>C
c.7788+6T>G

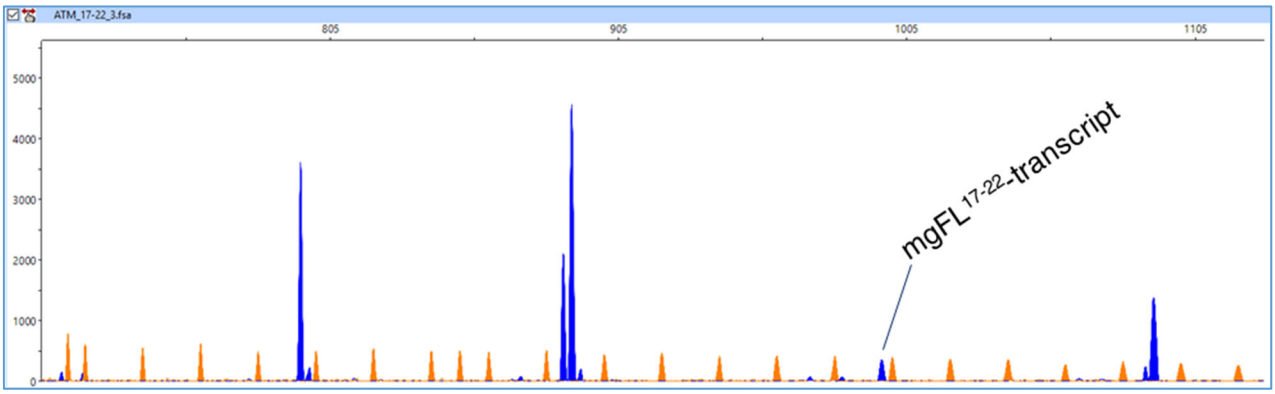
Other *ATM* minigenes

A minigene with exons 17 to 22 was constructed (mgATM_ex17-22). The insert was generated by gene synthesis (Genewiz, South Plainfield, NJ) and then cloned into the pSAD splicing vector. Its sequence and structure are shown in **Supplementary Figure S1** and it is represented in **Supplementary Figure S6A.** This construct was tested in MCF-7 where it produced at least five transcripts as it is shown in the capillary electrophoresis electropherogram of **Supplementary Figure S6B**. The minigene full-length transcript accounted for about 3% of all transcripts, so this construct should be disregarded, and 19 variants at exons 18 to 22 could not be assayed.

A)



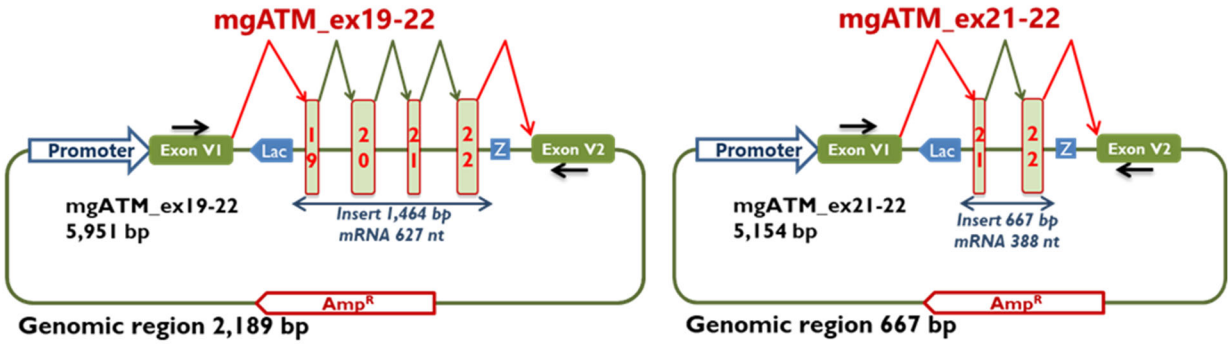
B)



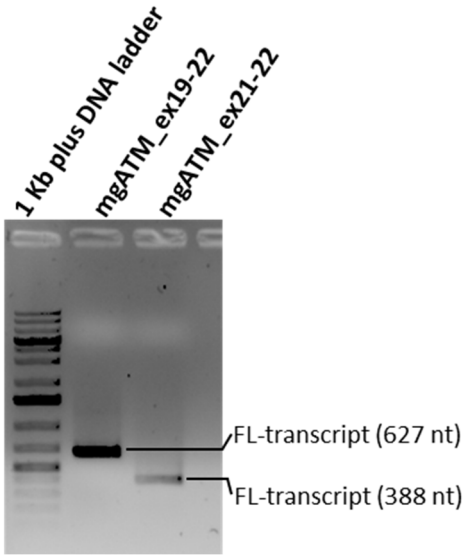
Supplementary Figure S6. Minigene mgATM_ex17-22. A) Minigene structure. Exons are indicated by boxes. B) Fluorescent fragment electrophoresis of the wild type minigene mgATM_ex17-22 in MCF-7 cells. FAM-labelled products (blue peaks) were run with LIZ1200 (orange peaks) as size standard. FL, Full-length transcript.

As indicated in Discussion, it is plausible that smaller constructs with less exons of this *ATM* region may functionally work. Then, to confirm this point we decided to test deletions of exons 17-18 and 17-20. Both deletions were introduced into the wt minigene by site-directed mutagenesis with primers 5' CGAATTGGAGCTCCACCGCGGTGGCGGCCGTTTTGTGTTTTTAGTAGAGATGGAGTT TCA 3' and 5' TGAAACTCCATCTCTACTAAAAACACAAAACGGCCGCCACCGCGGTGGAGCTCCAAT TCG 3' (for removing exons 17-18) and 5' CGAATTGGAGCTCCACCGCGGTGGCGGCCGGAGTATGTTGGCATATTCCACATAATG ACA 3' and 5' TGTCATTATGTGGAATATGCCAACATACTCCGGCCGCCACCGCGGTGGAGCTCCAAT TCG 3' (for removing exons 17 to 20). Both mutants were successfully obtained and assayed in MCF-7 cells (**Supplementary Figure S7**). Actually, the respective expected full-length transcripts were observed on agarose gel electrophoresis and their sequences were confirmed by Sanger sequencing (sequence files available at http://hdl.handle.net/10261/265669). Therefore, these minigenes are ready for future variant-splicing assays.

A)

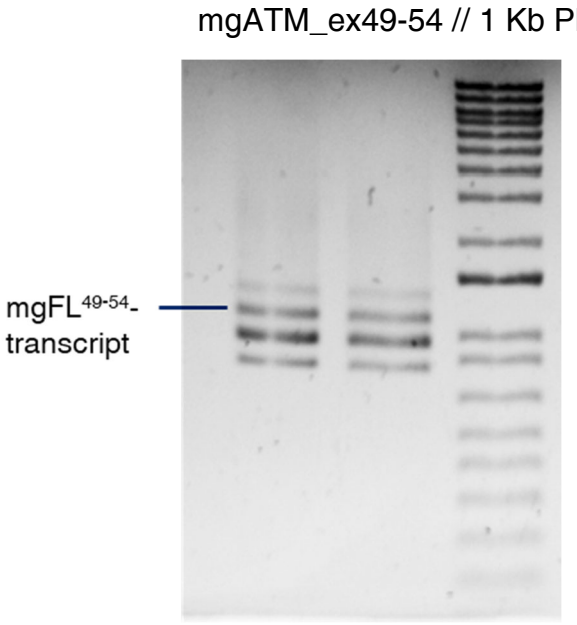


B)



Supplementary Figure S7. Structures and functional assays of the novel minigenes mgATM_ex19-22 and _ex21-22. A) Structure of the minigenes mgATM_ex19-22 (left) and _ex21-22 (right). Exons are indicated by boxes. B) Agarose gel electrophoresis of RT-PCR products of both minigenes.

Other minigenes, such as the insertions of exons 53 and 54 to mgATM_ex49-52 (other 5 additional candidate variants), were also tried and checked in MCF-7 cells. None of the clones (_ex49-53 and _ex49-54) yielded the expected mgFL-transcripts. The **Supplementary Figure S8** displays an agarose gel electrophoresis of the functional assays of two clones of mgATM_ex49-54 that revealed at least 4 transcripts where the mgFL$^{49\text{-}54}$-transcript is not the main one. In this situation, variant analysis is not feasible, so these clones were discarded and five variants at exons 53 and 54 could not be assayed.



Supplementary Figure S8. Agarose gel electrophoresis of RT-PCR products produced by minigene mgATM_ex49-54 (in duplicate).

ACMG/AMP-BASED TENTATIVE CLASSIFICATION OF 56 *ATM* VARIANTS.

To classify *ATM* variants according to the ACMG/AMP guidelines [27], we have taken into consideration two recently available *ATM* specifications:

(i) ATM ClinGen Hereditary Breast, Ovarian and Pancreatic Cancer (HBOPC) Expert Panel specifications (*HBOPC_ATMv1*), downloaded from https://www.clinicalgenome.org/affiliation/50039/

(ii) Spanish ATM Cancer Susceptibility Variant Interpretation Working Group (*Spanish_ATM_WG*) specifications [28].

If (I) and (II) demonstrated discrepancies, we have prioritized *HBOPC_ATMv1* specifications. If needed, we have developed *ad-hoc* rules as indicated.

ACMG/AMP evidences contributing to final classification.

1. PVS1_O_Variable/BP7_O_Variable (*well-established in vitro or in vivo functional studies support damaging/no-damaging effect*)**:**

As suggested by *HBOPC_ATMv1*, we have used PVS1_O_Variable and BP7_O_Variable codes (rather than PS3_Variable and BS3_Variable codes) to integrate mgATM data into the classification system.

Most *ATM* variants (33 out of 56) produce complex mgATM read-outs (≥ 2 transcripts). Neither generic ClinGen recommendations for functional evidence [53], nor HBOPC_ATMv1 (or *Spanish_ATM_WG)* provide guidance on how to interpret these complex splicing profiles. Consequently, to assign a pathogenic or benign code strength to mgATM read-outs we have decided to apply this *ad-hoc* strategy:

(i) De-convolute mgATM read-outs into individual transcripts.

(ii) Apply ACMG/AMP evidences to individual transcripts.

(iii) Produce an overall PVS1_O (or BP7_O) code strength based on the relative contribution of individual transcripts/evidences to the overall expression (summarized in Supplementary **Figure S3A**).

For that reason, we have applied some ACMG/AMP codes not to individual variants but to individual mgATM transcripts, as described in this section. Recently, we have applied a similar approach to equally complex mgPALB2 read-outs [17].

To determine the pathogenic code strength applicable to individual transcripts carrying premature termination codons (PTCs) or large in-frame indels, we have followed ClinGen Sequence variant interpretation (SVI) working group recommendations for interpreting the loss-of-function PVS1 ACMG/AMP criteria [54]**,** incorporating specific *ATM* adaptations (e.g. defining biologically relevant *ATM* transcripts and mapping regions critical to protein function).

To determine the pathogenic code strength of full-length transcripts carrying missense variants, we have considered PS1, PM5, PP3 and BP4 evidences (see section on full-length transcripts).

As far as we know, only one functional ATM protein of 3056 aa (UniProtKB Q13315) has been described [55]. Ensembl (ATM ENSG00000149311) reports >50 *ATM* transcripts, annotating 15 of them as protein coding biotype. Of these, only two transcripts differing at the 5'UTR

(ENST00000675843.1 and ENST00000452508.6) are predicted to encode Q13315. Other protein coding transcripts encode very short ORFs (ranging from 93 to 168 amino acids) not expected to be biologically relevant.

For the purpose of variant classification, we propose that ENST00000675843.1 (flagged as MANE Select v0.95, Ensembl Canonical, matching RefSEq NM_000051.4), 63 exons, 62 coding exons and ORF of 3056 amino acids) is the only biologically relevant *ATM* transcript (i.e. none of the variants here investigated is located in an exon absent from biologically relevant transcripts)

1.1. PTC transcripts.

In relation with sequence alterations disrupting reading-frame, the SVI working group makes a clear distinction between transcripts predicted to undergo NMD (premature termination codon not occurring in the last exon or the last 50 bp of the penultimate exon) and those not predicted [54]. The *HBOPC_ATMv1* specifications acknowledge PTC_NMD for PTCs upstream of p.Leu2980 in exon 62.

The most 3' disrupting reading-frame alteration identified in this study, ∆(E52p11), introduces a PTC p.(Leu2544Asnfs*23) in exon 52. Consequently, we have annotated all PTC transcripts as PTC_NMD, and we have assigned a very strong pathogenic evidence to them (see **Supplementary Figure S3B**).

1.2. No-FS transcripts.

According to the ACMG/AMP rationale, pathogenic code strength for in-frame deletions is based on the relevance of the targeted region for protein function (critical or unknown) and/or on the relative size of the region (>10% or <10% of the protein removed). The *Spanish_ATM_WG* does not address this specific topic. The *HBOPC_ATMv1* specifications recommends pathogenic strong for <u>any</u> in-frame exon skipping targeting the α-Solenoid region, and pathogenic very strong for <u>any</u> in-frame exon skipping targeting the FATKIN domain (note that exon-level granularity is not considered). Of note, *HBOPC_ATMv1* says nothing about in-frame alterations caused by de novo/cryptic sites (partial exon losses). To reach exon-level granularity (and tackle partial exon losses), we have developed *ad-hoc* rules that combine generic ClinGen recommendations for in-

frame exon skipping alterations [54], with *HBOPC_ATMv1* specifications, structural and clinical evidence as follows:

(i) Pathogenic strong if in-frame exon skipping targeting the α-Solenoid region.

(ii) Pathogenic very strong if in-frame exon skipping targeting the FATKIN domain.

(iii) Pathogenic strong if targeting an *ATM* region in which missense (or in-frame) pathogenic variants have been reported.

(iv) Pathogenic strong if targeting a region with a relevant structural role.

(v) Pathogenic strong if removing >10% of a well-defined *ATM* domain (Spiral, Pincer, or FATKIN) [55].

(vi) Pathogenic very strong if two of the above are fitted.

For (iii), we have considered only the subgroup of *ATM* missense (or in-frame) variants reported in ClinVar as pathogenic/likely pathogenic by multiple submitters (no conflict) (www.ncbi.nlm.nih.gov/clinvar, last accessed 16/12/2021). For in-frame deletions targeting the flexible N-terminal (Spiral and Pincer) region, we have based (iv) on two recent Cryo-EM analyses producing very detailed structural models in Chaetomium thermophilum Tel1ATM and human ATM [56,57]. For in-frame deletions targeting the C-terminal FATKIN region, we have retrieved complementary information from previous Cryo-EM analyses performed in human *ATM* [58,59] and Sacharomyces cerevisiae Tel1/ATM [60,61].

For in-frame deletions not fitting any of the above instances (i-vi), we have considered pathogenic moderate (PM4 rationale) or pathogenic supporting (PP3 rationale). The rationale supporting the pathogenic strength applied to each in-frame alteration is as follows:

1.2.1. Δ(E5): exon 5 skipping (p.Arg111_Glu166delinsLys) removes 55 aminoacids, 27 of which are strictly conserved in vertebrates (**Supplementary Figure S9**), which represent ~5% of the Spiral domain, including residue Asp126 and HEAT repeats H3 and H4. The residue Asp126 is involved in stabilization and charge distribution at the Spiral interaction hub and p.Asp126Glu is a cancer-associated *ATM* variant [57]. Tel1/ATM structural analysis reveals that HEAT repeats H3 and H4

interact with HEAT repeats H17 and H18 (hydrophilic interactions), stabilizing the highly bent Spiral α-solenoid conformation [56]. ClinVar does not report missense pathogenic variants (supported by multiple submitters) in this region, but reports c.332-1G>A, detected in A-T patients [37] and leading to in-frame exon 5 skipping, as likely pathogenic (multiple submitters, no conflicts VCV000231535.4). The *HBOPC_ATMv1* specifications support pathogenic strong for ∆(E5) transcripts. Based on structural considerations and clinical data, we have upgraded to **pathogenic_very strong evidence**. Interestingly, NM_000051.3 (ATM):c.496G>C (p.Glu166Gln) is reported in ClinVar as pathogenic (late onset A-T disease) via exon 5 skipping, but only by one submitter (VCV000627570.2).

1.2.2. ∆(E7_E9): skipping of exons 7 to 9 (p.Gln222_Trp412del) deletes 191 residues, 58 of which are strictly conserved in vertebrates (**Supplementary Figure S9**). This isoform eliminates ~16% of the Spiral domain, including residues Arg248, Ser333, Arg337 and Val410 that are involved in stabilization and charge distribution at the Spiral interaction hub. Actually, Arg248Gln, Ser333Phe, Arg 337Cys, Arg337His, and Val410Ala are cancer-associated *ATM* missense variants Residues Ser333 and Arg337 are located in a Spiral solvent-exposed loop near the TAN motif, which may be an important interaction site. Val410 forms an internal hydrophobic contact with Ile339 predicted relevant for Spiral stability [57]. Ser367 is a functionally important radiation-induced auto-phosphorylation site [62]. ClinVar reports one pathogenic missense variant in this region (p.Pro292Leu, see previous section). *HBOPC_ATMv1* specifications support pathogenic strong for ∆(E7_E9) transcripts. Based on structural considerations and size (>10% of Spiral) supported by p.Pro292Leu findings, we have upgraded to **pathogenic_very strong evidence** to ∆(E7_9) transcripts.

1.2.3. ∆(E11): exon 11 skipping (p.Pro537_Ser601del) deletes 65 amino acids (6 conserved in vertebrates, **Supplementary Figure S9**) that represent ~5% of the Spiral domain, including a region in the center of the solenoid structure where the HEAT repeat pattern is broken and the loop between helices 24 and 25 (amino acids 549 to 570) extends to interact with FATKIN domain loops (residues 2084–2092 and 2107–2123) [57]. ClinVar does not report pathogenic missense variants in this region. Yet, given the predicted relevance of Spiral-FATKIN interactions in the overall ATM structure and function, we have upgraded to **pathogenic very strong evidence**.

1.2.4. Δ(E12): exon 12 skipping (p.Asn602_Cys633del) targets 32 residues that constitute ~3% of the Spiral domain, including three short helices (25 to 28) that break the HEAT repeat pattern in the center of the solenoid ring. Pro604Ser is a cancer-associated human *ATM* missense variant. Pro604 is involved in stabilization and charge distribution at Spiral interaction hub [57]. Exon 12 skipping targets as well the first three residues of the "plug" loop (residues 630–643) which bridges through the center of the solenoid. Cys633, together with nearby His635, His636 and Cys790, probably coordinate a Zinc atom [57]. ClinVar does not report missense pathogenic variants in the region, but reports c.1898+1G>T as pathogenic/likely pathogenic (multiple submitters, no conflicts, accession VCV000453391.5). This variant causes exon 12 skipping, and has been reported in *trans* with an *ATM* multi-exon gross deletion in a patient with ataxia-telangiectasia (AT) [63]. Based on structural considerations supported by c.1898+1G>T findings, we have upgraded to **pathogenic_very strong evidence**.

1.2.5. Δ(E15), Δ(E16), and **Δ(E15_E16):** exon 15 skipping (p.Ser751_Lys792del, 42 residues, 11 strictly conserved), exon 16 skipping (p.Lys793_Leu822del, 30 residues, 3 strictly conserved) and exons 15 to 16 skipping (p.Ser751_Leu822del, 72 residues, 14 strictly conserved) remove 6%, 3%, and 9%, respectively, of the Spiral domain. We have identified neither specific structural features in the target region nor missense pathogenic variants. *HBOPC_ATMv1* specifications support **pathogenic_strong evidence** for all three transcripts. Based on the lack of additional evidence (we have not identified structural, clinical or functional evidences specifically applicable to exons 15 and/or 16), we have not upgraded to pathogenic very strong evidence.

1.2.6. ∆(E16p3): the alteration targets only one residue (p.Lys793del) located at the Spiral domain (loop between Helix 39 and Helix 40). We have not identified any critical structural feature associated with this residue. Lys is an exposed positive-charged hydrophilic residue not involved in H-bonds. ClinVar does not report pathogenic missense variant targeting this residue. *HBOPC_ATMv1* specifications do not provide specific recommendations for this type of alterations. Based on PROVEAN predictions (score -3.3) only, we have assigned a **pathogenic_supporting evidence** to ∆(E16p3) transcripts (*as per* PP3 rationale).

1.2.7. Δ(E25_E26): skipping of exons 25 and 26 (p.Val1193_Gln1331del) removes 139 residues, 37 of which are conserved, **Supplementary Figure S9**), affecting ~19% of the Pincer domain, including Ala1309. This residue is involved in flexibility and folding and p.Ala1309Thr is a cancer-associated human *ATM* missense variant (Stakyte et al, 2021). Further, p.Leu1283Pro is a ClinVar likely pathogenic mutation (multiple submitters, no conflicts, accession VCV000181994.10) supported by clinical (observed with at least three different pathogenic *ATM* mutations in a compound heterozygous state in multiple individuals diagnosed with ataxia telangiectasia) and functional (no detectable ATM protein and absent kinase activity) evidence [38,64–67]. *HBOPC_ATMv1* specifications support pathogenic strong for ∆(E25-26) transcripts. Based on structural considerations and size supported by p.Leu1283Pro findings, we have upgraded to **pathogenic very strong evidence**.

1.2.8. ∆(E25p159): the alteration eliminates 53 residues (15 conserved in vertebrates, **Supplementary Figure S9**) and targets <10% of the Pincer domain (p.Val1193_Glu1245del). We have identified neither critical structural features in this region, nor ClinVar (likely) pathogenic variants. *HBOPC_ATMv1* specifications do not provide specific recommendations for this type of alterations. Due to the lack of structural, clinical or functional evidences, we have assigned a **pathogenic evidence with moderate strength** to ∆(E25p159) transcripts (as per PM4 rationale).

1.2.9. ∆(E26q120): this alteration deletes 40 amino acids (5 conserved in vertebrates, **Supplementary Figure S9**) that represent <10% of the Pincer domain (p.Val1292_Gln1331del). We have not identified critical structural features other than Ala1309, a residue involved in flexibility and folding (Ala1309Thr is a cancer-associated mutation). *HBOPC_ATMv1* specifications do not provide specific recommendations for this type of alterations. Based on the relative small size of the alteration, limited structural evidences and lack of clinical or functional data, we have assigned to ∆(E26q120) transcripts a **pathogenic evidence with moderate strength** (as per PM4 rationale).

1.2.10. ∆(E28_29): skipping of exons 28 and 29 (p.Asp1371_Arg1479del, 109 amino acids, 23 of which are strictly conserved, **Supplementary Figure S9**) removes ~15% of the Pincer domain, including Leu1420, involved in flexibility and folding (Leu1420Phe is a cancer associated missense mutation) (Stakyte et al, 2021). *HBOPC_ATMv1* specifications support pathogenic strong for ∆(E28-

29) transcripts. ClinVar does not report pathogenic missense variants in this region. Based on size considerations (>10% of the Pincer domain removed) supported by Leu1420 evidences, we have upgraded to ∆(E28_29) transcripts to **pathogenic very strong evidence**.

1.10.11. ∆(E51): exon 51 skipping (p.Arg2506_Asn2543del, 38 residues, 23 strictly conserved in vertebrates, **Supplementary Figure S9**) removes ~27% of the FAT-HRD region (3% of the FATKIN domain), including the residue Gln2522, that is involved in a PIKK-specific and highly conserved FATKIN polar interaction (Gln2522-Gln2730) [58]. Indeed, the interface between the HRD and the Kinase domain frequently mutates in human cancer and HRD is directly involved in regulating ATM kinase activity, including transition from an inactive to an active state [56]. Further, p.Ala2524Pro is a ClinVar pathogenic/likely pathogenic variant (multiple submitters, no conflicts, accession VCV000187613.5) reported in three individuals affected with ataxia-telangiectasia (A-T) [37,68,69]. In two of them, this variant occurred with a different pathogenic variant in *ATM*, while the third individual was homozygous. Experimental studies in patient-derived lymphoblast cell lines indicate that the protein is expressed at significantly reduced levels and has defective kinase activity [69,70]. *HBOPC_ATMv1* specifications support pathogenic very strong for ∆(E51) transcripts. We think that structural considerations supported by p.Ala2524Pro findings fully support **pathogenic very strong evidence** to ∆(E51) transcripts.

1.10.12. ∆(E52): exon 52 skipping (p.Leu2544_Glu2596del, 53 residues, 23 conserved in vertebrates, **Supplementary Figure S9**) targets a very large proportion (~38%) of the FAT-HRD domain (4% of the FATKIN domain). ClinVar does not report missense pathogenic variants in this region, but Arg2579 and Arg2580 are essential for ATM binding to MRE11-RAD50 [71]. *HBOPC_ATMv1* specifications support pathogenic very strong for ∆(E52) transcripts. We think that the well-documented critical role of the FAT-HRD region on ATM function, fully supports **pathogenic_very strong evidence** to ∆(E52) transcripts.

1.3. Full-length transcripts coding missense variants.

Our data set includes four *ATM* variants (c.496G>A, c.902G>A, c.1898G>T and c.3577G>C) producing detectable amounts of full-length transcripts carrying a missense change [p.(Glu166Lys), p.(Gly301Asp), p.(Cys633Phe), and p.(Val1193Leu), respectively]**.**

1.3.1. PS1 *(Same amino acid or splice change as a previously established pathogenic variant regardless of nucleotide change)***. PM5** *(Missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before)*

We have searched ClinVar for pathogenic/likely pathogenic missense changes at residues 166, 301, 633, and 1193, checking that: (i) the variant is reported as pathogenic/likely pathogenic by multiple submitters (no conflicts) and (ii), the variant has not been reported to alter splicing and is not predicted to alter splicing (i.e. bona-fide missense change). We have not identified any missense change fitting (i) and (ii), concluding that none of the four full-length transcripts supports a pathogenic strong (as per PS1) or moderate (as per PM5) evidence (ClinVar, last accessed 04/01/2022).

Note that ClinVar reports NM_000051.4(*ATM*):c.496G>C (p.Glu166Gln) as pathogenic, but this classification is supported by one submitter only and exon 5 skipping is the alleged pathogenic mechanism (RCV000852350.2). Consequently, this piece of data does not support PM5 for NM_000051.4(*ATM*):c.496G>A (p.Glu166Lys).

1.3.2. PP3 (*Multiple lines of computational evidence support a deleterious effect)***/BP4** *(Multiple lines of computational evidence suggest no impact): HBOPC_ATMv1* specifications support PP3 for REVEL meta-predictor score >.773 and BP4 for REVEL score <.249. We have complemented the analysis with BayesDel meta-predictor gene level *ATM* thresholds (<-0.180 for BP4, >0.216 for PP3) as recently published [72], and with structural damage predictions. We have run REVEL and BayesDel with the Ensembl tool Variant Effect Predictor (VEP) (www.ensembl.org/Tools/VEP). To predict structural damage, we have used Missense3D (http://missense3d.bc.ic.ac.uk/missense3d/) with input files UniProt ID Q13315 and PDB code 7ni5.

1.3.2.1. The full-length transcript r.496G>A codes the missense change p.Glu166Lys, reported as a variant of uncertain significance in ClinVar (accession VCV000825334.6). We have not identified any

critical structural feature associated with this change. Located at Helix 9 (solenoid domain), Glu166 is an exposed hydrophilic residue (RSA 32.4%) interacting (ionic interaction) with Lys129 side-chain (3.52Å). Missense 3D does not predict structural damage for the p.Glu166Lys change. REVEL (0.184) and BayesDel (-0.345) score support benign. Based on these considerations, we have assigned a **benign supporting evidence** to r.496G>A transcripts

1.3.2.2. The full-length transcript r.902G>A codes the missense change p.Gly301Asp, reported as variant of uncertain significance in ClinVar (VCV000188359.23). Located at Helix 15 (solenoid domain), Gly is a buried uncharged residue (RSA 2.3%) interacting (H-bond) with Arg248 (3.36 Å). REVEL score for p.Gly301Asp (0.418) is inconclusive, but BayesDel no AF score (0.244) supports damaging. Further, Missense 3D predicts structural damage (buried uncharged Gly replaced by a charged Asp residue with disallowed phi/psi angles). Based on these predictions, we have assigned a **pathogenic supporting evidence** to r.902G>A transcripts.

1.3.2.3. The full-length transcript r.1898G>U codes the missense change p.Cys633Phe not reported in ClinVar (last accessed 21/12/2021). REVEL score (0.274) predicts a benign change. BayesDel score is inconclusive (-0.021). Missense3D predicts structural damage (the substitution leads to the expansion of cavity volume by 342.144 $Å^3$). Located in a loop connecting Helix 33 and Helix 34 (solenoid domain), p.Cys633 is a exposed (RSA 13.3%) uncharged residue interacting (H-bond) with His636 (2.58Å), Lys792 (3.11 Å), and a Zn atom (metal coordination, as part of a Zn-binding motif involving also His635, His636, and Cys790). In brief, despite REVEL and BayesDel scores not predicting damaging, we think that Missense3D output supported by structural considerations [57] is sufficient to assign a **pathogenic supporting evidence** to r.1898G>U transcripts.

1.3.2.4. The full-length transcript r.3577G>C encodes the missense change p.Val1193Leu, not reported in ClinVar (last accessed 21/12/2021). REVEL score (0.296) predicts a benign change. BayesDel score (-0.0065) is inconclusive. Missense3D does not detect structural damage. Located in Helix 59, p.Val1193 is a buried uncharged residue (RSA 3.5%) not involved in H-bonds(s). Based on REVEL score and Missense3D predictions, we have assigned a **benign supporting evidence** to r.3577G>C transcripts.

2. PS1 *(Same amino acid change as a previously established pathogenic variant regardless of nucleotide change)*

Since missense variants here investigated have a complete or partial effect on splicing, we have considered the PS1 evidence as part of the PVS1_O/BP7_O scheme to annotate full-length transcripts (see section 1.3.1).

3. PS4 (case-control association with BC disease).

HBOPC_ATMv1 specifications recommends PS4 for breast cancer case-control studies with OR ≥2, p≤.05, and lower 95% CI. The *Spanish_ATM_WG* specifications recommend PS4 if 4-15 AT probands have been identified, but does not specify any guidance for breast cancer case-control studies. We have used BRIDGES data (60,466 women with breast cancer and 53,561 matched controls) [6] to assess association with breast cancer disease (i.e. PS4 not applicable to any of the 56 *ATM* variants under investigation).

4. PM2.

The original ACMG/AMP guidelines defined the PM2 rarity code (rarity evidence) as absent from controls (ExAC, ESP, and/or 1000 genomes project). However, the availability of even larger control datasets (e.g. gnomAD) does not rule out of finding pathogenic alleles within them. We have followed *HBOPC_ATMv1* specifications (*Spanish_ATM_WG* are slightly different, as indicated):

(i) Using PM2 if variant absent, or frequency ≤.001% (≤.002% according to *Spanish_ATM_WG*) in any sub-population in which the variant has been detected.

(ii) Decreasing PM2 evidence strength to Supporting. The latter is recommended by generic ClinGen Sequence Variant Interpretation Recommendations for PM2 (Version 1.0 at https://clinicalgenome.org/working-groups/sequence-variant-interpretation/), and by *HBOPC_ATMv1* specifications. By contrast, *Spanish_ATM_WG* recommend sticking to the original moderate strength.

(iii) Not considered a conflicting piece of evidence for variants that otherwise are likely benign/benign (this particular rule, not acknowledge by the *Spanish_ATM_WG,* has no impact in our dataset).

For allele counting, we have interrogated gnomADv2.1 (global). For *ATM* variants absent from gnomADv2.1 (no counts), the actual number of interrogated alleles (allele number) was determined using as a proxy data on the closest available SNP (in all cases, ≤5nt apart from the variant of interest). On average, 96% of samples have a >20x coverage at the region of interest. See further details in **Supplementary Table S3.3**.

5. PM3.

ATM bi-allelic mutations cause Ataxia telangiectasia (A-T), a rare autosomal recessive disorder with a complex phenotype, including cerebellar degeneration accompanied by ataxia (movement dysfunction), immunodeficiency, thymic and gonadal atrophy, radiation sensitivity and predisposition to cancer [55]. We have applied PM3 with code strength as per ClinVar recommendations (ClinGen Sequence Variant Interpretation Recommendation for in trans Criterion (PM3) - Version 1.0 Working Group Page: https://clinicalgenome.org/working-groups/sequence-variant-interpretation/Date Approved: May 2, 2019).

Points extracted from the following references; c.332-1G>A [37]); c.496+5G>A [36,73]; c.902-1G>T [38]; c.1066-6T>G [74]; c.1898+2T>G [40,75–77]; c.2376+1G>A [78]; c.3993+1G>A [37,79]; c.4110-9C>G [42]; c.7307+1G>A [80]; c.7630-2A>C [43,66,81,82]; c.7788G>A [44,83]. See **Supplementary Table S3.1** for further details.

6. PM4 *(Protein length changes due to in-frame deletions/insertions in a non-repeat region or stop-loss variants).*

HBOPC_ATMv1 specifications recommends not using this rule for in-frame deletions/insertions not already PVS1-eligible. *Spanish_ATM_WG* specifications recommend using PM4 only for alterations targeting at least one residue in a critical functional region. We have not applied PM4 evidence to any individual variant, but we have used the rational to annotate some in-frame transcripts present in complex mgATM read-outs. See section 1.2 for further details.

7. PM5 *(Missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before)*

Since missense variants here investigated have a complete or partial effect on splicing, we have considered the PM5 evidence as part of the PVS1_O/BP7_O scheme to annotate full-length transcripts (see section 1.3.1).

8. BA1 *(GnomAD Filtering Allele Frequency)***.**

The generic SVI recommendation sets the threshold for the stand-alone benign at MAF>5% [84]. Yet, using the maximum credible frequency rationale [85], BA1 MAF threshold can be reduced for specific genes. Using this approach, the *HBOPC_ATMv1* specifications set stand-alone benign criteria for variants with allele frequency >0.5%. The *Spanish_ATM_WG* proposes cut-off of MAF >0.5% in any of the individual NFE, AFR, LAT, EAS, and SAS GnomADv2.1 (non-cancer) sub-populations. None of the variants investigated reaches *HBOPC_ATMv1* or *Spanish_ATM_WG* thresholds, although c.1066-6T>G is close, reaching 0.23% in the NFE sub-population.

9. BS1. *(GnomAD Filtering Allele Frequency greater than expected for disease)*

Both the *HBOPC_ATMv1* specifications and the *Spanish_ATM_WG* support strong benign criteria for variants with allele frequency >0.05%. Two variants only (c.1066-6T>G and c.3993+5G>T) reach this cutoff (see **Supplemental Table S3.1**).

10. BP2 *(Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder; or observed in cis with a pathogenic variant in any inheritance pattern)*

We have not identified any evidence in the scientific literature or ClinVar (last accessed 03/01/2022) supporting BP2 evidence (co-ocurrence in trans with a pathogenic or likely pathogenic variant *ATM* variant in an AT-unaffected individual) for any of the 56 variants under investigation.

11. BP7 *(A synonymous (silent) variant)*

The *HBOPC_ATMv1* specifications support the use of BP7 also for deep intronic variants (defined as +8 to -41 variants). The *Spanish_ATM_WG* restricts its use for synonymous variant at nucleotides not highly conserved (PhyloP score <6.66).

Our dataset does not include deep intronic variants, but includes four synonymous variants: c.903T>G p.(Gly301=), c.3993G>A p.(Gln1331=), c.7515G>Ap.(Lys2505=), and c.7788G>A p.(Glu2596=). Yet, mgATM analysis shows that all four synonymous are spliceogenic, so that none of them qualifies for BP7.

12. Predictive Splicing Evidences

Both *HBOPC_ATMv1* specifications and the *Spanish_ATM_WG* acknowledge the use of *predictive* splicing codes PVS1 (GT-AG splice site variants) and PP3/BP4 (other variants) to aid in the classification of *ATM* variants. Yet, neither *HBOPC_ATMv1* nor *Spanish_ATM_WG* specify rules for combining predictive and functional splicing evidences. In our opinion, once *functional* data is available (e.g. mgATM read-outs), predictive evidences should not contribute to variant classification, but being overridden by functional splicing evidence PVS1_O or BP7_O. Otherwise, internal inconsistences would arise in the ACMG/AMP classification system (e.g. +1 and +5 variants with identical splicing impact will score differently). Furthermore, the ACMG/AMP system implicitly assumes that each piece of evidence contributing to the final classification is independent [25], an assumption hardly met by predictive and functional splicing codes, as most splicing analyses (including our mgATM analyses) are performed in bioinformatically pre-selected variants. These issues have been extensively discussed elsewhere [15].

Said that, we have used predictive splicing codes for comparative purposes (i.e. to perform an alternative classification in absence of mgATM data, see **Supplementary Table S3.5 and S6**)

12.1. PVS1 *(Null GT-AG splice site variant in a gene where loss of function is a known mechanism of disease)*

To assign specific PVS1 code strengths to 19 GT-AG variants (and 3 variants located at the 3' end of exon), we have followed *HBOPC_ATMv1* specifications (displayed in **Supplementary Table S3.5**).

12.2. PP3/BP4 *(Multiple lines of splicing computational evidence support deleterious / no deleterious).*

To assign PP3/BP4 predictive codes (splicing) to non-GT-AG variants, the *HBOPC_ATMv1* does not provides any specific guidance other than concordance of ≥2 predictors. For the purpose of the present study, we have used SpliceAI [45] scores (**Supplementary Table S7**) supported by

MES (**Supplementary Table S1**). SpliceAI scores were calculated at https://spliceailookup.broadinstitute.org/ using the following setting (hg38, score type raw, and max distance ±4999 nucleotides). We have assigned a pathogenic supporting evidence PP3 if Splice AI score ≥0.2 (high recall), and benign supporting evidence BP4 if scores <0.1. For variant displaying SpliceAI scores in the 0.10-019 range (regardless of MES scores) we have considered that neither PP3 nor BP4 are applicable.

Pathogenic ACMG/AMP evidences not applicable to *ATM* variants.

PS2 *(de novo (paternity confirmed) in a patient with the disease and no family history)***.**

HBOPC_ATMv1 specifications recommend not using this ACMG-AMP evidenced for *ATM* variants (neither for autosomal dominant and breast cancer, nor four autosomal recessive and A-T). The Spanish_ATM_WG suggest that PS2 evidence might be applicable in the autosomal recessive setting. At any rate, BRIDGES data does not provide confirmed evidence of *de novo* findings. Further, we have not identified in the scientific literature or ClinVar (last accessed 04/01/2022) evidences in A-T patients supporting PS2 evidence for any of the 56 variants under investigation.

PM1 *(Located in a mutational hot spot and/or critical and well-established functional domain)*.

At present, we think that there is not enough data supporting the use of PM1 for the classification of *ATM* missense variants (i.e. neither evidence of mutational hot spots, nor well-established functional domain without benign variants). The HBOPC_ATMv1 specifications recommend not using this ACMG-AMP evidenced for *ATM* variants. The Spanish_ATM_WG proposes that PM1 is applicable only to variants targeting residues p.S1981 (auto-phosphorylation codon) and p.Arg3008 (mutational hotspot). None of our alterations targets these residues.

PM6 *(Confirmed de novo without confirmation of paternity and maternity).*

The HBOPC_ATMv1 specifications recommend not using this ACMG-AMP evidenced for *ATM* variants (neither for autosomal dominant and breast cancer, nor four autosomal recessive and A-T). The Spanish_ATM_WG proposes that PM6 evidence might be applicable in the autosomal recessive setting. At any rate, we have not identified in the scientific literature or ClinVar (last

accessed 03/01/2022) evidences in A-T patients supporting PM6 evidence for any of the 56 variants under investigation.

PP1 *(co-segregation with disease in multiple affected family).*

The HBOPC_ATMv1 specifications recommend not using this ACMG-AMP evidenced for *ATM* variants (neither for autosomal dominant and breast cancer, nor four autosomal recessive and A-T). The Spanish_ATM_WG suggests that PP1 might be applicable to A-T families (co-segregation with AT in multiple affected family members with 3-4 meiosis observed). At any rate, we have not identified in the scientific literature or ClinVar (last accessed 03/01/2022) evidence in A-T patients supporting PP1 evidence for any of the 56 variants under investigation.

PP2 *(missense variant in a gene that has a low rate of benign missense variation and where missense variants are a common mechanism of disease).*

Neither the HBOPC_ATMv1 specifications nor the Spanish_ATM_WG recommend using PP2 for *ATM* variants. Current data do not support a low rate of *ATM* benign missense variants, nor *ATM* missense changes as a common mechanism of disease.

PP4 *(Phenotype specific for disease with single genetic etiology)*

Both the HBOPC_ATMv1 specifications and the Spanish_ATM_WG recommend not using this ACMG-AMP evidenced for *ATM* variants.

PP5: The reputable source evidence has been discontinued by the ClinGen Sequence Variant Interpretation Working Group [86].

Benign ACMG/AMP evidences not applicable to *ATM* variants.

BS2 *(Observed in a healthy adult individual for a dominant (heterozygous) disorder with full penetrance expected at an early age).*

The HBOPC_ATMv1 specifications recommend not using BS2 for *ATM* variants, since *ATM* has incomplete BC penetrance [6]. The *Spanish_ATM_WG* supports that BS2 might be applicable to variants observed in the homozygous state in an A-T unaffected individual. At any rate, we have not identified any evidence in the scientific literature or ClinVar (last accessed 03/01/2022) supporting BS2 for any of the 56 variants under investigation.

BS4 *(Lack of segregation in affected members of a family)*.

HBOPC_ATMv1 specifications recommend not using BS4 for *ATM* variants (neither for autosomal dominant and breast cancer, nor four autosomal recessive and A-T). The *Spanish_ATM_WG* indicates that BS4 might be applicable to A-T families (lack of segregation in affected member of two or more AT families). At any rate, we have not identified in the scientific literature or ClinVar (last accessed 03/01/2022) evidences in A-T patients supporting BS4 for any of the 56 variants under investigation.

BP1 *(Missense variant in gene where only LOF causes disease).*

Both *HBOPC_ATMv1* and *Spanish_ATM_WG* specifications recommend not using BP1 for *ATM* variants (various *ATM* missense pathogenic variants have been described)

BP3 *(In-frame deletions/insertions in a repetitive region without a known function).*

Both *HBOPC_ATMv1* and *Spanish_ATM_WG* specifications recommend not using BP3 for *ATM* variants. Indeed, no repetitive region (without function) has been reported in ATM [55].

BP5 *(Variant found in a case with an alternate molecular basis).*

Both *HBOPC_ATMv1* and *Spanish_ATM_WG* specifications recommend not using BP5 for *ATM* variants. We fully agree, as several BC cases with multiple pathogenic variants have been described.

BP6 *(Reputable source recently reports variant as benign but the evidence is not available to the laboratory to perform an independent evaluation).*

Both *HBOPC_ATMv1* and *Spanish_ATM_WG* specifications acknowledge that the ClinGen Sequence Variant Interpretation Working Group has discontinued this benign evidence [86].

Supplementary Table S1. Bioinformatics analysis of 381 BRIDGES *ATM* variants with Max Ent Scan.

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.-31+4G>C	IVS1/EX1	4.2	6.61	57.4			
c.-31+6G>T	IVS1/EX1	4.2	7.23	72.1	-1.47	6.17	+519.73 2 nt downstream
c.-31+6G>C	IVS1/EX1	4.2	5.59	33.1			
c.-31+8C>T	IVS1/EX1	4.2	-	-			
c.-31+10C>T	IVS1/EX1	4.2	-	-			
c.-31+14G>A	IVS1/EX1	4.2	-	-			
c.-31+19C>T	IVS1/EX1	4.2	-	-			
c.72+8del	IVS2/EX2	5.69	-	-			
c.74A>G	EX3	9.72	9.45	-2.8			
c.184A>G	EX3	4.57	1.95	**-57.3**			
c.185+8A>G	IVS3/EX3	4.57	-	-			
c.332-9T>C	IVS4/EX5	7.19	6.87	-4.5			
c.332-5A>G	IVS4/EX5	7.19	1.58	**-78.0**			
c.332-1G>A	IVS4/EX5	7.19	-1.58	**-121.6**			
c.496G>A	EX5	7.76	4.58	**-41.0**			
c.496+4T>C	IVS5/EX5	7.76	8.17	5.3			
c.496+5G>A	IVS5/EX5	7.76	4.89	**-37.0**			
c.496+8T>C	IVS5/EX5	7.76	-	-			
c.497-7T>C	IVS5/EX6	7.84	7.54	-3.8			
c.497-4T>A	IVS5/EX6	7.84	6.83	-13.0			
c.497-3A>T	IVS5/EX6	7.84	10.29	31.3			
c.901G>T	EX7	7.08	-3.6	**-150.9**			
c.901+2T>C	IVS7/EX7	7.08	-0.66	**-109.3**			
c.901+3A>T	IVS7/EX7	7.08	1.88	**-73.5**			
c.902-7A>G	IVS7/EX8	7.63	8.57	12.3			
c.902-1G>T	IVS7/EX8	7.63	-0.96	**-112.6**			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.902G>A (1st exon-nt)	EX8	7.63	6.9	-9.6			
c.903T>G	EX8	7.63	6	**-21.4**			
c.903T>C	EX8	7.63	6.13	-19.7			
c.903T>A	EX8	7.63	6.16	-19.3			
c.1065+1G>T	IVS8/EX8	8.68	0.18	**-97.9**			
c.1065+3A>G	IVS8/EX8	8.68	4.95	**-43.0**			
c.1065+4C>T	IVS8/EX8	8.68	8.73	0.6			
c.1066-6T>G	IVS8/EX9	10.82	8.33	**-23.0**			
c.1066-3C>T	IVS8/EX9	10.82	10.35	-4.3			
c.1234T>C	EX9	4.05	6.18	52.6			
c.1235+4_1235+5del	IVS9/EX9	4.05	-1.32	**-132.6**			
c.1235+5A>G	IVS9/EX9	4.05	8.92	120.3			
c.1235+8G>A	IVS9/EX9	4.05	-	-			
c.1236-10T>C	IVS9/EX10	12.07	12.12	0.4			
c.1236-8T>C	IVS9/EX10	12.07	12.45	3.2	1.82	3.07	+268.68 1 nt downstream
c.1236-3T>C	IVS9/EX10	12.07	11.78	-2.4			
c.1607G>C	EX10	11.01	10.44	-5.2			
c.1608-10T>C	IVS10/EX11	9.78	9.29	-5.0			
c.1608-9T>A	IVS10/EX11	9.78	8.97	-8.3			
c.1608-8T>C	IVS10/EX11	9.78	9.77	-0.1			
c.1608-4A>G	IVS10/EX11	9.78	9.61	-1.7			
c.1608-3T>C	IVS10/EX11	9.78	10.51	7.5			
c.1609C>T	EX11	9.78	9.96	1.5			
c.1802+7T>G	IVS11/EX11	8.55	-	-			
c.1803-7T>C	IVS11/EX12	8.85	9.14	3.3			
c.1803-6T>C	IVS11/EX12	8.85	9.04	2.2			
c.1897T>C	EX12	8.35	9.94	19.0			
c.1898G>T	EX12	8.35	4.34	**-48.0**			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.1898+3_1898+4del	IVS12/EX12	8.35	-5.14	**-161.6**			
c.1898+2T>G	IVS12/EX12	8.35	0.7	**-91.6**			
c.1898+3A>G	IVS12/EX12	8.35	4.84	**-42.0**			
c.1898+3A>T	IVS12/EX12	8.35	4.88	**-41.6**			
c.1898+9T>C	IVS12/EX12	8.35	-	-			
c.1899-10T>G	IVS12/EX13	No	-	-			
c.1899-7C>G	IVS12/EX13	No	-	-			
c.1900G>A	EX13	No	-	-			
c.2124+12dup	IVS13/EX13	7.66	-	-			
c.2124+7T>A	IVS13/EX13	7.66	-	-			
c.2124+9T>A	IVS13/EX13	7.66	-	-			
c.2125-4dup	IVS13/EX14	5.66	7.27	28.5			
c.2125-6T>C	IVS13/EX14	5.66	4.78	-15.6			
c.2125A>G	EX14	5.66	7.41	30.9			
c.2251-4A>G	IVS14/EX15	6.62	3.96	**-40.2**	-1.41	7.33	+619.86 3 nt upstream
c.2251-1G>C	IVS14/EX15	6.62	-1.43	**-121.6**	3.92	4.49	+14.54 20 nt downstream
c.2376+1G>A	IVS15/EX15	10.57	2.39	**-77.4**			
c.2376+3A>G	IVS15/EX15	10.57	8.68	-17.9			
c.2376+3A>T	IVS15/EX15	10.57	6.12	**-42.1**			
c.2376+6A>C	IVS15/EX15	10.57	10.22	-3.3			
c.2376+15_2376+17del	IVS15/EX15	10.57	-	-			
c.2377-7A>C	IVS15/EX16	8.06	9.42	16.9			
c.2377-6T>A	IVS15/EX16	8.06	6.11	**-24.2**			
c.2377-5T>G	IVS15/EX16	8.06	5.49	**-31.9**			
c.2377-2A>G	IVS15/EX16	8.06	0.11	**-98.6**			
c.2377A>G	EX16	8.06	8.71	8.1			
c.2465T>A	EX16	7.79	8.07	3.6			
c.2466A>G	EX16	7.79	10.47	34.4			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.2466+8_2466+14del	IVS16/EX16	7.79	-	-	2.72	4.58	+268.38 2 nt downstream
c.2466+7A>G	IVS16/EX16	7.79	-	-			
c.2466+8T>C	IVS16/EX16	7.79	-	-	2.72	3.13	+215.07 2 nt downstream
c.2466+9dup	IVS16/EX16	7.79	-	-			
c.2466+8T>A	IVS16/EX16	7.79	-	-	2.72	5.66	+308.09 2 nt downstream
c.2466+10_2466+13del	IVS16/EX16	7.79	-	-			
c.2467-7del	IVS16/EX17	8.81	8.14	-7.6			
c.2467-8C>T	IVS16/EX17	8.81	8.78	-0.3			
c.2467-7C>T	IVS16/EX17	8.81	8.68	-1.5			
c.2467-3A>G	IVS16/EX17	8.81	0.94	**-89.3**			
c.2638+3A>G	IVS17/EX17	6.62	3.81	**-42.5**			
c.2638+7A>G	IVS17/EX17	6.62	-	-			
c.2638+8C>T	IVS17/EX17	6.62	-	-			
c.2639-3T>C	IVS17/EX18	9.9	10.14	2.4			
c.2639G>T	EX18	9.9	7.14	**-27.9**	4.05	4.59	+13.33 19 nt downstream
c.2639G>A	EX18	9.9	8.6	-13.1			
c.2838+2T>C	IVS18/EX18	10.13	2.37	**-76.6**			
c.2838+4A>T	IVS18/EX18	10.13	4.84	**-52.2**			
c.2838+9C>A	IVS18/EX18	10.13	-	-			
c.2838+9C>G	IVS18/EX18	10.13	-	-			
c.2838+9C>T	IVS18/EX18	10.13	-	-			
c.2839-7A>G	IVS18/EX19	8.07	8.04	-0.4			
c.2839-4T>C	IVS18/EX19	8.07	7.87	-2.5			
c.2839-2A>T	IVS18/EX19	8.07	-0.29	**-103.6**	3.73	5.61	+50.4 19 nt downstream
c.2921C>T	EX19	6.29	4.37	**-30.5**			
c.2921C>G	EX19	6.29	9.48	50.7			
c.2922-8T>C	IVS19/EX20	8.55	8.98	5.0			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.2922-1G>A	IVS19/EX20	8.55	-0.19	**-102.2**			
c.2922C>T	EX20	8.55	7.86	-8.1			
c.3077G>A	EX20	7.03	2.18	**-69.0**			
c.3077+3A>C	IVS20/EX20	7.03	0.79	**-88.8**	4.67	3.7	-20.77 2 nt downstream
c.3077+4G>A	IVS20/EX20	7.03	10.47	48.9			
c.3078-10T>G	IVS20/EX21	4.4	3.12	**-29.1**			
c.3078-4T>C	IVS20/EX21	4.4	3.83	-13.0			
c.3078-1G>A	IVS20/EX21	4.4	-4.34	**-198.6**			
c.3078G>T	EX21	4.4	2.88	**-34.6**			
c.3153G>T	EX21	10.03	7.77	**-22.5**			
c.3153+4A>G	IVS21/EX21	10.03	7.07	**-29.5**			
c.3154-9T>G	IVS21/EX22	8.27	7.35	-11.1			
c.3154-7C>T	IVS21/EX22	8.27	7.85	-5.1			
c.3154-7C>A	IVS21/EX22	8.27	6.32	**-23.6**			
c.3154-6C>T	IVS21/EX22	8.27	6.28	**-24.1**			
c.3154-5C>T	IVS21/EX22	8.27	7.47	-9.7			
c.3154-4G>A	IVS21/EX22	8.27	8.25	-0.2			
c.3154-4G>T	IVS21/EX22	8.27	7.5	-9.3			
c.3284G>C	EX22	7.57	0.76	**-90.0**			
c.3284G>A	EX22	7.57	-0.57	**-107.5**			
c.3284+1G>A	IVS22/EX22	7.57	-0.6	**-107.9**			
c.3284+4del	IVS22/EX22	7.57	7.37	-2.6			
c.3284+4A>G	IVS22/EX22	7.57	3.11	**-58.9**			
c.3284+8dup	IVS22/EX22	7.57	6.49	6.5			
c.3284+7G>T	IVS22/EX22	7.57	-	-			
c.3284+9T>C	IVS22/EX22	7.57	-	-			
c.3285-10T>A	IVS22/EX23	7.43	6.22	-16.3			
c.3286T>C	EX23	7.43	5.51	**-25.8**			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.3402+4A>G	IVS23/EX23	6.49	2.4	**-63.0**			
c.3402+5T>C	IVS23/EX23	6.49	6.06	-6.6			
c.3402+9A>G	IVS23/EX23	6.49	-	-			
c.3403-10T>C	IVS23/EX24	7.9	7.38	-6.6			
c.3403-9C>T	IVS23/EX24	7.9	6.9	-12.7			
c.3403-3A>C	IVS23/EX24	7.9	10.39	31.5			
c.3576G>A	EX24	7.84	1.45	**-81.2**			
c.3576+8T>C	IVS24/EX24	7.84	-	-			
c.3577-10G>A	IVS24/EX25	8.52	8.01	-6.0			
c.3577-6G>A	IVS24/EX25	8.52	8.76	2.8			
c.3577-1G>C	IVS24/EX25	8.52	0.45	**-94.7**	-1.46	4.69	+421.23 8 nt downstream
c.3577-1G>A	IVS24/EX25	8.52	-0.22	**-102.6**	-4.32	3.62	+183.8 1 nt downstream
c.3577G>C (1st exon-nt)	EX25	8.52	7.35	-13.7			
c.3746+1G>A	IVS25/EX25	9.66	1.48	**-84.7**			
c.3746+4A>C (+4A)	IVS25/EX25	9.66	8.17	-15.4			
c.3746+5G>A	IVS25/EX25	9.66	6.62	**-31.5**			
c.3746+7T>G	IVS25/EX25	9.66	-	-			
c.3747-9T>C	IVS25/EX26	9.94	10.75	8.2			
c.3993G>A	EX26	9.99	4.99	**-50.1**			
c.3993+1G>A	IVS26/EX26	9.99	1.8	**-82.0**			
c.3993+5G>T (+5G)	IVS26/EX26	9.99	8.35	-16.4			
c.3993+6G>T	IVS26/EX26	9.99	9.8	-1.9			
c.3993+8T>C	IVS26/EX26	9.99	-	-			
c.3994-7C>T	IVS26/EX27	11.38	11.11	-2.4			
c.3994-6T>C	IVS26/EX27	11.38	10.96	-3.7			
c.3994-4G>T	IVS26/EX27	11.38	11.23	-1.3			
c.3994-3C>T (-3C)	IVS26/EX27	11.38	9.89	-13.1			
c.3994-2A>G	IVS26/EX27	11.38	3.43	**-69.9**			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.3995T>G	EX27	11.38	10.37	-8.9			
c.4109+1G>T	IVS27/EX27	8.34	-0.16	**-101.9**			
c.4109+3A>G	IVS27/EX27	8.34	4.51	**-45.9**			
c.4109+4T>C	IVS27/EX27	8.34	9.15	9.7			
c.4109+5G>A	IVS27/EX27	8.34	3.89	**-53.4**			
c.4109+6T>G (+6T)	IVS27/EX27	8.34	7	-16.1			
c.4109+6T>C	IVS27/EX27	8.34	7.47	-10.4			
c.4109+9A>C	IVS27/EX27	8.34	-	-			
c.4110-9C>G	IVS27/EX28	5.58	-0.09	**-101.6**	-2.47	5.58	+325.91 8 nt upstream
c.4110-9C>T	IVS27/EX28	5.58	5.73	2.7			
c.4110-2A>C	IVS27/EX28	5.58	-2.45	**-143.9**			
c.4110-2A>G	IVS27/EX28	5.58	-2.37	**-142.5**			
c.4111G>C	EX28	5.58	4.96	-11.1			
c.4235C>T	EX28	7.52	8	6.4			
c.4236+1G>A	IVS28/EX28	7.52	-0.65	**-108.6**			
c.4236+5G>A	IVS28/EX28	7.52	-0.63	**-108.4**			
c.4236+6T>C	IVS28/EX28	7.52	3.69	**-50.9**			
c.4236+9A>T	IVS28/EX28	7.52	-	-			
c.4237-10T>C	IVS28/EX29	10.5	10.22	-2.7			
c.4238A>T	EX29	10.5	10.46	-0.4			
c.4436+4A>G (+4A)	IVS29/EX29	8.88	7.24	-18.5			
c.4436+7A>G	IVS29/EX29	8.88	-	-			
c.4436+8A>G	IVS29/EX29	8.88	-	-			
c.4437-10C>T	IVS29/EX30	9.92	10.75	8.4			
c.4437-7A>G	IVS29/EX30	9.92	9.8	-1.2			
c.4437G>T	EX30	9.92	8.03	-19.1			
c.4442_4444del	EX30	9.92	-	-			
c.4612-7_4612del	IVS30/EX31	8.27	-10.59	**-228.1**			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.4612G>T	EX31	8.27	5.22	**-36.9**			
c.4776G>A	EX31	7.96	0.4	**-95.0**			
c.4776+2T>C	IVS31/EX31	7.96	0.2	**-97.5**			
c.4909+3G>A	IVS32/EX32	-2.25	3.46	253.8			
c.4909+4C>T	IVS32/EX32	-2.25	-	-			
c.4910-8A>G	IVS32/EX33	5.52	5.74	4.0			
c.4910-7T>C	IVS32/EX33	5.52	5.41	-2.0			
c.5005+5A>G	IVS33/EX33	5.88	8.89	51.2			
c.5005+7_5005+8del	IVS33/EX33	5.88	-	-			
c.5005+7T>C	IVS33/EX33	5.88	-	-			
c.5006-8A>G	IVS33/EX34	No	-	-			
c.5006-5T>C	IVS33/EX34	No	-	-			
c.5176T>G	EX34	9.27	8.73	-5.8			
c.5177+3G>A	IVS34/EX34	9.27	10.47	12.9			
c.5177+4A>C	IVS34/EX34	9.27	6.18	**-33.3**			
c.5319G>A	EX35	0.9	-	-			
c.5319+1G>T	IVS35/EX35	0.9	-	-			
c.5319+2T>C	IVS35/EX35	0.9	-	-			
c.5319+5C>G	IVS35/EX35	0.9	6.43	814.4			
c.5319+6_5319+7del	IVS35/EX35	0.9	-	-			
c.5320-10T>C	IVS35/EX36	8.6	9.78	13.7			
c.5320-6T>C	IVS35/EX36	8.6	8.66	0.7			
c.5320-5_5320-2del	IVS35/EX36	8.6	0.04	**-99.5**	2.8	8	+385.71 8 nt downstream
c.5320-3T>C	IVS35/EX36	8.6	8.77	2.0			
c.5320T>C	EX36	8.6	8.61	0.1			
c.5496+3A>G	IVS36/EX36	6.42	3.22	**-49.8**			
c.5496+8G>A	IVS36/EX36	6.42	-	-			
c.5496+8G>T	IVS36/EX36	6.42	-	-			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.5497-4C>T	IVS36/EX37	10.05	10.21	1.6			
c.5497-2A>G	IVS36/EX37	10.05	2.09	**-79.2**			
c.5497-2A>C	IVS36/EX37	10.05	2	**-80.1**			
c.5675-4T>A	IVS37/EX38	4.9	5.59	14.1			
c.5762+6G>A	IVS38/EX38	8.73	7.96	-8.8			
c.5763-8T>C	IVS38/EX39	8.8	9.22	4.8			
c.5763-7T>C	IVS38/EX39	8.8	8.99	2.2			
c.5763A>G	EX39	8.8	9.9	12.5			
c.5764C>T	EX39	8.8	9.94	13.0			
c.5918G>A	EX39	8.99	1.16	**-87.1**			
c.5918+8A>T	IVS39/EX39	8.99	-	-	-4.33	3.85	+188.91 4 nt downstream
c.5919-8A>G	IVS39/EX40	3.03	2.81	-7.3			
c.6006+8T>C	IVS40/EX40	8.76	-	-			
c.6007-10A>G	IVS40/EX41	10.44	10.84	3.8			
c.6007-7T>G	IVS40/EX41	10.44	9.37	-10.3			
c.6007G>A	EX41	10.44	9.5	-9.0			
c.6095G>A	EX41	6.62	1.48	**-77.6**			
c.6095+4A>G	IVS41/EX41	6.62	3.95	**-40.3**			
c.6095+5A>G	IVS41/EX41	6.62	9.66	45.9			
c.6095+6T>C	IVS41/EX41	6.62	5.88	-11.2			
c.6095+8G>T	IVS41/EX41	6.62	-	-			
c.6096-3T>C	IVS41/EX42	8.44	9.49	12.4			
c.6096-2A>G	IVS41/EX42	8.44	0.48	**-94.3**			
c.6096A>G	EX42	8.44	9.47	12.2			
c.6198+1G>A	IVS42/EX42	7.79	-0.38	**-104.9**			
c.6199-10T>C	IVS42/EX43	4.27	4.15	-2.8			
c.6199-6G>A	IVS42/EX43	4.27	4.36	2.1			
c.6200C>T	EX43	4.27	5.52	29.3			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.6200C>A	EX43	4.27	5.12	19.9			
c.6347+4A>G	IVS43/EX43	10.77	9.79	-9.1			
c.6348-10T>A	IVS43/EX44	7.92	5.33	**-32.7**			
c.6348-6_6348-5del	IVS43/EX44	7.92	5.96	**-24.8**			
c.6348-8T>C	IVS43/EX44	7.92	7.44	-6.1			
c.6348-4A>G	IVS43/EX44	7.92	8.06	1.8			
c.6348-2A>T	IVS43/EX44	7.92	-0.43	**-105.4**			
c.6451A>G	EX44	7.26	3.78	**-47.9**			
c.6452+5T>A	IVS44/EX44	7.26	7.46	2.8			
c.6452+7T>C	IVS44/EX44	7.26	-	-			
c.6453-3T>C	IVS44/EX45	7.83	8.45	7.9			
c.6571A>G	EX45	9.79	8.34	-14.8			
c.6572+1G>A	IVS45/EX45	9.79	1.61	**-83.6**			
c.6572+4T>C	IVS45/EX45	9.79	10.75	9.8			
c.6572+7A>G	IVS45/EX45	9.79	-	-			
c.6573-5T>C	IVS45/EX46	6.7	6.08	-9.3			
c.6573-4T>C	IVS45/EX46	6.7	5.97	-10.9			
c.6574T>C	EX46	6.7	5.42	-19.1			
c.6807+5del	IVS46/EX46	8.76	8.55	-2.4			
c.6807+5A>G	IVS46/EX46	8.76	10.86	24.0			
c.6808-7A>G	IVS46/EX47	8.17	8.8	7.7			
c.6808-2A>T	IVS46/EX47	8.17	-0.18	**-102.2**	0.03	5.72	+19166.67 12 nt downstream
c.6809T>C	EX47	8.17	7.29	-10.8			
c.6974C>T	EX47	4.02	2.7	**-32.8**			
c.6975G>A	EX47	4.02	-2.87	**-171.4**			
c.6975G>C	EX47	4.02	-5.95	**-248.0**			
c.6975+1G>A	IVS47/EX47	4.02	-4.15	**-203.2**			
c.6975+7T>A	IVS47/EX47	4.02	-	-			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.6976-3C>T	IVS47/EX48	9.04	8.38	-7.3			
c.6976A>G	EX48	9.04	10.23	13.2			
c.7088A>G	EX48	10.57	9.21	-12.9			
c.7089+3A>G	IVS48/EX48	10.57	8.67	-17.9			
c.7089+4A>G	IVS48/EX48	10.57	9.45	-10.6			
c.7307+1G>A	IVS49/EX49	8.63	0.44	**-94.9**			
c.7307+4A>G	IVS49/EX49	8.63	6.52	**-24.5**			
c.7307+4A>T	IVS49/EX49	8.63	7.33	-15.1			
c.7307+8G>C	IVS49/EX49	8.63	-	-			
c.7308-10T>G	IVS49/EX50	1.87	-	-			
c.7308-3A>G	IVS49/EX50	1.87	-	-			
c.7308A>C	EX50	1.87	-	-			
c.7308A>G	EX50	1.87	3.88	307.5			
c.7515G>A	EX50	3.24	1.97	**-39.2**			
c.7515+6T>C	IVS50/EX50	3.24	2.46	**-24.1**			
c.7515+7G>A	IVS50/EX50	3.24	-	-			
c.7516-10T>C	IVS50/EX51	6.03	5.71	-5.3			
c.7516-2_7517del	IVS50/EX51	6.03	6.45	7.0			
c.7516A>G	EX51	6.03	7.18	19.1			
c.7629T>C	EX51	8.62	10.74	24.6			
c.7629+2T>G	IVS51/EX51	8.62	0.98	**-88.6**			
c.7629+2T>C	IVS51/EX51	8.62	0.87	**-89.9**			
c.7629+9A>C	IVS51/EX51	8.62	8.62	-			
c.7630-3C>T	IVS51/EX52	6.96	5.51	**-20.8**			
c.7630-2A>C	IVS51/EX52	6.96	-1.07	**-115.4**	0.87	5.36	+716.09 12 nt downstream
c.7630C>G	EX52	6.96	8.4	20.7			
c.7787A>T	EX52	7.63	3.1	**-59.4**			
c.7788G>A	EX52	7.63	0.87	**-88.6**			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.7788+1G>C	IVS52/EX52	7.63	-0.63	**-108.3**			
c.7788+6T>G (+6T)	IVS52/EX52	7.63	7.44	-2.5			
c.7788+7G>A	IVS52/EX52	7.63	-	-			
c.7788+8G>T	IVS52/EX52	7.63	-	-			
c.7789-3T>G	IVS52/EX53	9.87	-0.54	**-105.5**	-4.55	4.04	+188.79 2 nt upstream
c.7789G>A	EX53	9.87	8.2	-16.9			
c.7927+2_7927+3del	IVS53/EX53	6.97	-6.01	**-186.2**			
c.7927+1G>A	IVS53/EX53	6.97	-1.2	**-117.2**			
c.7927+6T>G	IVS53/EX53	6.97	3.5	**-49.8**			
c.7928-10T>C	IVS53/EX54	5.38	4.91	-8.7			
c.7928-10T>G	IVS53/EX54	5.38	3.33	**-38.1**			
c.7928-8A>T	IVS53/EX54	5.38	6.13	13.9			
c.8010+6T>C	IVS54/EX54	8.83	7.46	-15.5			
c.8010+8G>A	IVS54/EX54	8.83	-	-			
c.8010+9C>T	IVS54/EX54	8.83	-	-			
c.8011-6T>G	IVS54/EX55	7.11	6.09	-14.4			
c.8011-2A>G	IVS54/EX55	7.11	-0.84	**-111.8**			
c.8011-2A>C	IVS54/EX55	7.11	-0.93	**-113.1**	0.3	3.09	+1130 14 nt downstream
c.8150A>C	EX55	9.6	9.44	-1.7			
c.8151+3G>A	IVS55/EX55	9.6	10.22	6.5			
c.8152-9A>C	IVS55/EX56	4.97	5.73	15.3			
c.8152-4G>A	IVS55/EX56	4.97	5.38	8.3			
c.8152G>T	EX56	4.97	2.71	**-45.5**			
c.8268+6T>A	IVS56/EX56	9.01	8.92	-1.0			
c.8268+8T>C	IVS56/EX56	9.01	-	-			
c.8269-7A>G (Pyr)	IVS56/EX57	6.65	1.86	**-72.0**	-5.46	3.28	+160.07 6 nt upstream
c.8269-2A>T	IVS56/EX57	6.65	-1.7	**-125.6**	5.88	6.93	+17.86 19 nt downstream

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.8417T>C	EX57	10.13	11.45	13.0			
c.8418+5_8418+8del	IVS57/EX57	10.13	0.78	**-92.3**			
c.8418+5G>T	IVS57/EX57	10.13	3.18	**-68.6**			
c.8418+5G>A	IVS57/EX57	10.13	3.54	**-65.1**			
c.8418+7G>A	IVS57/EX57	10.13	-	-			
c.8419-8A>G	IVS57/EX58	6.53	6.81	4.3			
c.8584+1G>A	IVS58/EX58	3.39	-4.78	**-241.0**			
c.8584+2T>C	IVS58/EX58	3.39	-4.35	**-228.3**			
c.8584+4A>G	IVS58/EX58	3.39	-1.15	**-133.9**			
c.8584+9del	IVS58/EX58	3.39	-	-			
c.8584+9C>T	IVS58/EX58	3.39	-	-			
c.8585-9T>C	IVS58/EX59	10.09	10.86	7.6			
c.8585-4C>A	IVS58/EX59	10.09	10.05	-0.4			
c.8585-4C>T	IVS58/EX59	10.09	9.67	-4.2			
c.8585-3C>T	IVS58/EX59	10.09	9.78	-3.1			
c.8671+1G>A	IVS59/EX59	9.66	1.48	**-84.7**			
c.8671+2T>A	IVS59/EX59	9.66	1.48	**-84.7**			
c.8671+4A>G	IVS59/EX59	9.66	7.92	-18.0	-1.04	6.84	+757.69 2 nt downstream
c.8671+9T>G	IVS59/EX59	9.66	-	-	-1.04	5.17	+597.12 2 nt downstream
c.8672-10T>C	IVS59/EX60	9.74	9.08	-6.8			
c.8672-7A>T	IVS59/EX60	9.74	10.47	7.5			
c.8672-5T>C	IVS59/EX60	9.74	9.03	-7.3			
c.8672-4T>C	IVS59/EX60	9.74	10.2	4.7			
c.8672-3T>G	IVS59/EX60	9.74	2.77	**-71.6**	6.14	6.51	+6.03 18 nt downstream
c.8672-3T>C	IVS59/EX60	9.74	9.18	-5.8			
c.8673T>C	EX60	9.74	8.03	-17.6			
c.8786+1G>A	IVS60/EX60	11	2.82	**-74.4**			
c.8786+4A>T	IVS60/EX60	11	9.79	-11.0			

ATM VARIANTS[1]	EXON/ INTRON	MES wt	MES mut	MES score change[2] (%)	MES *de novo* SS-wt[3]	MES *de novo* SS-mut[3]	MES score change[2] (%)
c.8786+8A>C	IVS60/EX60	11	-	-			
c.8787-9T>C	IVS60/EX61	8.29	8.45	1.9			
c.8787-8G>T	IVS60/EX61	8.29	9.43	13.8			
c.8787-6C>T	IVS60/EX61	8.29	8.46	2.1			
c.8787-3T>C	IVS60/EX61	8.29	8.01	-3.4			
c.8850G>T	EX61	8.27	0.19	**-97.7**			
c.8850+4A>C	IVS61/EX61	8.27	6.62	**-20.0**			
c.8850+5A>C	IVS61/EX61	8.27	9.65	16.7			
c.8850+7T>G	IVS61/EX61	8.27	-	-			
c.8850+12del	IVS61/EX61	8.27	-	-			
c.8851-7T>C	IVS61/EX62	11.33	10.72	-5.4			
c.8851-3T>G	IVS61/EX62	11.33	1.92	**-83.1**			
c.8851-1G>T	IVS61/EX62	11.33	2.73	**-75.9**			
c.8851-1G>C	IVS61/EX62	11.33	3.27	**-71.1**			
c.8852T>C	EX62	11.33	10.08	-11.0			
c.8987+4_8987+15dup	IVS62/EX62	9.6	-	-			
c.8987+3G>A	IVS62/EX62	9.6	9.88	2.9			
c.8987+5G>C	IVS62/EX62	9.6	6.54	**-31.9**			
c.8988-9T>C	IVS62/EX63	10.46	10.82	3.4			
c.8988-7_8988-5del	IVS62/EX63	10.46	5.27	**-49.6**			
c.8988-7T>C	IVS62/EX63	10.46	10.04	-4.0			

[1] Bioinformatically selected variants (MES score change ≤ -20%) are shown in red. Tested variants are yellow-shadowed. The conserved nucleotide of tested variants that did not meet the -20% threshold is shown between brackets.
[2] MES score changes (Δ%), wild type (wt) vs. mutant (mut). Values below the cutoff (≤ -20%) are in red.
[3] *De novo*: predicted creation of new alternative splice sites.

Supplementary Table S2. Mutagenesis primers for *ATM* variants.

Variant	Primers (5'→3')
c.332-5A>G	TCTTTATTTGTTTATTTTGAGATAGGAGCACCTAGGCT
	AGCCTAGGTGCTCCTATATCAAAATAAACAAATAAAGA
c.332-1G>A	TGTTTATTTTGAAATAAGAGCACCTAGGCTAAAATGT
	ACATTTTAGCCTAGGTGCTCTTATTTCAAAATAAACA
c.496G>A	ATCTCAGCAACAGTGGTTAAGTATGTTTTGAAGGTTGT
	ACAACCTTCAAAACATACTTAACCACTGTTGCTGAGAT
c.496+5G>A	AACAGTGGTTAGGTATATTTTGAAGGTTGTTGTTTGTG
	CACAAACAACAACCTTCAAAATATACCTAACCACTGTT
c.901G>T	GCCAAAACCCAAGAAAAATGTATAAAGGAAATGTTTAC
	GTAAACATTTCCTTTATACATTTTTCTTGGGTTTTGGC
c.901+2T>C	GAGCCAAAACCCAAGAAAAAGGCATAAAGGAAATGTTT
	AAACATTTCCTTTATGCCTTTTTCTTGGGTTTTGGCTC
c.901+3A>T	ACCCAAGAAAAAGGTTTAAAGGAAATGTTTACTGTTTTGA
	TCAAAACAGTAAACATTTCCTTTAAACCTTTTTCTTGGGT
c.902-1G>T	TTTAATTTTTTGGATTACATGTGCTTATGAATCAACAA
	TTGTTGATTCATAAGCACATGTAATCCAAAAAATTAAA
c.902G>A	TTTAATTTTTTGGATTACAGATGCTTATGAATCAACAA
	TTGTTGATTCATAAGCATCTGTAATCCAAAAAATTAAA
c.903T>G	TTTTGGATTACAGGGGCTTATGAATCAACAAAATGGAG
	CTCCATTTTGTTGATTCATAAGCCCCTGTAATCCAAAA
c.1065+1G>T	TGTCACCAGTTACAGTAAGTAGGTCATGTCACATTTAG
	CTAAATGTGACATGACCTACTTACTGTAACTGGTGACA
c.1065+3A>G	TGTCACCAGGTGCAGTAAGTAGGTCATGTCACATTTAG
	CTAAATGTGACATGACCTACTTACTGCACCTGGTGACA
c.1066-6T>G	AAGTGGATTTATTTTTATTGTACAGGTTTTTAATGAAG
	CTTCATTAAAAACCTGTACAATAAAAATAAATCCACTT
c.1235+4_1235+5del	CAGAATGATTTTGATCTTGTGCCTTGGTAGTGTTAC
	GTAACACTACCAAGGCACAAGATCAAAATCATTCTG
c.1898G>T	TTTTTCCAAAGCGTGCCAGAATTGTATGTTATCTAATA
	TATTAGATAACATACAATTCTGGCACGCTTTGGAAAAA
c.1898+3_1898+4del	CAAAGCGTGCCAGAATGGTGTTATCTAATAATGCTCTT
	AAGAGCATTATTAGATAACACCATTCTGGCACGCTTTG
c.1898+2T>G	TTTTTCCAAAGCGTGCCAGAATGGGATGTTATCTAATA
	TATTAGATAACATCCCATTCTGGCACGCTTTGGAAAAA
c.1898+3A>T	AAAGCGTGCCAGAATGGTTTGTTATCTAATAATGCTCT
	AGAGCATTATTAGATAACAAACCATTCTGGCACGCTTT
c.2251-1G>C	TTTTTCCTTCTATTCACAATACTCTCTAATGCAATGTG
	CACATTGCATTAGAGAGTATTGTGAATAGAAGGAAAAA
c.2376+1G>A	GCTTGAGCAACTGTACCAAGATAAGATTTTCTTCTTCT
	AGAAGAAGAAAATCTTATCTTGGTACAGTTGCTCAAGC
c.2376+3A>T	TGAGCAACTGTACCAAGGTTAGATTTTCTTCTTCTTGT
	ACAAGAAGAAGAAAATCTAACCTTGGTACAGTTGCTCA
c.2377-6T>A	GCTTGGTTCTTTGTTTGTCTTAAATGCAGAAGAGTCCA
	TGGACTCTTCTGCATTTAAGACAAACAAAGAACCAAGC
c.2377-2A>G	GGTTCTTTGTTTGTCTTAATTGCGGAAGAGTCCAAATA
	TATTTGGACTCTTCCGCAATTAAGACAAACAAAGAACC
c.2467-3A>G	TTTGTTTCTCTTCCTTGGAGGCATCCTTCATCAAAAAG
	CTTTTTGATGAAGGATGCCTCCAAGGAAGAGAAACAAA
c.2638+3A>G	GAACCTGGAGAGAGCCAAAGTACCATAGGTGAATACAT
	ATGTATTCACCTATGGTACTTTGGCTCTCTCCAGGTTC
c.3577-1G>A	CACATTGACTTTTTGGTTCGTGCAAGTTTTAGAGAAAG
	CTTTCTCTAAAACTTGCACGAACCAAAAAGTCAATGTG

c.3577G>C	ATTGACTTTTTGGTTCGTGCAGCTTTTAGAGAAAGTTT
	AAACTTTCTCTAAAAGCTGCACGAACCAAAAAGTCAAT
c.3746+1G>A	ATTGAGGATTTCTATAGATAAGTTTATACATGACATAT
	ATATGTCATGTATAAACTTATCTATAGAAATCCTCAAT
c.3746+4A>C	ATTGAGGATTTCTATAGGTACGTTTATACATGACATAT
	ATATGTCATGTATAAACGTACCTATAGAAATCCTCAAT
c.3746+5G>A	ATTGAGGATTTCTATAGGTAAATTTATACATGACATAT
	ATATGTCATGTATAAATTTACCTATAGAAATCCTCAAT
c.3993G>A	CTTATTGGGAAAACAAGTATGGCTTCAATTTTTATGTA
	TACATAAAAATTGAAGCCATACTTGTTTTCCCAATAAG
c.3993+1G>A	AAAACAGATATGGCTTCAATTTTTATGTACTTTTCATT
	AATGAAAAGTACATAAAAATTGAAGCCATATCTGTTTT
c.3993+5G>T	AAAACAGGTATTGCTTCAATTTTTATGTACTTTTCATT
	AATGAAAAGTACATAAAAATTGAAGCAATACCTGTTTT
c.3994-3C>T	AATATATTTTAATTTTGTGCCCTTGTAGATTGATCACT
	AGTGATCAATCTACAAGGGCACAAAATTAAAATATATT
c.3994-2A>G	TATTTTAATTTTGTGCCCTTGCGGATTGATCACTTATT
	AATAAGTGATCAATCCGCAAGGGCACAAAATTAAAATA
c.4109+1G>T	CCTCTGTGACTTTTCAGGTTATGTACATTTTAAACTTA
	TAAGTTTAAAATGTACATAACCTGAAAAGTCACAGAGG
c.4109+3A>G	GACTTTTCAGGGTGTGTACATTTTAAACTTAGAGAACT
	AGTTCTCTAAGTTTAAAATGTACACACCCTGAAAAGTC
c.4109+5G>A	TTTTCAGGGTATATACATTTTAAACTTAGAGAACTAGC
	GCTAGTTCTCTAAGTTTAAAATGTATATACCCTGAAAA
c.4109+6T>G	ACTTTTCAGGGTATGGACATTTTAAACTTAGAGAACTA
	TAGTTCTCTAAGTTTAAAATGTCCATACCCTGAAAAGT
c.4110-9C>G	ACTGTATTTTTTCCCTTAAGTCTGTTAGGGATTTGGAT
	ATCCAAATCCCTAACAGACTTAAGGGAAAAAATACAGT
c.4110-2A>C	ACTGTATTTTTTCCCTTAACTCTGTTCGGGATTTGGAT
	ATCCAAATCCCGAACAGAGTTAAGGGAAAAAATACAGT
c.4236+1G>A	TTTCCAAAAGCCCTATAAGTATACATGATGAGTTTAAT
	ATTAAACTCATCATGTATACTTATAGGGCTTTTGGAAA
c.4236+5G>A	TTCCAAAAGCCCTGTAAATATACATGATGAGTTTAATA
	TATTAAACTCATCATGTATATTTACAGGGCTTTTGGAA
c.4236+6T>C	TTCCAAAAGCCCTGTAAGCATACATGATGAGTTTAATA
	TATTAAACTCATCATGTATGCTTACAGGGCTTTTGGAA
c.4436+4A>G	CAACCAAAGGTAGATAACATATTTAGACCAATATATAA
	TTATATATTGGTCTAAATATGTTATCTACCTTTGGTTG
c.7307+1G>A	TAAAATTCAGACAAACAGATAACTAGGTTTCTACAAG
	CTTGTAGAAACCTAGTTATCTGTTTGTCTGAATTTTA
c.7307+4A>G	CATAAAATTCAGACAAACAGGTAGCTAGGTTTCTACA
	TGTAGAAACCTAGCTACCTGTTTGTCTGAATTTTATG
c.7515G>A	AAGTCAATGGCATGATGAAAGCAAGTGTTACTCAGCC
	GGCTGAGTAACACTTGCTTTCATCATGCCATTGACTT
c.7515+6T>C	ATGAAGGCAAGCGTTACTCAGCCCAATATTCTACCCT
	AGGGTAGAATATTGGGCTGAGTAACGCTTGCCTTCAT
c.7629+2T>G	GAAGTCCTCAATAATGGAAGTAAACCTGAAAATCAAAC
	GTTTGATTTTCAGGTTTACTTCCATTATTGAGGACTTC
c.7630-3C>T	TAATAGTTCTTTTCTTATAGCTAATCTCTAGAATTTCA
	TGAAATTCTAGAGATTAGCTATAAGAAAAGAACTATTA
c.7630-2A>C	TAATAGTTCTTTTCTTACCGCTAATCTCTAGAATTTCA
	TGAAATTCTAGAGATTAGCGGTAAGAAAAGAACTATTA
c.7787A>T	CTCTCAGCTTGATGTGGTATTTGGATTAAACATACGTA
	TACGTATGTTTAATCCAAATACCACATCAAGCTGAGAG

c.7788G>A	CTCAGCTTGATGAAGTATTTGGATTAAACATACGTACC
	GGTACGTATGTTTAATCCAAATACTTCATCAAGCTGAG
c.7788+1G>C	CAGCTTGATGAGCTATTTGGATTAAACATACGTACCTT
	AAGGTACGTATGTTTAATCCAAATAGCTCATCAAGCTG
c.7788+6T>G	CTTGATGAGGTATTGGGATTAAACATACGTACCTTTTA
	TAAAAGGTACGTATGTTTAATCCCAATACCTCATCAAG

Supplementary Table S3.5

dbSNP	c.HGVS[1]	p.HGVS[1]	ACMG/AMP-based tentative classification Bayesian Framework point-based [2] (experimental splicing codes)		PVS1[3]	PS1	PS2	PS4	PM1
	c.332-5A>G		VUS	(+1+1)=2	N/A	N/A	(-)	N/A	ATM(N/A)
rs747855862	c.332-1G>A		LP	(+4+1+2)=7	**PVS1_Strong (+4)**	N/A	(-)	(-)	ATM(N/A)
	c.496G>A	p.(Glu166Lys)	VUS	(+1)	N/A	(-)	(-)	(-)	ATM(N/A)
rs796051858	c.496+5G>A[10]		LP	(+1+4+1)=6	N/A	N/A	(-)	(-)	ATM(N/A)
	c.901G>T	p.(Gly301Cys)	VUS	(+4+1)=5	**PVS1_Strong (+4)**	N/A	(-)	(-)	ATM(N/A)
	c.901+2T>C		LP	(+8+1)=9	PVS1	N/A	(-)	(-)	ATM(N/A)
rs786203070	c.901+3A>T[11]		VUS	(+1+1)=2	N/A	N/A	(-)	(-)	ATM(N/A)
rs1064793518	c.902-1G>T[11]		P	(+8+1+2)=11	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
rs202208861	c.902G>A	p.(Gly301Asp)	VUS	(+1)	N/A	(-)	(-)	(-)	ATM(N/A)
rs876659335	c.903T>G	p.(Gly301=)	VUS	(+1+1)=2	N/A	N/A	(-)	(-)	ATM(N/A)
rs201089102	c.1065+1G>T		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
rs1282724169	c.1065+3A>G		VUS	(+1+1)=2	N/A	N/A	(-)	N/A	ATM(N/A)
rs201686625	c.1066-6T>G[12]		LB	(-4+1)=-3	N/A	N/A	(-)	(-)	ATM(N/A)
rs770033355	c.1235+4_1235+5del		VUS	(+1+1)=2	N/A	N/A	(-)	N/A	ATM(N/A)
	c.1898G>T	p.(Cys633Phe)	VUS	(+1+1)=2	N/A	(-)	(-)	N/A	ATM(N/A)
rs587782124	c.1898+2T>G		P	(+4+1+8)=+13	**PVS1_Strong (+4)**	N/A	(-)	(-)	ATM(N/A)
	c.1898+3A>T		LB	0	N/A	N/A	(-)	(-)	ATM(N/A)
rs879254094	c.1898+3_1898+4del[11]		VUS	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
rs876659710	c.2251-1G>C		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
rs730881347	c.2376+1G>A		LP	(4+1+2)=+7	**PVS1_Strong (+4)**	N/A	(-)	(-)	ATM(N/A)
rs758083563	c.2376+3A>T		VUS	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
rs876660963	c.2377-6T>A		LB	(-1+1)=0	N/A	N/A	(-)	(-)	ATM(N/A)
rs1057516553	c.2377-2A>G		VUS	(+4+1)=+5	**PVS1_Strong (+4)**	N/A	(-)	(-)	ATM(N/A)
	c.2467-3A>G		LB	(-1)	N/A	N/A	(-)	(-)	ATM(N/A)
rs876660552	c.2638+3A>G		VUS	(-1+1)=0	N/A	N/A	(-)	(-)	ATM(N/A)
	c.3577-1G>A		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
	c.3577G>C	p.(Val1193Leu)	VUS	(+1-1)=0	N/A	(-)	(-)	(-)	ATM(N/A)

	c.3746+1G>A		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
	c.3746+4A>C		VUS	(+1-1)=0	N/A	N/A	(-)	(-)	ATM(N/A)
rs876658419	c.3746+5G>A		LP	(+8+1)=+9	N/A	N/A	(-)	N/A	ATM(N/A)
rs863224566	c.3993G>A	p.(Gln1331=)	VUS	(+2+1)=+3	**PVS1_Moderate (+2)**	N/A	(-)	(-)	ATM(N/A)
rs200196781	c.3993+1G>A[11]		LP	(+4+4)=+8	**PVS1_Strong (+4)**	N/A	(-)	N/A	ATM(N/A)
rs3092842	c.3993+5G>T		LB	(-1)	N/A	N/A	(-)	N/A	ATM(N/A)
	c.3994-3C>T		VUS	(+1-1)=0	N/A	N/A	(-)	(-)	ATM(N/A)
rs587782276	c.3994-2A>G		LP	(+8)	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
rs879254034	c.4109+1G>T		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	N/A	ATM(N/A)
rs1388807238	c.4109+3A>G		VUS	(+1-1)=0	N/A	N/A	(-)	(-)	ATM(N/A)
	c.4109+5G>A		VUS	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
rs368606937	c.4109+6T>G		VUS	(+1)	N/A	N/A	(-)	(-)	ATM(N/A)
rs730881367	c.4110-9C>G[11]		VUS	(+1+2+1)=+4	N/A	N/A	(-)	(-)	ATM(N/A)
	c.4110-2A>C		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
	c.4236+1G>A		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
rs1441281768	c.4236+5G>A		VUS	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
	c.4236+6T>C		LP	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
rs764662712	c.4436+4A>G		VUS	(+1)	N/A	N/A	(-)	N/A	ATM(N/A)
	c.7307+1G>A		P	(+8+1+2)=+11	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
rs730881316	c.7307+4A>G		VUS	(+1)	N/A	N/A	(-)	(-)	ATM(N/A)
	c.7515G>A	p.(Lys2505=)	VUS	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
rs1250327887	c.7515+6T>C		VUS	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
	c.7629+2T>G		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
rs587782448	c.7630-3C>T		VUS	(+1)	N/A	N/A	(-)	(-)	ATM(N/A)
rs587779866	c.7630-2A>C[11]		P	(+8+8)=+16	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
	c.7787A>T	p.(Glu2596Val)	VUS	(+1+1)=+2	N/A	N/A	(-)	(-)	ATM(N/A)
rs587780639	c.7788G>A[11]	p.(Glu2596=)	LP	(+4+1+4)=+13	**PVS1_Strong (+4)**	N/A	(-)	(-)	ATM(N/A)
	c.7788+1G>C		LP	(+8+1)=+9	**PVS1 (+8)**	N/A	(-)	(-)	ATM(N/A)
	c.7788+6T>G		VUS	(+1)	N/A	N/A	(-)	N/A	ATM(N/A)

The table summarizes ACMG/AMP evidences applied to 56 individual ATM variants under investigation to reach a final ACMG-AMP point-based classification. Evidences contr
genetic variant under consideration (e.g. a missense related evidence PS1 is not applicable to an intronic variant). (-) After careful review of the relevant data, the individual v
(VUS) Variant of Uncertain Significance, (LB) Likely Benign. 3) PVS1 code strength as recently proposed in ClinGen Hereditary Breast, Ovarian and Pancreatic Cancer (HBC
(clinicalgenome.org/working-groups/sequence-variant-interpretation/). Points are shown in brackets. See Supplemental Methods for further details. 5) PP3 (missense) not ap
0.154). For c.3993+5G>T gnomADV2 MAF(%) = 0.091 (99%CI 0.077-0.107)

PM2	PM3[4]	PM4	PM5	PM6	PP1	PP2	PP3 (missense)[5]	PP3[6] (splicing)	PP4	PP5	BA1
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	[1]PM3(+2)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	(-)	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	[2]PM3_Strong (+4)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	[1]PM3 (+2)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	N/A	(-)	(-)	(-)	ATM(N/A)	PP3(+1)	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	(-)	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	[4]PM3_VS (+8)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	[1]PM3 (+2)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	PP3(+1)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	PM4	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	(-)	(-)	(-)	ATM(N/A)	(-)	(-)	ATM(N/A)	ATM(N/A)	(-)

PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
(-)	**[3]PM3_Strong (+4)**	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	**[1]PM3 (+2)**	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	**[1]PM3 (+2)**	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
(-)	**[4]PM3_VS (+8)**	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	**PP3(+1)**	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	**[2]PM3_Strong (+4)**	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
PM2_P(+1)	(-)	N/A	N/A	(-)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)

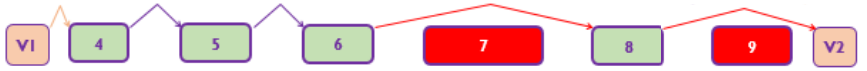
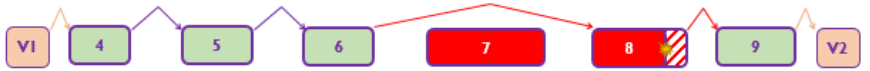
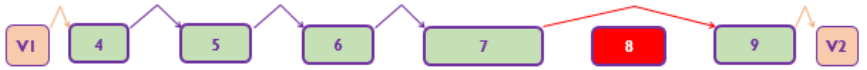
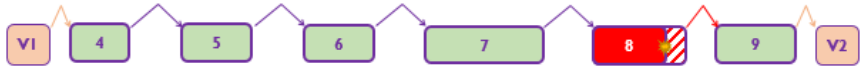
ibuting to the final classification have been gray highlighted (see Supplementary Methods for further details). (ATM N/A) Evidence not applicable to the gene (e.g. at present, PI
variant under investigation does not meet criteria for this particular evidence (e.g., a missense variant does not meet criteria for PM5 if no previous missense pathogenic change
)PC) Expert Panel Specifications for ATM Version 1 (https://www.clinicalgenome.org/affiliation/50039/), including some G to non-G variants at last nucleotide of exon. 4) PM
pplicable (N/A) if predictive splicing codes PVS1 or PP3 are applicable. 6) PP3 splicing if SpliceAI score >0.20. BP4 splicing if SpliceAI score <0.10. PP3_N/A and BP4_N/A if 0.10 <

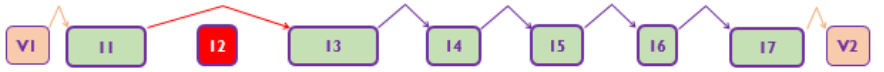
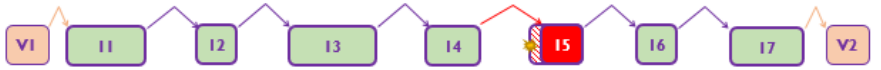
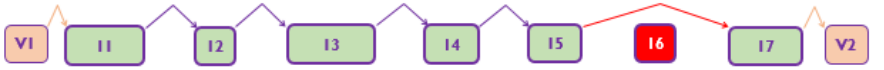
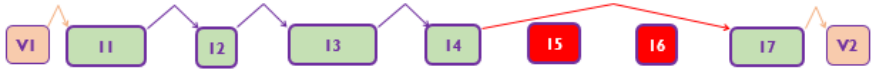
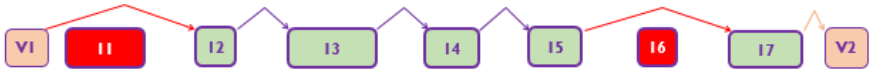
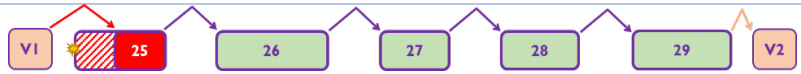
BS1[7]	BS2	BS4	BP1	BP2	BP3	BP4 (missense)	BP4[6] (splicing)	BP5 ATM(N/A)	BP6 ATM(N/A)	BP7
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	**BP4**	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	(-)	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	**(-)**	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(N/A)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
BS1(-4)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	(-)	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	**BP4 (-1)**	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	**BP4 (-1)**	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	**BP4 (-1)**	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	**BP4 (-1)**	(N/A)	ATM(N/A)	ATM(N/A)	N/A

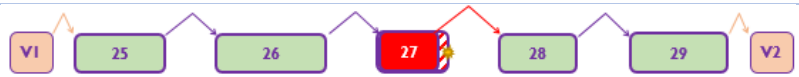
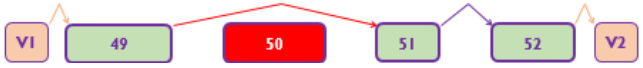
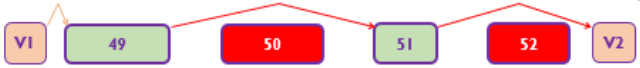
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	**BP4 (-1)**	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
BS1 (-4)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	**BP4 (-1)**	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	**BP4 (-1)**	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	(-)
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	N/A	ATM(N/A)	ATM(N/A)	N/A
(-)	(-)	(-)	ATM(N/A)	(-)	ATM(N/A)	N/A	(-)	ATM(N/A)	ATM(N/A)	N/A

P2 not applicable to ATM, as low rate of benign missense variants has not been proven). (N/A) Evidence not applicable to the specific
e at the same residue observed). 1) HGVS nomenclature using NM_000051.3 as a reference. 2) (P) Pathogenic, (LP) Likely Pathogenic,
M3 code strength according to ClinGen Sequence Variant Interpretation Recommendation for in trans Criterion (PM3) - Version 1.0
< Splice AI score < 0.20. See Supplemental Methods for further details. 7) For c.1066-6T>G gnomADV2 MAF(%) = 0.135 (99%CI 0.118-

Supplementary Table S4. RNA and protein HGVS descriptions according to the Genbank sequence NM_000051.4.

TRANSCRIPT	RNA-HGVS	PROTEIN-HGVS	DIAGRAMS OF THE SPLICING EVENTS
▼(E5p4) Δ(E7) - PTC	r.[331_332ins(332-4_332-1); 663_901del]	p.Arg111Asnfs*4	
Δ(E5p1) - PTC	r.332delG	p.Arg111Lysfs*5	
Δ(E5)	r.332_496del	p.Arg111_Glu166delinsLys	
Δ[(E5)(E7)] - PTC	r.(332_496del; 663_901del)	p.[Arg111_Glu166delinsLys;Gln222Cysfs*3]	
Δ(E7) - PTC	r.663_901del	p.Gln222Cysfs*3	
Δ[(E7)(E9)] - PTC	r.(663_901del;1066_1235del)	p.Gln222Cysfs*3	
Δ(E7_8) - PTC	r.663_1065del	p.Gln222Phefs*34	
[Δ(E7)▼(E8q5)] - PTC	r.[663_901del; 1235_1236ins(1235+1_1235+5)]	p.Gln222Cysfs*3	
Δ(E7_9)	r.663_1235del	p.Gln222_Trp412del	
Δ(E8) - PTC	r.902_1065del	p.Ala302Phefs*2	
▼(E8q5) - PTC	r.[1065_1066ins(1065+1_1065+5)]	p.Val356Valfs*36	
Δ(E9) - PTC	r.1066_1235del	p.Val356Alafs*17	

Δ(E11)	r.1608_1802del	p.Pro537_Ser601del	
Δ(E12)	r.1803_1898del	p.Asn602_Cys633del	
Δ(E12q1) - PTC	r.1898_1899ins(1898+1)	p.Cys633TrpFs*2	
Δ(E15p19) - PTC	r.2251_2269del	p.Ser751Glufs*20	
Δ(E15)	r.2251_2376del	p.Ser751_Lys792del	
Δ(E16)	r.2377_2466del	p.Lys793_Leu822del	
Δ(E15_16)	r.2251_2466del	p.Ser751_Leu822del	
Δ(E16p3)	r.2377_2379del	p.Lys793del	
Δ(E11)Δ(E16)	r.[1608_1802del; 2377_2466del]	p.[Pro537_Ser601del;Lys793_Leu822del]	
Δ(E25) - PTC	r.3577_3746del	p.Val1193Ilefs*4	
Δ(E25_26)	r.3577_3993del	p.Val1193_Gln1331del	
Δ(E25p159)	r.3577_3735del	p.Val1193_Glu1245del	
Δ(E26) - PTC	r.3747_3993del	p.Ser1250Leufs*17	
Δ(E26q120)	r.3874_3993del	p.Val1292_Gln1331del	

Δ(E27) - PTC	r.3994_4109del	p.Ile1332Glyfs*7	
Δ(E27q1) - PTC	r.4109del	p.Asp1371Ilefs*15	
Δ(E28) - PTC	r.4110_4236del	p.Asp1371Ilefs*38	
▼(E28p8) - PTC	r.4110-1_4110ins(4110-8_4110-1)	p.Asp1371Leufs*18	
Δ(E28p53) - PTC	r.4110_4162del	p.Asp1371Asnfs*10	
Δ(E28_29)	r.4110_4436del	p.Asp1371_Arg1479del	
Δ(E29) - PTC	r.4237_4436del	p.Asp1413Alafs*11	
Δ(E49) - PTC	r.7090_7307del	p.Ala2364Ilefs*13	
Δ(E49q38) - PTC	r.7270_7307del	p.Val2424Ilefs*13	
[Δ(E49q38) Δ(E52)] - PTC	r.[(7270_7307del)(7630_7788del)]	p.Val2424Ilefs*13	
[Δ(E49) Δ(E52)] - PTC	r.[(7090_7307del)(7630_7788del)]	p.Ala2364Ilefs*13	
Δ(E50) - PTC	r.7308_7514del	p.Tyr2437Glufs*4	
[Δ(E50) Δ(E52)] - PTC	r.[(7308_7514del)(7630_7788del)]	p.Tyr2437Glufs*4	
Δ(E51)	r.7516_7629del	p.Arg2506_Asn2543del	
Δ(E52)	r.7630_7788del	p.Leu2544_Glu2596del	

Δ(E52p11) - PTC	r.7630_7640del	p.Leu2544Asnfs*23	

Supplementary Table S5. Comparative classification of 21 *ATM* variants.

		ACMG/AMP-based tentative classification (this study)			
		P/LP	**VUS**	**LB/B**	
ClinVar (two stars review status)	**P/LP**	10	3[a]		**13**
	VUS	3[b]	3		**6**
	LB/B		1[c]	1	**2**
		13	**7**	**1**	

Out of the 56 ATM variants investigated in the present study, ClinVar reports 37 (66%), but only 21 (37%) with a clinical classification supported by multiple submitters with no conflicts, (ClinVar two stars review status). The table compares the ClinVar classification in the subgroup of 21 ATM variants with our own classification. We have red-highlighted clinically relevant discrepancies. Apparently, none of the ClinVar submitters based its classification on a formal ACMG/AMP scheme, so that it is difficult to identify the exact source of discrepancy.

a) Up to four variants reported as likely pathogenic (LP) in ClinVar (c.496+5G>A, c.2251-1G>C, c.2376+1G>A, c.2377-2A>G) are downgraded to Uncertain Significance (VUS) in our study.

For c.496+5G>A, it is the evidence of leakiness in the mgATM assay that prevent us from classifying the variant as LP. Existing experimental data in carriers coincides with mgATM analysis in demonstrating exon 5 skipping, but leakiness was not tested (Dörk et al., 2004; Verhagen et al., 2009). Apparently, ClinVar submitters have discarded the possibility of leakiness, ending up in a LP classification.

For c.2251-1G>C, we detect Δ(E15p19) transcripts, but also Δ(E15) and Δ(E15_16) transcripts in a significant proportion (25% of the overall expression). It is the uncertain functional impact of Δ(E15) and Δ(E15_16) that leads us to classify the variant as uncertain significance. As far as we know, there is no experimental data in carriers. Apparently, ClinVar submitters base the likely pathogenic classification on existing experimental data in carriers of c.2251-1G>A (Wright et al., 1996; Hacia et al., 1998) that show Δ(E15p19) only. Whether Δ(E15) and Δ(E15_16) transcripts are not expressed by c.2251-1G>A, or not detectable by the experimental approach, is not obvious.

For c.2376+1G>A, it is the uncertain functional impact of Δ(E15) that leads us to classify the variant as VUS, despite the evidence of an A-T patient (PM3). Apparently, ClinVar submitters assume a clear-cut pathogenic splicing outcome (as far as we know, there is no experimental data in carriers) to reach a LP classification.

For c.2377-2A>G, it is the uncertain functional impact of Δ(E16) and Δ(E16p3) that leads us to classify the variant as uncertain significance. Apparently, ClinVar submitters reach a likely pathogenic classification by assuming a clear-cut pathogenic splicing outcome (as far as we know, there is no experimental data in carriers) together with very unconvincing clinical data: a recent study reports a Sézary Syndrome patient carrying the ATM c.2377-2A>G variant (heterozygous carrier) (Prasad et al., 2016). Indeed, the authors of the original study state that they cannot conclude that this particular germline mutation is involved in disease susceptibility.

b) Up to three variants reported as VUS in ClinVar (c.901+3A>T, c.3577G>C, and c.7307+4A>G) are upgraded to LP in our study.

For c.901+3A>T, existing experimental data in carriers is identical to the observed mgATM read-out, demonstrating exon 7 skipping (Laake et al., 2000). Apparently, ClinVar submitters have missed the experimental data.

For c.3746+5G>A and c.7307+4A>G, the mgATM analysis shows clear-cut deleterious splicing effects that allows us reaching a LP classification. In absence of these evidences (as far as we know, there is no experimental splicing data in carriers), the VUS classification provided by ClinVar submitters is fully justified.

c) The synonymous variant c.903T>G p.(Gly301=) is reported as likely benign by ClinVar submitters, albeit the exact rationale supporting the classification is not specified. Apparently, the possibility of a splicing effect has not been considered. The mgATM shows a significant proportion of Δ(E7_8) transcripts (17%) that justify, in our opinion, a VUS classification.

Supplementary Table S6. Impact of mgATM data on the ACMG/AMP-based tentative classification of 56 *ATM* variants.

ACMG-AMP classification (PVS1_O/BP7_O codes)					
		P	**LP**	**VUS**	**LB**
ACMG-AMP classification (PVS1/PP3/BS4 splicing predictive codes)	**P**	**3**	1[a]			**4**
	LP	2	**14**	2[a]		**18**
	VUS	1	8	**18**	2	**29**
	LB				**5**	**5**
		6	**23**	**20**	**7**	

In absence of experimental splicing data, our approach classifies 27 *ATM* variants as pathogenic/likely pathogenic or likely benign (P/LP/LB). The remaining variants (up to 29) do not classify (VUS). Incorporating mgATM data into the classification scheme has a positive effect, classifying up to 36 ATM variants (P/LP/LB), and reducing uncertainty (VUSs rate reduced from 52% to 36%).

A more detailed analysis revels that the positive effect of incorporating mgATM data into the classification scheme is mainly observed in non GT-AG variants (37 variants, VUSs reduced from 28 to 18), while the impact in GT-AG variants is less obvious (19 variants, VUSs increased from 1 to 2).

Interestingly, for three specific variants, mgATM data increases uncertainty. Specifically:

- c.1898+2T>G is downgraded from P to LP because a splicing predictive code PVS1_Strong is overridden by mgATM data not reaching PVS1_O at any strength (experimental data shows, unexpectedly, that 1898+2T>G is a leaky variant producing a non-negligible proportion of full-length transcripts).
- c.496+5G>A is downgraded from LP to VUS because a splicing predictive code PP3 is overridden by mgATM data not reaching PVS1_O at any strength (experimental data demonstrates a leaky effect, with the variant allele expressing aberrant transcripts, but also a significant proportion of full-length transcripts).
- c.2251-1G>C is downgraded from LP to VUS because a splicing predictive code PVS1 is overridden by PVS1_O_Strong (experimental data shows that, in addition to Δ(E15p19), a PTC_NMD transcripts, a significant proportion of in-frame Δ(E15) and Δ(E15_16) transcripts are expressed).

Supplementary Table S7. Comparative analysis of SpliceAI predictions, mgATM read-outs, and experimental splicing data in carriers

ATM variant[A]	ACMG_AMP predictive code[B]	SpliceAI[C]					mgATM read-out (>10%)[D]	Experimental data in RNA from carriers[E]
		AL	DL	AG	DG	Predicted splicing outcome		
							mgATM_ex4_9[F]	
c.332-5A>G	PP3	.23(5)	.32(169)	.40(1)	-	▼(E5p4)+Δ(E5)	▼(E5p4) leaky	
c.332-1G>A	PVS1	.93(1)	.78(165)	.04(2)	-	Δ(E5)+Δ(E5p1)	Δ(E5)+Δ(E5p1)	
c.496G>A	PP3	.34(-164)	.66(0)	-	-	Δ(E5)	(-)	
c.496+5G>A	PP3	.64(-169)	.82(-5)	-	-	Δ(E5)	Δ(E5) leaky	Δ(E5)[G]
c.901G>T	PP3	.81(-238)	.94(0)	-	-	Δ(E7)	Δ(E7)	
c.901+2T>C	PVS1	.81(-240)	.99(-2)	-	-	Δ(E7)	Δ(E7)	
c.901+3A>T	PP3	.73(-241)	.55(-3)	-	-	Δ(E7)	Δ(E7)	Δ(E7)
c.902-1G>T	PVS1	.96(1)	.95(164)	-	-	Δ(E8)	Δ(E8)	Δ(E8)
c.902G>A	BP4	-	-	-	-	(-)	Δ(E8) leaky	
c.903T>G	PP3	.11 (-1)	.19(167)	-	-	Δ(E8)[H]	Δ(E8) leaky	
c.1065+1G>T	PVS1	-	.97(-1)	-	.27(4)	▼(E8q5)	▼(E8q5)	
c.1065+3A>G	PP3	.13(-166)	.38(-3)	-	.38(6)	▼(E8q9)+Δ(E8)	▼(E8q5)+Δ(E8) leaky	
c.1066-6T>G	PP3	.62(6)	.63(175)	-	-	Δ(E9)	Δ(E7_E9)+Δ(E9) leaky	Δ(E9)[G]
c.1235+4_1235+5del	PP3	.45(-172)	.70(-3)	-	-	Δ(E9)	Δ(E9) leaky	
							mgATM_ex11_17[I]	
c.1898G>T	PP3	.21(-95)	.29(0)		.44(-62)	Δ(E12)+Δ(E12q62)	Δ(E12) leaky	
c.1898+2T>G	PVS1	.57(-97)	1(-2)		.49(-64)	Δ(E12)+Δ(E12q62)	Δ(E12) leaky	Δ(E12)[G]
c.1898+3A>T	BS3[J]	-	-	-	.41(-65)	(-)[J]	(-)	

c.1898+3_+4del	PP3	.56(-97)	.99(-2)		.51(-64)	Δ(E12)+Δ(E12q62)	Δ(E12)	Δ(E12)
c.2251-1G>C	PVS1	.97(1)		.35(20)		Δ(E15p19)	Δ(E15p19)+Δ(E15))+Δ(E15_E16)	
c.2376+1G>A	PVS1	.96(-126)	.95(-1)			Δ(E15)	Δ(E15)	
c.2376+3A>T	PP3	.65(-128)	.78(-3)			Δ(E15)	Δ(E15)	
c.2377-2A>G	PVS1	.99(2)	.39(91)	.07(5)		Δ(E16)+ Δ(E16p3)	Δ(E16)+ Δ(E16p3)	
c.2377-6T>A	BP4					(-)	(-)	
c.2467-3A>G	BP4					(-)	(-)	
c.2638+3A>G	BP4					(-)	(-)	
							mgATM_ex25_29[K]	
c.3577-1G>A	PVS1	.99(1)	.27(170)	.39(2)		Δ(E25p1)+Δ(E25)	Δ(E25)	
c.3577G>C[I]	PP3	.03(0)	.05(169)	.07(159)		Δ(E25)+Δ(E25p159)	Δ(E25)+Δ(E25p159) leaky	
c.3746+1G>A	PVS1	.65(-170)	.98(-1)			Δ(E25)	Δ(E25)	
c.3746+4A>C	BP4					(-)	Δ(E25) leaky	
c.3746+5G>A	PP3	.17(-174)	.24(-5)			Δ(E25)	Δ(E25)	
c.3993G>A	PP3		.19(0)		.68(-120)	Δ(E26q120)	Δ(E26q120)+Δ(E26)	
c.3993+1G>A	PVS1		1(-1)		.76(-121)	Δ(E26q120)	Δ(E26q120)+Δ(E26)	Δ(E26q120)
c.3993+5G>T	PP3				.18(-125)	Δ(E26q120)	Δ(E26q120) leaky	
c.3994-3C>T	BP4					(-)	(-)	
c.3994-2A>G	PVS1	.99(2)	.63(117)	.42(22)		Δ(E27)+Δ(E27p20)	Δ(E27)	
c.4109+1G>T	PVS1	.06(-116)	1(-1)		.79(-2)	Δ(E27)+Δ(E27q1)	Δ(E27)+Δ(E27q1)	
c.4109+3A>G	PP3	.04(-118)	.11(-3)			Δ(E27)	Δ(E27) leaky	
c.4109+5G>A	PP3	.45(-120)	.83(-5)			Δ(E27)	Δ(E27)	
c.4109+6T>G	PP3	.02(-121)	.09(-6)			Δ(E27)	Δ(E27) leaky	
c.4110-9C>G	PP3	.58(9)	.02(135)	.91(1)		▼(E28p8)+Δ(E28)	▼(E28p8)+Δ(E28)	▼(E28p8)
c.4110-2A>C	PVS1	.99(2)	.41(128)	.11(55)		Δ(E28)+Δ(E28p53)	Δ(E28)+Δ(E28p53)	
c.4236+1G>A	PVS1	.66(-127)	1(-1)			Δ(E28)	Δ(E28)	
c.4236+5G>A	PP3	.60(-131)	.98(-5)			Δ(E28)	Δ(E28)	
c.4236+6T>C	PP3	.37(-132)	.51(-6)			Δ(E28)	Δ(E28)	
c.4436+4A>G	PP3	.07(-203)	.1(-4)			Δ(E29)	Δ(E29) leaky	
							mgATM_ex49_52[L]	
c.7307+1G>A	PVS1	.21(-136)[M]	.96(-1)		.12(-39)	Δ(E49q38)	Δ(E49q38)+Δ(E49)	

c.7307+4A>G	PP3	.20(-139)[M]	.63(-4)		.12(-42)	Δ(E49q38)	Δ(E49q38)	
c.7515G>A	PP3	.37(-199)[N]	.88(0)		.32(-70)	Δ(E50q70)	Δ(E50)	
c.7515+6T>C	PP3	.31(-205)[N]	.54(-6)		.28(-76)	Δ(E50q70)	Δ(E50) leaky	
c.7629+2T>G	PVS1	.34(-115)	.99(-2)			Δ(E51)	Δ(E51)	
c.7630-3C>T	PP3	.38(3)	.35(161)			Δ(E52)	Δ(E52) leaky	
c.7630-2A>C	PVS1	.91(2)	.13(160)	.53(13)		Δ(E52p11)+Δ(E52)	Δ(E52p11)+Δ(E52)	Δ(E52p11)+Δ(E52)
c.7787A>T	PP3	.76(-157)	.77(1)			Δ(E52)	Δ(E52)	
c.7788G>A	PP3	.73(-158)	.81(0)			Δ(E52)	Δ(E52)	Δ(E52)
c.7788+1G>C	PVS1	.82(-159)	.82(-1)			Δ(E52)	Δ(E52)	
c.7788+6T>G	PP3	.09(-164)	.12(-6)			Δ(E52)	Δ(E52) leaky	

A) HGVS

B) ACMG-AMP splicing predictive evidence codes based on location (PVS1) and SpliceAI scores (PP3/BP4). Splicing predictive codes do not contribute to our final ACMG-AMP classification (see Supplementary Table S1), as we think that experimental mgATM splicing evidence (PS3/BS3) overrides predictive evidence (see Supplemental Methods for further details).

C) SpliceAI parameters were as follow (genome version hg38; score type raw; max distance 10000nt, Illumina's pre-computed scores yes). The table shows Acceptor Loss (AL), Donor Loss (DL), Acceptor Gain (AG), and Donor Gain (DG) scores (and positions) corresponding to annotated splice sites (NM_000051.3). Positions are annotated as (-) if upstream of the variant, or (+) if downstream. A minimum of scores are required to predict a specific aberrant outcome (e.g. for a variant damaging a donor site, an acceptor loss scoring at the right position predicts exon skipping, while a donor gain will predict use of a cryptic/de novo site). Otherwise (one or no scores), the full-length transcript (FL) is the default prediction. Scores below the high recall score (<20%) are grey-highlighted. Apparently, scores <20% might be relevant, identifying several events confirmed by experimental data. False predictions (not supported by experimental data) are red-highlighted.

D) mgATM transcripts representing <10% of the overall expression (and/or unannotated transcripts) are not shown. Leaky if full-length transcript (>10% of the overall signal) is detected.

E) If needed, exon numbering (as reported in the indicated manuscripts) has been adapted to NM_000051.3. c.496+5G>A (Dörk et al., 2004); c.901+3A>T (Laake et al., 2000); c.902-1G>T (Teraoka et al., 1999); c.1066-6T>G (Soukupova et al., 2008); c.1898+2T>G (Stankovic et al., 1998); c.1898+3_+4del (Laake et al., 2000); c.3993+1G>A (Laake et al., 2000; Cavaciuti et al., 2005); c.4110-9C>G (Demuth et al., 2011). For c.7630-2A>C some studies detect Δ(E52) only, some Δ(E52p11) only, and some detect both transcripts (Sandoval et al., 1999; Teraoka et al., 1999; Laake et al., 2000; Cavaciuti et al., 2005), c.7788G>A (Broeks et al., 1998).

F) The wt minigene mgATM_ex4_9 produces a high proportion of Δ(E7) transcripts (34% of the overall expression). For the purpose of this comparative analysis, Δ(E7) transcripts (if representing ≤40% of the overall signal) have been excluded from the mgATM_ex4_9 read-out.

G) To what extent these studies did not detect or did not asses leakiness is hard to say

H) SpliceAI predicts AL at -1(.11) and DL at +167 (.19) suggesting Δ(E8), but scores are <20%, and SpliceAI fails predicting the exact exon 8 size (164nt), predicting instead an exon of 168nt. Splice AI error mapping exon 8 boundaries is caused by a very strong cryptic donor site (MES score=9.49) nearby the native site (MES=8.68). Indeed, this strong cryptic site explains ▼(E8q5) observed in c.1065+1G>T and c.1065+3A>G.

I) The wt minigene mgATM_ex11_17 produces a significant proportion of Δ(E11) transcripts (16% of the overall expression). For the purpose of this comparative analysis, Δ(E11) transcripts (if representing ≤16% of the overall signal) have been excluded from the read-out.

J) SpliceAI predicts for c.1898+3A>T a DG at positon -65, suggesting Δ(E12q62). Yet, SpliceAI does not predict any DL at the native donor site, supporting (in our opinion) no splicing alteration.

K) The wt minigene mgATM_ex25_29 produces only full-length transcripts. Consequently, all read-out transcripts representing >10% of the overall expression) have been included in the comparative analysis.

L) The wt minigene mgATM_ex49_52 produces a high proportion of Δ(E52) transcripts (24% of the overall expression). For the purpose of this comparative analysis, Δ(E52) transcripts (if representing ≤25% of the overall signal) have been excluded from th read-out.

M) For variants targeting exon 49 donor site, acceptor loss score >.20 suggests exon 49 skipping. Yet, SpliceAI fails predicting ATM exon 49 size (218nt), predicting instead a much shorter version of 136nt (i.e. SpliceAI misplaces exon 49 acceptor site at an internal cryptic site). This is probably related to the fact that ATM exon 49 native acceptor site is very weak (MES score=0). Use of a cryptic donor site is righty predicted Δ(E49q38), but below the 20% threshold.

N) SpliceAI predictions (AL) suggests exon 50 skipping. Yet, SpliceAI fails predicting ATM exon 50 size (208nt), predicting instead a shorter version of 200nt (i.e. SpliceAI misplaces exon 50 acceptor site at an internal cryptic site). This is probably related to the fact that ATM exon 50 native acceptor site is very weak (MES score=1.87

GGATCCagaccagatattaaattggtcttgtaggagttaggccttgaaagagagatttaattgttttatttgttt
ttttcagctgatgtagtaatctaagcaaggtggtttaaaagttgctctttgtgatggcatgaacagcttttgaaa
ttattataatttaagtattcaacgagtttctgaaattgcattttgttttcttgaagATTTTTACAGAAATATATT
CAGAAAGAAACAGAATGTCTGAGAATAGCAAAACCAAATGTATCAGCCTCAACACAAGCCTCCAGGCAGAAAAAG **Ex4**
ATGCAGGAAATCAGTAGTTTGGTCAAATACTTCATCAAATGTGCAAACAGAAgtaagtgatgttataaattataa
ataaatggcttaacagattactgtcgcgtgagttttttttttttttcagatcattttaaggacttaactgttgca
taagtttgttctaaatagaataagataaaagttgagtgtaagtacatataatgatttttatttttattaaaaggt
ttttttttagggctagtcaagtgaagcgactatgttccaataaaatgttatttgcaaagaaaaaaaaaatgaatt
ttgcttggggccataatttgccaatttcttctctacaaaagaagtttagttaatagtaatttcccaaatggaatt
atttaaatagttgccattccaagtgtcttatttttgttcaaatttatgtttttctttatttgtttatttgaaat
agGAGCACCTAGGCTAAAATGTCAAGAACTCTTAAATTATATCATGGATACAGTGAAAGATTCATCTAATGGTGC
TATTTACGGAGCTGATTGTAGCAACATACTACTCAAAGACATTCTTTCTGTGAGAAAATACTGGTGTGAAATATC **Ex5**
TCAGCAACAGTGGTTAGgtatgttttgaaggttgttgtttgtgaatttttcctcatgaaatgaaacttcaccaaa
gaaagcactctgtctgtatctgtctatatcccccaagtgacctgacagtttaacagtactttagtaaaattatat
ggttatcgaactgacccttaatttttatttattatgtagcttttgaataaagtcatgaataatatatagttattc
ctgtcttaaaggcattttatataagataaaattaaatactgatggagtacttttactatgttaaaaataaattag
tgcagttttaaaatccttttttctgtatgggattatggaatatttaagttaaattgtaacatttaatacattttga
ttttttaaaaaatcatgactaataatttttttttttttttttaagAATTGTTCTCTGTGTACTTCAGGCTCTATCTGA
AACCTTCACAAGATGTTCATAGAGTTTTAGTGGCTAGAATAATTCATGCTGTTACCAAAGGATGCTGTTCTCAGA **Ex6**
CTGACGGATTAAATTCCAAATTTTTGGACTTTTTTTCCAAGGCTATTCAGTGTGCGAGgtaatctaatctctttt
tcttttgttttgtattgaaatactttgatcttgcaagaccatgttttagactcagtaactaaaaattctacctt
aaaataaaacattgatccatcataacagaactagtggattcctaaagagacaaccaagtccaacactttctgaat
atccaatatgcagaacactacgtgaagttttcacaccaaataagaactaatttttttgtcagtgtgaagtaatgct
gtgatttttttttttaatgaatagttttgaaattaagactactgtttgaaaattagggttttgttttttttttctt
tcagcataccacttcataactgttcagtttgtacagtttgttccccctgttatacccagttgagcttgtttgttt
cttcacagACAAGAAAAGAGCTCTTCAGGTCTAAATCATATCTTAGCAGCTCTTACTATCTTCCTCAAGACTTTG
GCTGTCAACTTTCGAATTCGAGTGTGTGAATTAGGAGATGAAATTCTTCCCACTTTGCTTTATATTTGGACTCAA **Ex7**
CATAGGCTTAATGATTCTTTAAAAGAAGTCATTATTGAATTATTTCAACTGCAAATTTATATCCATCATCCGAAA
GGAGCCAAAACCCAAGAAAAAGgtataaaggaaatgtttactgttttgaatttgcttcttcattcaaacatagaa
gtctaagtataaaattagtgttctttaggaggatatgactttcctctggatttctctggttgataatgttactta
gccatgagaatgtttttcatagagtttttagaatttgaagcagtggtagcagattctgttatgcctttctatagt
tatctaaatgtaaaatggacagttatgaaatagagtttgtcaaggatcttgtcagaagaggcattagtactgaat
ggaaggttggaccaggtgtcttctaacgctgatgcagcttgacagctgaataattttgtgggagctagcagtgta
aacagagtacatacataaaaattacattttaatttttggattacagGTGCTTATGAATCAACAAAATGGAGAAG
TATTTTATACAACTTATATGATCTGCTAGTGAATGAGATAAGTCATATAGGAAGTAGAGGAAAGTATTCTTCAGG **Ex8**
ATTTCGTAATATTGCCGTCAAAGAAAATTTGATTGAATTGATGGCAGATATCTGTCACCAGgtacagtaagtagg
tcatgtcacatttagaaatttcctgttaatttttttttttaaactgggcattttgggcttttaaaacctgtgttct
cacaaaaagcctataaaatgactctgtacatgcaactattcctttcaaactatcagaaatatttggaattaccct
tttaacttaaaagttaatgcttttgcagatatttgaAAGCTTtcttgctttatgagttccctgtatcttcatttc
agaaagaatttttttttcccccctaatgaaatattagaagacaatagcaatgaaaaaagtgtatttggccaggcgtg
gtggctcacacctgtaatcccGgcactttgggaggccaaggcgggcagatcacctgaggttgggagttcaagacc
agtctgaccaacatggagaaactctgtccctactaaaaatacaaagttagctgggcatggtggtgcaagcctgta
atcccagctactcgggaggctgaggcaggagaatcacttgaacctgggaggtagaggttgtggtgatacgagatc
gtgctgttccactccaacctgggcaacaacagcgaaactctggctcaaaaaaaaaaaaaagaaaaaagtggattt
attttattttacagGTTTTTAATGAAGATACCAGATCCTTGGAGATTTCTCAATCTTACACTACTACACAAAGA
GAATCTAGTGATTACAGTGTCCCTTGCAAAAGGAAGAAAATAGAACTAGGCTGGGAAGTAATAAAAGATCACCTT **Ex9**
CAGAAGTCACAGAATGATTTTGATCTTGTGCCTTGgtaaagtgttaccattttctcattcagtgtcattttaatc
tcttgtatgttatttttcagaaaactttcagtggaatcctttcatctcaaccagaactaagtcatttgtctaccc
ccaaacctattactagcaaagggatatgtgattgccatgacaaatgagatcaatcattaatggctcatttgcttg
ggccaagtgcagggccacctattttaatcatattgtcatagcaagttaagaaggag

Supplementary Figure S1-A. Insert sequence of minigene mgATM_ex4-9. Exons 4 to 9 are indicated in upper case. Structure of the insert (3,731 bp): BamHI - ivs3 (200 bp) – ex4 (146 bp) – ivs4-1 (200bp) // ivs4-2 (200 bp) – ex5 (165 bp) – ivs5-1 (200 bp) // ivs5-2 (200 bp) – ex6 (166 bp) – ivs6-1 (200 bp) // ivs6-2 (200 bp) – ex7 (237 bp) – ivs7-1 (200 bp) // ivs7-2 (200 bp) – ex8 (164 bp) – ivs8-1 (200 bp) -HindIII- ivs8-2 (423 bp) – ex9 (170 bp) - ivs9 (246 bp) - SalI (shortened introns are indicated by a double slash //)

tttaaatttcttttatgtgcaatttatcattatttattaaatagccatgtttaaattgtagtactatgcactgtt
aataaacgagctatttttttaatcaagaatcttcccaaatgtaatcagacttttaacagtttttatgttcatttag
tcaccttaactaaatgtatgtgccaggcactgtcctgatagataaagtctttgcccctccaatagcttgcttttc
acaattgtcctttgttttgttatagTCCTGCAGTATGCTGTTTGACTTTGGCACTGACCACCAGTATAGTTCCAG
GAACGGTAAAAATGGGAATAGAGCAAAATATGTGTGAAGTAAATAGAAGCTTTTCTTTAAAGGAATCAATAATGA **Ex11**
AATGGCTCTTATTCTATCAGTTAGAGGGTGACTTAGAAAATAGCACAGAAGTGCCTCCAATTCTTCACAGgtaat
ttaagttcattagcatgctgctgttttttttgtttgttttatcaggctctctccacttatttgatgccagatggc
tttatttttataataataatgcagaatttcccagatctaaccttaattattaaatattatgtttgtttttacagtt
atctgtgtctttatgcctgattgcttctgaaataaagggttgtcttggaaatgataataacaatggttgtcctcc
ttaaattgtccttttagatattaagaaatttagtatagatgaaagcaattttaatctaggatccaaattttagaa
gtcaagatttatagctaaacatggatgttaaagtttaaagtattctttacatggcttttggtcttctaagtgaag
cttttgtttttctttgtagTAATTTTCCTCATCTTGTACTGGAGAAAATTCTTGTGAGTCTCACTATGAAAAAC **Ex12**
TGTAAAGCTGCAATGAATTTTTTCCAAAGCGTGCCAGAATGgtatgttatctaataatgctctttatcattttaa
gctatagctttaattacaaagatgataattttagctgggtagtagctgcatcttaataattgtaaactaaattgg
tccaaaaaaaattgcaactgttagccagggaagaggttgttttaattcagtgattgtaatctatgttatataacat
tagaccaagcttactatgaataatgacatttgatataagtaggtctcaaagtccgaagaagagaagcatttaaaa
gaataatctattaattatataagtagtctttgaatgatgtagatactaggttaatgttttcctttgtaatatatt
gctaatacatataaggcaaagcattaggtacttggtttatatattaaagatcttactttcttgaagTGAACACCA
CCAAAAAGATAAAGAAGAACTTTCATTCTCAGAAGTAGAAGAACTATTTCTTCAGACAACTTTTGACAAGATGGA
CTTTTTAACCATTGTGAGAGAATGTGGTATAGAAAAGCACCAGTCCAGTATTGGCTTCTCTGTCCACCAGAATCT **Ex13**
CAAGGAATCACTGGATCGCTGTCTTCTGGGATTATCAGAACAGCTTCTGAATAATTACTCATCTGAGgtgagatt
ttttaaaaaaagaactaagcttatatatgattcaactttggtaaactgttaggaaggagaaataggggcaggaaa
aacagcaaggatggtgggaggcttcattttaaaagcaaagtggcagtaaagggctctaaattggacaacttagca
taattaaaggaaaactcaagaataataatttgagtacttcctgctatttttcttgagaatcctggttataattct
acagtgatctcctagttgtttttagagctatccaggatatgccacctttaactcagttaactgaacttttgtttt
ttaatatgtatgtagaatttgttcttacaaaagatagagtatactaaattatttatgaaatatatatatttttat
ttgtggtttactttaagATTACAAATTCAGAAACTCTTGTCCGGTGTTCACGTCTTTTGGTGGGTGTCCTTGGCT **Ex14**
GCTACTGTTACATGGGTGTAATAGCTGAAGAGGAAGCATATAAGTCAGAATTATTCCAGAAAGCCAAGgtaggag
aatttatactaataaagtttcggataaatttgaatgaaatgtattcctgtgaaaattattacatttgtttggaag
acattaaattttatgcaggttaacccttctctttttatttatgtaatgtgagaagaaattatactatgtatttt
taaattgttttaattgtttaatttttaattattattatactttatttgccattttaaaaaattgttaatgagttt
tgcttatactgtatgactacgtggaacttctaaaaacatttcatttttttctcttaagtgcactttatttttatt
ttatagtatgtccaagatcaaagtacactgtaaaaagcaatactaaactataatttttaactggaatttgcatttt
tccttctattcacaatagTCTCTAATGCAATGTGCAGGAGAAAGTATCACTCTGTTTAAAAATAAGACAAATGAG **Ex15**
GAATTCAGAATTGGTTCCTTGAGAAATATGATGCAGCTATGTACACGTTGCTTGAGCAACTGTACCAAGgtaaga
ttttcttcttcttgttttgtttttgagataggatctttctctgtcacccaggctggagtgcagtgggattgtca
caactcattgtagccttgacctcctggtttccagcaattctcctgcctcagcctcccaagtaattgggactacag
gcatgtaccacctagctaaaatttttcttttttacttgaaagtgtaatgttgttaaagcacaatgaaagatgtacag
tatagttattataactctaagaaaagatgtgtttttgaagcagcatatatattggccctaatagtaaactattta
tctacattccattcaagatagagaaaacactgtctgccaagaataattgtttttatttctttgttgcttggttct
ttgtttgtcttaattgcagAAGAGTCCAAATAAGATTGCATCTGGCTTTTTCCTGCGATTGTTAACATCAAAGCT **Ex16**
AATGAATGACATTGCAGATATTTGTAAAAGTTTAgtaagtatgcttcctgttttgctatcatattttgattctaa
taggcataatttttttgttgaaatatctttgtaaataaggatgcatctcacaacatatagctcttaacattttta
caaatgtggaaattaaggccaggtgcggtggctcatgcctgtaaccccagcactttgggaggccgaggtgggtgg
atcacaaggttttaaatagagacaaggtttcaccatgttggccaggctggtttcgaactcccgacctcaggtgat
ccacctggctctgcctcccaaattgctgagattacagatgtgagccactgtgcccagcctgattaggtaaatttt
gactacagcatgctcctgcaagaagccatcttgaacatctttgtttctcttccttgaagGCATCCTTCATCAAAA
AGCCATTTGACCGTGGAGAAGTAGAATCAATGGAAGATGATACTAATGGAAATCTAATGGAGGTGGAGGATCAGT **Ex17**
CATCCATGAATCTATTTAACGATTACCCTGATAGTAGTGTTAGTGATGCAAACGAACCTGGAGAGAGCCAAAGTA
CCATAGgtaaatacatatttactacttgggatttctttacttctttatattgatttggcagtataagaggcctc
attgatatcaattttgtgcttatttcattttctcttagtatagccttttaggattgttcctttcttatatacttt
attttttttttatttttacttgaatttattagtttcatattttattcttcatagaaggaacttaagataactatt
aaagaaataaa

Supplementary Figure S1-B. Insert sequence of minigene mgATM_ex11-17. Exons 11 to 17 are indicated in upper case. Structure of the insert (3,911 bp): SacII – ivs10 (250 bp) – ex11 (195 bp) – ivs11-1 (200bp) // ivs11-2 (200 bp) – ex12 (96 bp) – ivs12-1 (200 bp) // ivs12-2 (200 bp) – ex13 (226 bp) – ivs13-1 (200 bp) // ivs13-2 (200 bp) – ex14 (126 bp) – ivs14-1 (200 bp) // ivs14-2 (200 bp) – ex15 (126 bp) – ivs15-1 (200 bp) // ivs15-2 (200 bp) – ex16 (90 bp) – ivs16-1 (200 bp) // ivs16-2 (200 bp) – ex17 (172 bp) – ivs17 (230 bp) - SalI (shortened introns are indicated by a double slash //)

ttttaaatagagacaaggtttcaccatgttggccaggctggtttcgaactcccgacctcaggtgatccacctggc
tctgcctcccaaattgctgagattacagatgtgagccactgtgcccagcctgattaggtaaattttgactacagc
atgctcctgcaagaagccatcttgaacatctttgtttctcttccttgaagGCATCCTTCATCAAAAAGCCATTTG
ACCGTGGAGAAGTAGAATCAATGGAAGATGATACTAATGGAAATCTAATGGAGGTGGAGGATCAGTCATCCATGA **Ex17**
ATCTATTTAACGATTACCCTGATAGTAGTGTTAGTGATGCAAACGAACCTGGAGAGAGCCAAAGTACCATAGgta
aatacatatttactacttgggatttcttttacttctttatattgatttggcagtataagaggcctcattgatatc
aattttgtgcttatttcattttctcttagtatagccttttaggattgttcctttcttatatactttatttttttt
ttatttttacttgaatttattagtttcatattttattcttcatagaatattttaaattatttcttgacaacagaa
tcttggataatttttcaaaaagacttttgaagctttcagtatataattaatttcactataatttgcttttcata
tacttttttttgtgaagaggaggaaatttgagttaatatgactatatatggctgttgtgcccttctcttagtgtt
aatgagtgcttttttatttttagGTGCCATTAATCCTTTAGCTGAAGAATATCTGTCAAAGCAAGATCTACTTTTC
TTAGACATGCTCAAGTTCTTGTGTTTGTGTGTAACTACTGCTCAGACCAATACTGTGTCCTTTAGGGCAGCTGAT **Ex18**
ATTCGGAGGAAATTGTTAATGTTAATTGATTCTAGCACGCTAGAACCTACCAAATCCCTCCACCTGCATATGgtg
agttacgttaaatgaagaagctcttggattttatctgatgttgctgactaaatgtaatgagttgacatgtaagaa
tcacatggtgtctttgaagaattgaaattgctttcttgagaaatgaacctgagactagttggaaaataacacttt
taacgtgctgtgagcaaatttaagtggatgctgaaatattaaaacttttttgtgtttttagtagagatggagttt
caccatattggccaggctgttctcaaactcctgaccttgtgatctgcctgcttcagcctcccaaagtgctgggat
tacaggtgtgagccactgcacccggcctatgtttatatacttttaaagtaaatgatttgtggataaacctgatt
ttttccctcctaccatcttagTATCTAATGCTTTTAAAGGAGCTTCCTGGAGAAGAGTACCCCTTGCCAATGGA **Ex19**
AGATGTTCTTGAACTTCTGAAACCACTATCgtaagaaattaaaaccttatgttatgttcactttaaagttataaa
ataactgatgtgttctgttaagcttataaagttgaacttttttttttttttaccacagCAATGTGTGTTCTTTG
TATCGTCGTGACCAAGATGTTTGTAAAACTATTTTAAACCATGTCCTTCATGTAGTGAAAAACCTAGGTCAAAGC **Ex20**
AATATGGACTCTGAGAACACAAGGGATGCTCAAGGACAGTTTCTTACAGTAATTGGAGCATTTTGgtaggtacag
tctattttgtggtcctatttttcttttgctatctgtggatacgaatgcaagttttgtatccacatcagtgatttc
ttctgatcttcctacatagctaatacatcttttaagaatagcagaatgtaatttgtgtttccctcagtcgcttga
agaactacattgctttttgtttaaggcttggctttctaaagctaaaatgaattcttttacactaatttcttttag
cttgaatttttggcaaggtgagtatgttggcatattccacataatgacaaataagtttagcacagaaagacatat
tggaagtaacttacaataacctttcagtgagttttctgagtgcttttatcagaatgattatttaactttggaaaa
cttacttgatttcagGCATCTAACAAAGGAGAGGAAATATATATTCTCTGTAAGAATGGCCCTAGTAAATTGCCT **Ex21**
TAAAACTTTGCTTGAGgtgagtttttgcatttttttagtaagatctccattgaaaattttaaagcagtctttgtt
tgttaatgagtaatttttctctatttcatatttaaccacagttcttttcccgtagGCTGATCCTTATTCAAAATG
GGCCATTCTTAATGTAATGGGAAAAGACTTTCCTGTAAATGAAGTATTTACACAATTTCTTGCTGACAATCATCA **Ex22**
CCAAGTTCGCATGTTGGCTGCAGAGTCAATCAATAGgtaatgggtcaaatattcatgaagtatttggaatgctgc
agatggcagtagaatgtcttacatagtaacagctcacagttgcaatattaaaaatagctaacacttgttgagtat
atacggtgtgcctggcatttatgtttattcttaattcttatacttctgtcacttagattctattatttccttcaa
tttataaatga

Supplementary Figure S1-C. Insert sequence of minigene mgATM_ex17-22. Exons 17 to 22 are indicated in upper case. Structure of the insert (2,636 bp): EagI – ivs16 (200 bp) – ex17 (172 bp) – ivs17-1 (200bp) // ivs17-2 (200 bp) – ex18 (200 bp) – ivs18-1 (200 bp) // ivs18-2 (200 bp) – ex19 (83 bp) – ivs19 (104 bp) – ex20 (156 bp) – ivs20-1 (200 bp) // ivs20-2 (200 bp) – ex21 (76 bp) – ivs21 (114 bp) – ex22 (131 bp) – ivs22 (200 bp)- BamHI (shortened introns are indicated by a double slash //)

caaatgttggaaactcttagcatatttttatatttttatttacttaacatttattaagggtttgctatgtgtcag
atactgtgccagttgagtacattttcttaattattattcccatctcatagatgaggaaatcaagaaaagttgaat
gaatgttgtttctaggtcctactctaaataatattaacaagcatttaaatgatttatttttttcatttttcttaa
cacattgacttttttggttcgtgcagGTTTTAGAGAAAGTTTCTGAAACTTTTGGATATAGACGTTTAGAAGACTT
TATGGCATCTCATTTAGATTATCTGGTTTTGGAATGGCTAAATCTTCAAGATACTGAATACAACTTATCTTCTTT
TCCTTTTATTTTATTAAACTACACAAATATTGAGGATTTCTATAGgtaagtttatacatgacatatgtgaaattt
gtttaatttaaaattagttaacaatacttagcaagtcccctcaccagcaacacacataccatacccatacacatg
tgtgtgtgggagcctacatagtatgagaagcaggacagcttcttttaataagaatgtattgaagggagtcactgg
acttcagatctggtcctagctaagctttctaatttttttaatgtgactatttagaatttacttaatttttccatt
tataaaattaaagaatgtttaataatctggataaagtatgatactttaatgctgatggtattaaaacagttttta
agaactattttataaaattttacttggaaaagttatatataacctgtattttaaatttttctatttttagATCTT
GTTATAAGGTTTTGATTCCACATCTGGTGATTAGAAGTCATTTTGATGAGGTGAAGTCCATTGCTAATCAGATTC
AAGAGGACTGGAAAAGTCTTCTAACAGACTGCTTTCCAAAGATTCTTGTAAATATTCTTCCTTATTTTGCCTATG
AGGGTACCAGAGACAGTGGGATGGCACAGCAAAGAGAGACTGCTACCAAGGTCTATGATATGCTTAAAAGTGAAA
ACTTATTGGGAAAACAGgtatggcttcaatttttatgtacttttcattccctgaatgatatgagatataacctttt
aagttttaaggctatttattcgatttattcgtatttatatattgaaacttagcttgtggtaatcattatctagca
tagccaacccatgaatttttttggttatgtcgtgttgtctccctctgattggcttttaactaagtagttcctctt
agtctacaggttggctgcatagaagaaaaaggtagagttatttataatcttgtaaatcttggactttgagtcatc
tattttcttttacagtcatcgaatacttttggaaataaggtaatatatgccttttgagctgtcttgacgttcaca
gatataaaatattaaatatattttaattttgtgcccttgcagATTGATCACTTATTCATTAGTAATTTACCAGAG
ATTGTGGTGGAGTTATTGATGACGTTACATGAGCCAGCAAATTCTAGTGCCAGTCAGAGCACTGACCTCTGTGAC
TTTTCAGGgtatgtacattttaaacttagagaactagctctaacttcacaagttttaaagaagtttattggttg
acaccttcaatgtctatttcaatttatagacatcactcttttaaaaaattttctttaaaaatagccacctttga
attgaggtaattacctatcctctttaccatagtctccccacagtaattactctttctaaggtatttagataatc
tgatttatatatctggactgtgatatgtcatttgtgattttattgaaagtatagttttcagtagaaaaatggtt
tttgaatttgggggttattaaaatctaaattttcattttggaagttcactggtctatgaacaaaacttttaaaa
cgatgactgtatttttttcccttaactctgttagGGATTTGGATCCTGCTCCTAATCCACCTCATTTTCCATCGCA
TGTGATTAAAGCAACATTTGCCTATATCAGCAATTGTCATAAAACCAAGTTAAAAAGCATTTTAGAAATTCTTTC
CAAAAGCCCTgtaagtatacatgatgagtttaataatagaacattccttcttttttagctaaaaaaactttgtaa
atacatcttaaagaggaaaagtaaacaaatgaaaaatttatctcataattaaaaaggaaaacattcatttacaag
tttaaatggtattttacttgtcagcattaattgaaatatgttacatatgagaacagaatcttgtgacactttagt
gatatattagctcagggaatatatctacttttttcataggaatatactatttaattgtagtttactttctgaaaat
taaataaattggcaatagtttaagatagtaatttttcttaatgtaacattttgtacttgatatcaaacccaaatct
aaattctgttatttagttatttaaatataaaatgtgtaggtattcaaatatttgaagaaaaaatataaagtgta
tttattgtagccgagtatctaattaaacaagtttttactaaatctgtttatttttctagGATTCCTATCAGAAAAT
TCTTCTTGCCATATGTGAGCAAGCAGCTGAAACAAATAATGTTTATAAGAAGCACAGAATTCTTAAAATATATCA
CCTGTTTGTTAGTTTATTACTGAAAGATATAAAAAGTGGCTTAGGAGGAGCTTGGGCCTTTGTTCTTCGAGACGT
TATTTATACTTTGATTCACTATATCAACCAAAGgtaaataacatatttagaccaatatataagcagtctttctat
cctgttcttcctgtttttttgctttgttttgttttgttttgagacaaagactcactctgtccgcccaggctgggt
gcagtcacggctcactgcatcctcaacctcctgggttcagattatccttcctcctcaccctgccgggtagctggg
gctacaggagcataccaccatgcctgactaatttttttgtgtttttttgtagaggtgggg

Ex25

Ex26

Ex27

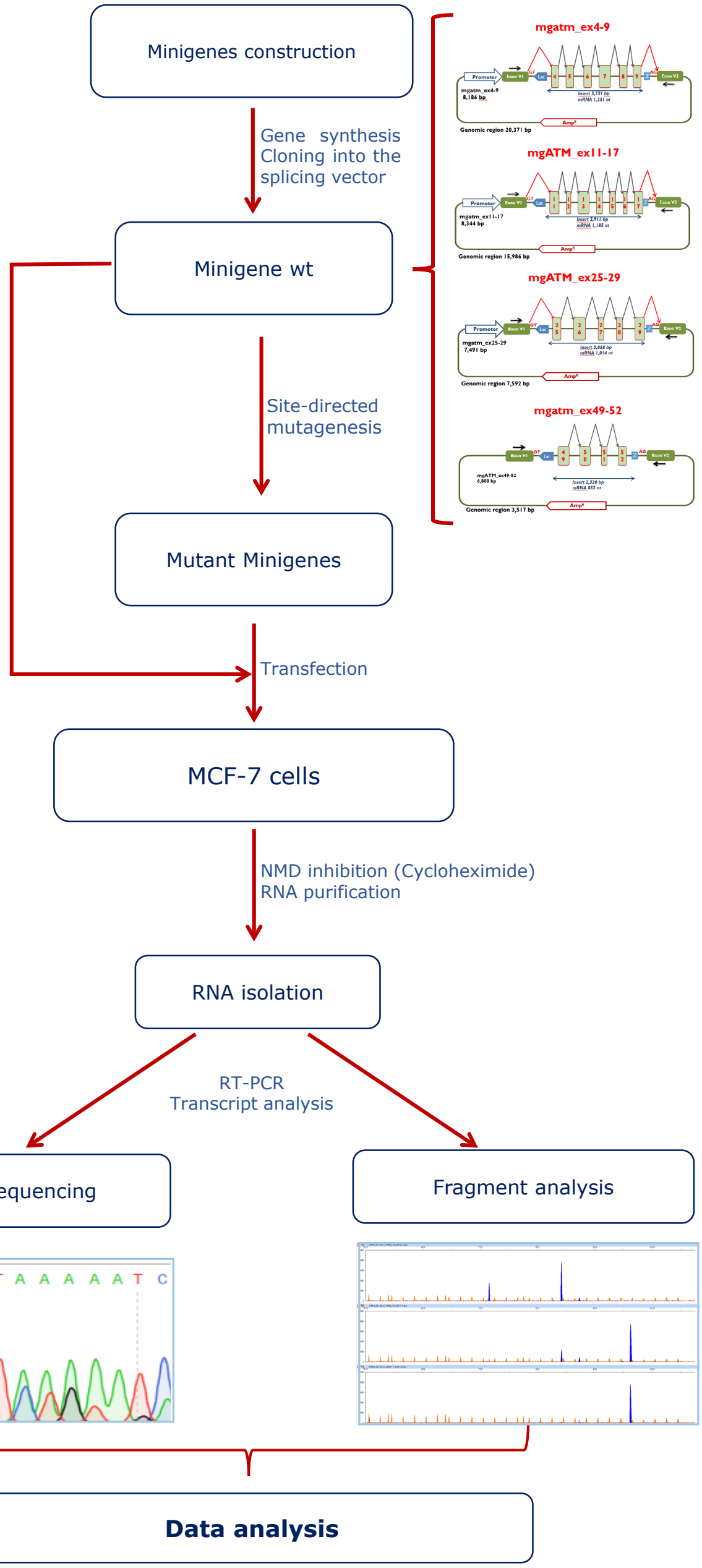
Ex28

Ex29

Supplementary Figure S1-D. Insert sequence of minigene mgATM_ex25-29. Exons 25 to 29 are indicated in upper case. Structure of the insert (3,058 bp): SacII - ivs24 (250 bp) – ex25 (170 bp) – ivs25-1 (200 bp) // ivs25-2 (200 bp) – ex26 (247 bp) – ivs26-1 (200 bp) // ivs26-2 (200 bp) – ex27 (116 bp) – ivs27-1 (200 bp) // ivs27-2 (200 bp) – ex28 (127 bp) – ivs28 (498 bp) – ex29 (200 bp) – ivs29 (250 bp) - SalI (shortened introns are indicated by a double slash //)

aaaatcacaaaagaaatctcataatgttttaagaaaatgtacgaatttgtgttgggccacattcaaagccgtcct
gggccacatgcggcccatgggccgtgggttggacaagtttgcaatagttcatataatttagctagcttttatatg
tatataagttaaattttagtgtattaccttaatttgagtgattctttagatgtatttagtatttgtaaatataat
ttaaattggttgtgttttcttgaagGCAGTAGAAGTTGCTGGAAATTATGATGGAGAAAGTAGTGATGAGCTAAG
AAATGGAAAAATGAAGGCATTTCTCTCATTAGCCCGGTTTTCAGATACTCAATACCAAAGAATTGAAAACTACAT Ex49
GAAATCATCGGAATTTGAAAACAAGCAAGCTCTCCTGAAAAGAGCCAAAGAGGAAGTAGGTCTCCTTAGGGAACA
TAAAATTCAGACAAACAGgtaactaggtttctacaagtgacaattttatgttcaccagttaactgagtgagtgtt
tttgcatagaaagagtgacttggtctttttatctgatatagttttgagctctaaaggtcggcttaactatatata
gattatcttggtctttgggttcttttcggtttttgttttttgttttttttttttgagacaaggtctcatcctttg
atgcttaggaaggtgtgtgaattgcacagttaagacaaaagtaagtttattccctttataatccttagaagtttg
ctttttttccctgggataaaaacccaacttttttcattaaatgttgtatatcatgtgtgattttgtagttctgtta
aagttcatggcttttgtgttttaccttaattattctatgcaagATACACAGTAAAGGTTCAGCGAGAGCTGGAGT
TGGATGAATTAGCCCTGCGTGCACTGAAAGAGGATCGTAAACGCTTCTTATGTAAAGCAGTTGAAAATTATATCA Ex50
ACTGCTTATTAAGTGGAGAAGAACATGATATGTGGGTATTCCGACTTTGTTCCCTCTGGCTTGAAAATTCTGGAG
TTTCTGAAGTCAATGGCATGATGAAGgcaagtgttactcagcccaatattctaccctgtgcttgaaaaacttaga
cataagccccttgatgtcaggaatcgtgtatacctctttgtattcctagcacttggtccagtgctctacacataa
gtagcattttgtagttttctaaactttgatccatatttaggattatttacaagttctagtcttgtttctacaaaa
gtaccaatgcattaatctagagtacccattagaaagaccttcagataagaaaagaaatgaaggaaaacaatatag
ttagtgaagttttgttaaccacttgtgctaatagaggagcactgtcttaaaataacttacttgcttagatgtgag
aatatttgaaatacctttgtttcttaattttgtgtctttttttttaatggtagAGAGACGGAATGAAGATTCCAACA
TATAAATTTTTGCCTCTTATGTACCAATTGGCTGCTAGAATGGGGACCAAGATGATGGGAGGCCTAGGATTTCAT Ex51
GAAGTCCTCAATAATgtaagtaaacctgaaaatcaaaccacaataattatttttattctattattactatatatt
atataaagtatatataccattccctctaagaaatggaaatacaaaatttgtatttttttgtcttctcacatcaca
taagttactcattttctctctctaattcctcataggcctctgccttttttctcacacatgcaggcatacacgctct
acccactgcagtatctagacagtaatacacattttaatgttaagcaaaatgaaaaatatggattatatttttttg
tttatttgcataaatctaatagttcttttcttacagCTAATCTCTAGAATTTCAATGGATCACCCCCATCACACT
TTGTTTATTATACTGGCCTTAGCAAATGCAAACAGAGATGAATTTCTGACTAAACCAGAGGTAGCCAGAAGAAGC Ex52
AGAATAACTAAAAATGTGCCTAAACAAAGCTCTCAGCTTGATGAGgtatttggattaaacatacgtacctttag
aagtgtgatattcagtctttcctagaatatttctttttaaaatcttgtgttattaagatgccatctaaaatcggt
tcaaggctggcacggtggctcacgcctgtaatcccagcactttgggaggctgaggcgggtggattacttgaggtc
agaagttcgagaccatcctggctgaccgacacagcaaaaccctgtctctactaaaaatgcaaaaaacagc

Supplementary Figure S1-E. Insert sequence of minigene mgATM_ex49-52. Exons 49 to 52 are indicated in upper case. Structure of the insert (2,320 bp): BamHI – ivs49 (250 bp) – ex49 (218 bp) – ivs49-1 (200 bp) // ivs49-2 (200 bp) – ex50 (208 bp) – ivs50-1 (200 bp) // ivs50-2 (200 bp) – ex51 (114 bp) – ivs51 (321 bp) – ex52 (159 bp) – ivs52 (250 bp) - EcoRI (shortened introns are indicated by a double slash //)



Supplementary Figure S2. Workflow of the minigene protocol. The basic assay includes the following steps: 1) Minigenes construction; 2) Site-directed mutagenesis; 3) Transfection of the wild type and mutant minigenes; 3) Inhibition of Nonsense-mediated decay and RNA purification; 4) Transcript sequencing and fragment analysis by fluorescent capillary electrophoresis; 5) Data interpretation.

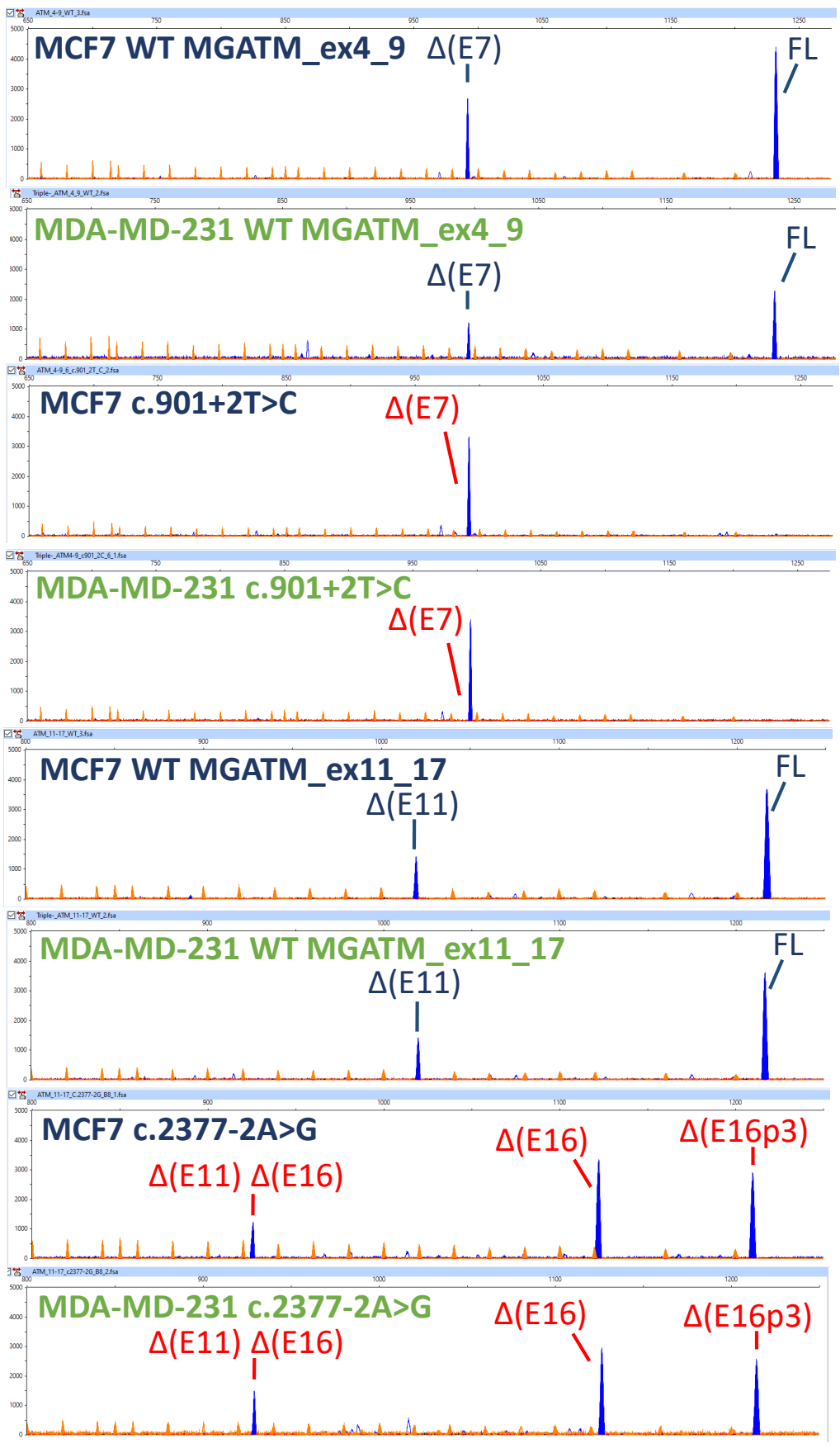
Supplementary Figure S3-A. Proposed decision tree assigning a PVS1_O/BP7_O code strength to mgATM minigene read-outs. Most *ATM* genetic variants tested in the present study produce two or more altered transcripts, and/or a substantial residual FL component (leaky variants). To tackle this issue, we propose the following approach: (i) read-out signal deconvolution into individual components (transcripts), (ii) assigning a pathogenic or bening code strenght to each individual transcript (acording to ACMG-AMP guidelines, see Supplementary Methods for further details), and (iii) apply *ad-hoc* combinatorial rules (taking into consideration the relative contribution of pathogenic and benign code strengths to the overall expression) to produce an overall PVS1_O/BP7_O code strength. Note that for leaky exon variants, the variant itself will be present in the FL transcript, determining its pathogenic or benign code strength.

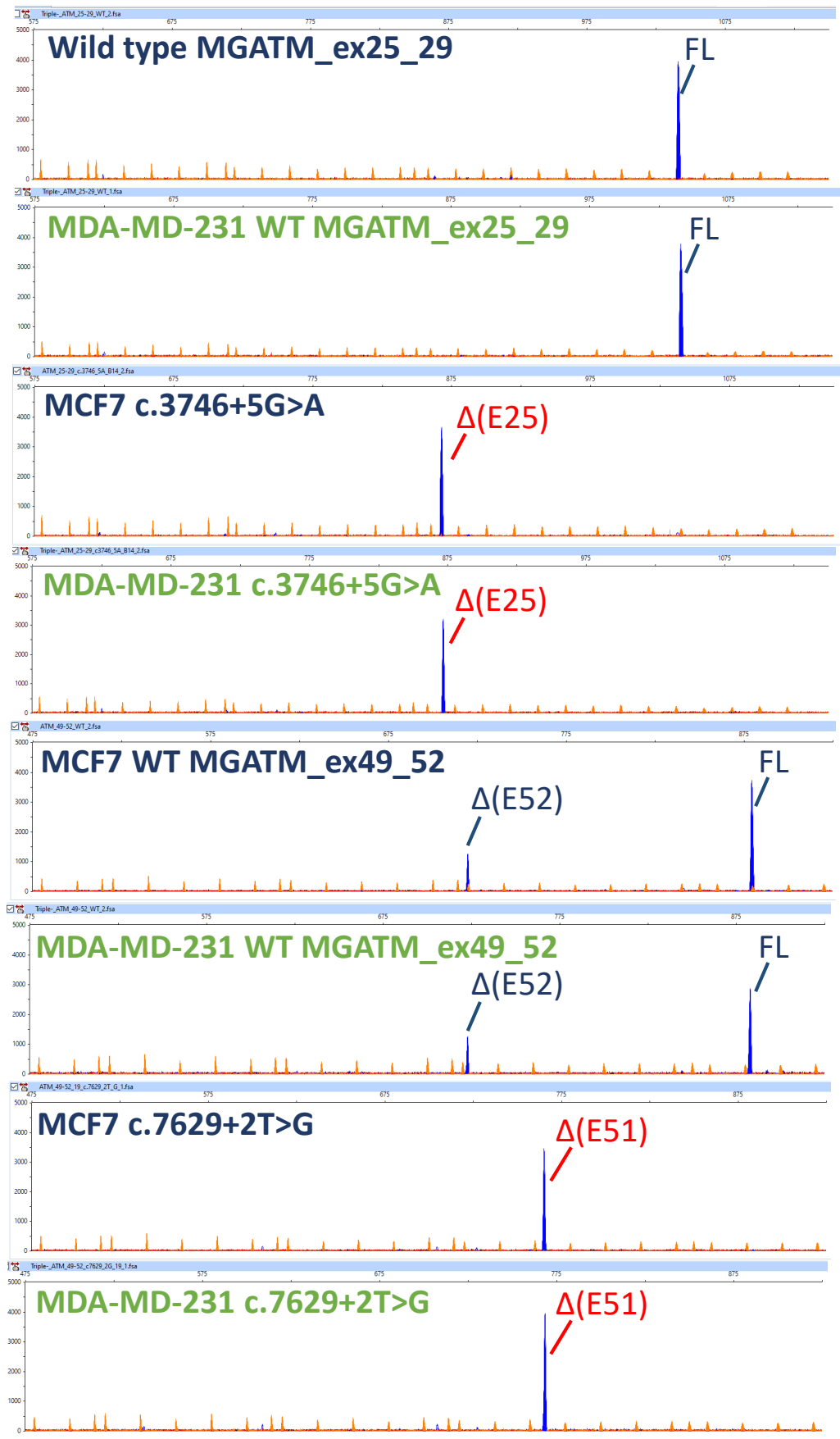
Variant-induced splicing event introducing a nonsense or frameshift alteration	**Pathogenic Very Strong** *PTC_NMD transcripts*	▼**(E5p4)** p.(Arg111Asnfs*4) Δ**(E5p1)** p.(Arg111Lysfs*5) Δ**(E7)** p.(Gln222Cysfs*3) Δ**(E7_E8)** p.(Gln222Phefs*34) Δ**(E8)** p.(Ala302Phefs*2) ▼**(E8q5)** p.(Arg111Asnfs*4) Δ**(E9)** p.(Val356Alafs*17) Δ**(E15p19)** p.(Cys633TrpFs*2) Δ**(E25)** p.(Val1193Ilefs*4) Δ**(E26)** p.(Ser1250Leufs*17) Δ**(E27)** p.(Ile1332Glyfs*7) Δ**(E27q1)** p.(Asp1371Ilefs*15) Δ**(E28)** p.(Asp1371Ilefs*38) ▼**(E28p8)** p.(Asp1371Leufs*18) Δ**(E28p53)** p.(Asp1371Asnfs*10) Δ**(E29)** p.(Asp1413Alafs*11) Δ**(E49)** p.(Ala2364Ilefs*13) Δ**(E49q38)** p.(Val2424Ilefs*13) Δ**(E50)** p.(Tyr2437Glufs*4) Δ**(E52p11)** p.(Leu2544Asnfs*23)
Variant-induced splicing event preserves reading-frame	**Pathogenic Very Strong** *in-frame exon skipping targeting the α-solenoid domain* *(deleted region: pathogenic missense variants reported and/or structural impact expected)* ***-or-*** in-frame exon skipping targeting the FATKIN domain	Δ**(E6_7)** p.(Glu166_Lys300del) Δ**(E7_9)** p.(Gln222_Trp412del) Δ**(E11)** p.(Pro537_Ser601del) Δ**(E12)** p.(Asn602_Cys633del) Δ**(E25_26)** p.(Val1193_Gln1331del) Δ**(E51)** p.(Arg2506_Asn2543del) Δ**(E52)** p.(Leu2544_Glu2596del)
	Pathogenic Strong *in-frame exon skipping targeting the α-solenoid domain* *(deleted region: no pathogenic missense variant reported and no obvious structural impact)*	Δ**(E5)** p.(Arg111_Glu166delinsLys) Δ**(E15)** p.(Ser751_Lys792del) Δ**(E16)** p.(Lys793_Leu822del) Δ**(E15_16)** p.(Ser751_Leu822del) Δ**(E28_29)** p.(Asp1371_Arg1479del)
	Pathogenic Moderate (PM4 rationale) *In-frame partial exon skipping* *(deleted region: no pathogenic missense reported and no obvious structural impact)*	Δ**(E25p159)** p.(Val1193_Glu1245del) Δ**(E26q120)** p.(Val1292_Gln1331del)
	Pathogenic Supporting *Very short in-frame with PP4 rationale applicable*	Δ**(E16p3)** p.(Lys793del)
Canonical transcripts	**Pathogenic Supporting** *PP3 applicable to the missense change*	**r.902G>A (p.Gly301Asp)** **r.1898G>U (p.Cys633Phe)**
	Benign Supporting *BP4 applicable to the missense change*	**r.496G>A (p.Glu166Lys)** **r.3577G>C (p.Val1193Leu)**
	Benign Strong	***r.903U>G (pGly301=)*** ***WT sequence***

Supplementary Figure S3-B. Pathogenic/Benign code strengths applicable to individual transcripts produced by mgATM minigenes.
We have adapted the PVS1 decision tree rationale proposed by the ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI) (**PMID: 30192042**) to the specific purpose of determining the strength of the loss-of-function evidence for all mgATM transcripts. For in-frame splicing, see Supplementary Methods for further details.

variant	PVS1_O/BP7_O	mgATM read-out	
		Transcript 1	Transcript 2
c.901+2T>C	**PVS1_O**	Δ(E7) (100%) **P_VS**	
c.1898+3A>T	**BP7_O_Strong**	FL WT (100%) **B_S**	
c.496G>A	**BP7_O_P**	FL p.(Glu166Lys) (72%) **B_P**	Δ(E7) (28%)
Simple read-outs. c.901+2T>C produces only exon 7 skipping (PTC-NMD) transcripts, while c.1898+3A>T produces only FL (WT sequence) transcripts. In both cases, code strength is straightforward. c.496GA produces FL and exon 7 skipping transcripts. Yet, the contribution of Δ(E7) to the overall expression (28%) is very similar to that observed in the corresponding mgATM ex4-9 WT. For that reason, Δ(E7) is not considered a specific c.496G>A outcome. Therefore, mgATM analysis is interpreted as as proving that c.496G>A is a bona-fide missense variant (i.e. non-spliceogenic missense variant). Based on REVEL/BayesDel scores and structural analysis, we have assigned a benign supporting strength evidence to the FL transcript coding for the missense change. To reflect the splicing assay performed, we have assigned a BP7_O_P code. Alternatively, the variant can be assigned a BP4. Note that it is evidence strength, and not code type, that contributes to the final classification.			
c.4109+1G>T	**PVS1_O**	(E27) (33%) P_VS	Δ(E27q1) (67%) P_VS
c.3993+1G>A	**PVS1_O_M**	Δ(E26) (33%) P_VS	Δ(E26q120) (67%) P_M
Complex read-outs supporting pathogenic. c.3993+1G>A produces two different transcripts, both supporting pathogenic with very strong strenght. PVS1_O evidence strenght is straightforward. c.3993+1G>A produces two different transcripts suporting pathogenic (albeit with different strenght). Very strong represents 33% of the overall signal, while Moderate represents 67%. Based on SVI conservative recommendations, we select Moderate as appropiate strenght for overall PVS1_O evidence strenght.			
c.7788+6T>G	PVS1_O N/A BP7_O N/A	FL WT (51%) B_S	Δ(E52) (49%) P_VS
c.1898+2T>G	PVS1_O N/A BP7_O N/A	Δ(E12) (87%) P_VS	FL WT (13%) B_S
c.3993+5G>T	PVS1_O N/A BP7_O N/A	FL WT (76%) B_S	Δ(E26q120) (23%) P_M
Conflicting read-outs: c.7788+6T>G is a leaky variant producing similar amounts of FL transcripts (supporting benign) and exon 52 skipping transcripts that preserve the reading-frame, but target the critical FAT-HRD domain (supporting pathogenic). Based on the conflicting evidence, neither PVS1_O nor BPS7_O evidence is applied to c.7788+6T>G. Similarly, c.1898+2T>G and c.3993+5G>T are leaky, producing pathogenic supporting transcripts, but also benign, supporting FL transcripts. For c.1898+2T>G, the signal supporting pathogenic (87% of the overall signal) is much higher than the signal supporting benign (13%). For c.3993+5G>T, the opposite is true. Yet, based on our conservative *ad-hoc* rule (neither pathogenic nor benign signal reaches 90%), both read-outs are considered conflicting and neither PVS1_O (at any strength) nor BP7_O (at any strength) are applied.			

Supplementary Figure S3-C. Pathogenic/Benign annotation of *ATM* transcripts. Representative examples of overall PVS1_O/BP7_O code strengths after combining pathogenic or benign code strengths assigned to individual transcripts. If both pathogenic and benign supporting transcripts represent >10% of the overall expression level (Table 1), the minigene read-out is considered a conflicting evidence, and neither PVS1_O nor BP7_O codes are incorporated into the final classification scheme. If minigene read-out supports pathogenic, but different transcripts support different strength, the lowest strength (representing ≥10% of the overall expression) is incorporated into the final classification scheme. For clarity, code strengths have been color-coded.





Supplementary Figure S4. Splicing functional assays of four selected splice-site variants and WT minigenes in MDA-MB-231 (green) and MCF-7 (blue) cells. Cell growth, transfections and cDNA amplifications were conducted as described in Materials and Methods. FAM-labelled products (blue peaks) were run with LIZ1200 (orange peaks) as size standard. For transcript descriptions see Supplemental Table S3, FL, Full-length transcript.

Clinical Data in 3 *ATM* leaky variants (leakiness supported by mgATM data) support a dosage-sensitive expression model

***ATM* c.1898+2T>G**

Somehow unexpectedly, this variant produces a non-negligible proportion of full-length transcripts (13%) supporting benign. The remaining expression (87%) supports pathogenic. Neither pathogenic nor benign supporting transcripts contribute ≥90% to the overall expression. Based on that (Supplementary Figure S3A), mgATM data does not contribute to the final classification of this variant. Yet, we have been able to classify c.1898+2T>G as **LP**, largely due to a very strong clinical evidence **PM3** (Supplemental Table S2). Further, ClinVar consistently reports the variant as **LP/P** (criteria provided, nine submitters, no conflicts, VCV000141939.22). Taking together, we infer that leaky alleles producing ≤13% are pathogenic.

***ATM* c.496+5G>A**

Previous studies in RNA from carriers have shown that *ATM* variants c.332-1G>A and c.496+5G>A cause identical in-frame exon 5 skipping (Laake et al., 2000; Dörk et al., 2004b). The mgATM data confirms these findings but, contrary to previous RNA studies in carriers, shows some differences; c.496+5G>A, but not c.332-1G>A, is a leaky variant expressing a significant proportion of full-length transcripts (~25% of the overall expression). Interestingly, c.332-1G>A has been reported in classical A-T patients (Laake et al., 2000), while c.496+5G>A causes a form of A-T with slow progression (Dörk et al., 2004b). Taken together, the data suggests that in-frame exon 5 skipping is, indeed, damaging and that milder phenotype observed in c.496+5G>A is probably caused by leakiness.

***ATM* c.1066-6T>G**

Currently, there is a conflicting interpretation for the pathogenicity of this variant in ClinVar (10 submitters supporting benign/likely benign, and 6 submitters supporting uncertain significance; accession VCV000003038.34). First described as pathogenic based on: identification in the homozygous state in a German A-T patient and a RNA study in carriers supporting that it is a truncating variant causing exon 9 skipping (Dörk et al., 2001). Subsequently, a small case-control study suggested that c.1066-6T>G is not associated with breast cancer risk (and that it is too common in control population to be pathogenic), but confirmed that it is a truncating variant causing exon 9 skipping (Soukupova et al., 2008). Later, Tavtigian and col. informed on a second *ATM* variant able to explain the A-T phenotype in the original German patient (Tavtigian et al., 2009). Ding and col. conducted a large meta-analysis confirming lack of association of *ATM* c.1066-6T>G with breast cancer risk (Ding et al., 2011). BRIDGES case-control data further confirms lack of association (see Supplementary Figure 2). mgATM analysis identifies a high proportion of truncating exon 9 skipping transcripts (as previously described in RNA from carriers), a small proportion of truncating transcripts skipping exons 7 to 9 (not described in previous studies) and, more relevant, a significant proportion of full-length transcripts (~30% of the overall expression). Previous studies in RNA from carriers were not able to detect this leaky effect that, most likely, explain the non-pathogenic nature of the c.1066-6T>G variant. Taken together, the data suggests that leaky variants producing ≥30% of full-length transcripts are benign.

Assigning an overall PVS1_O/BP7_O code strength to ATM leaky variant producing pathogenic supporting transcripts

If full-length transcripts contribute <13% to the overall expression → PVS1_O_variable strength
(variable strength depending on the nature and relative expression of pathogenic supporting transcripts)

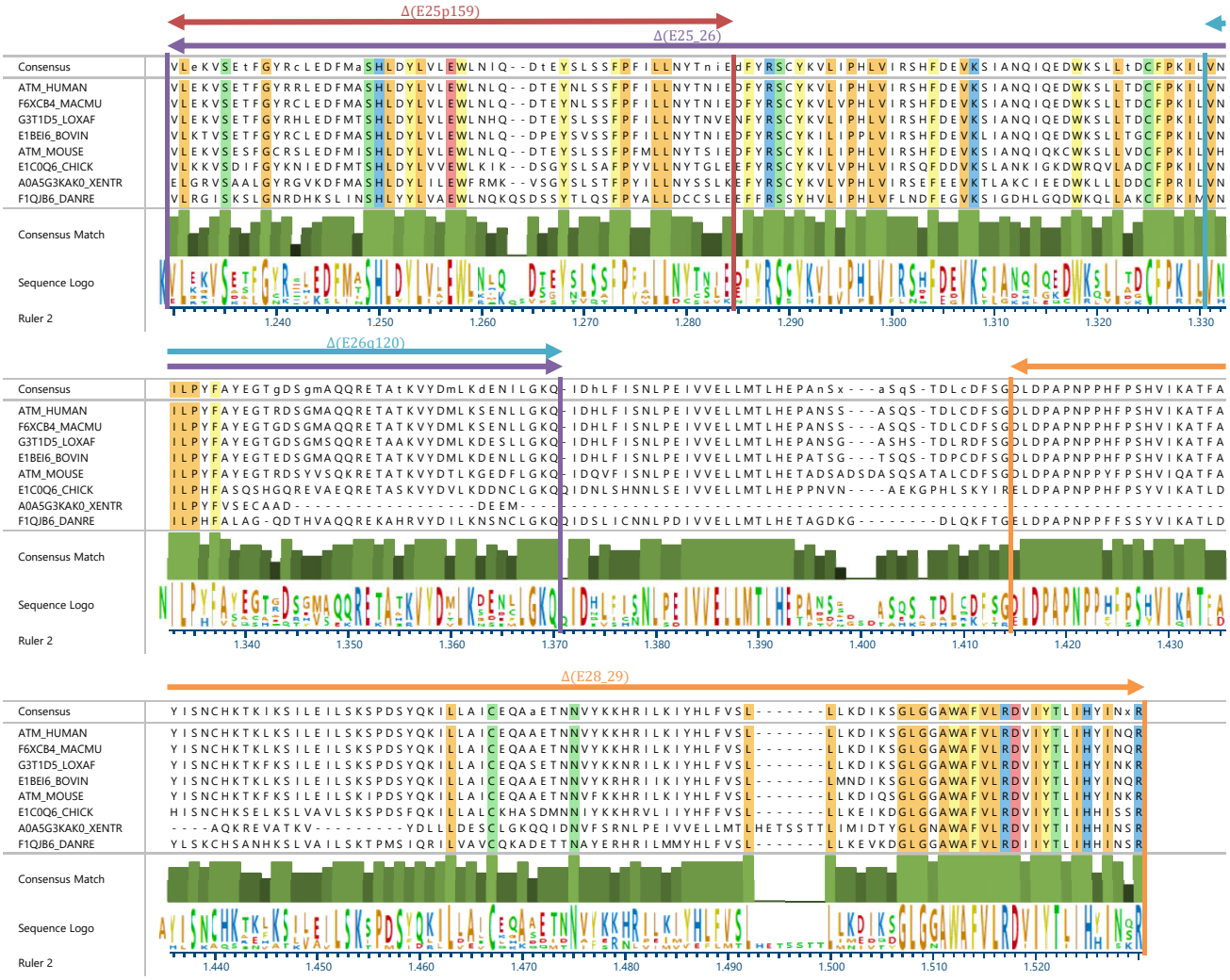
If full-length transcripts contribute >30% to the overall expression→ BP7_O_Strong
(regardless of the code strength supported by pathogenic supporting transcripts)

If full-length expression in between (13%-30% range) → PVS1_O N/A and BP7_O N/A

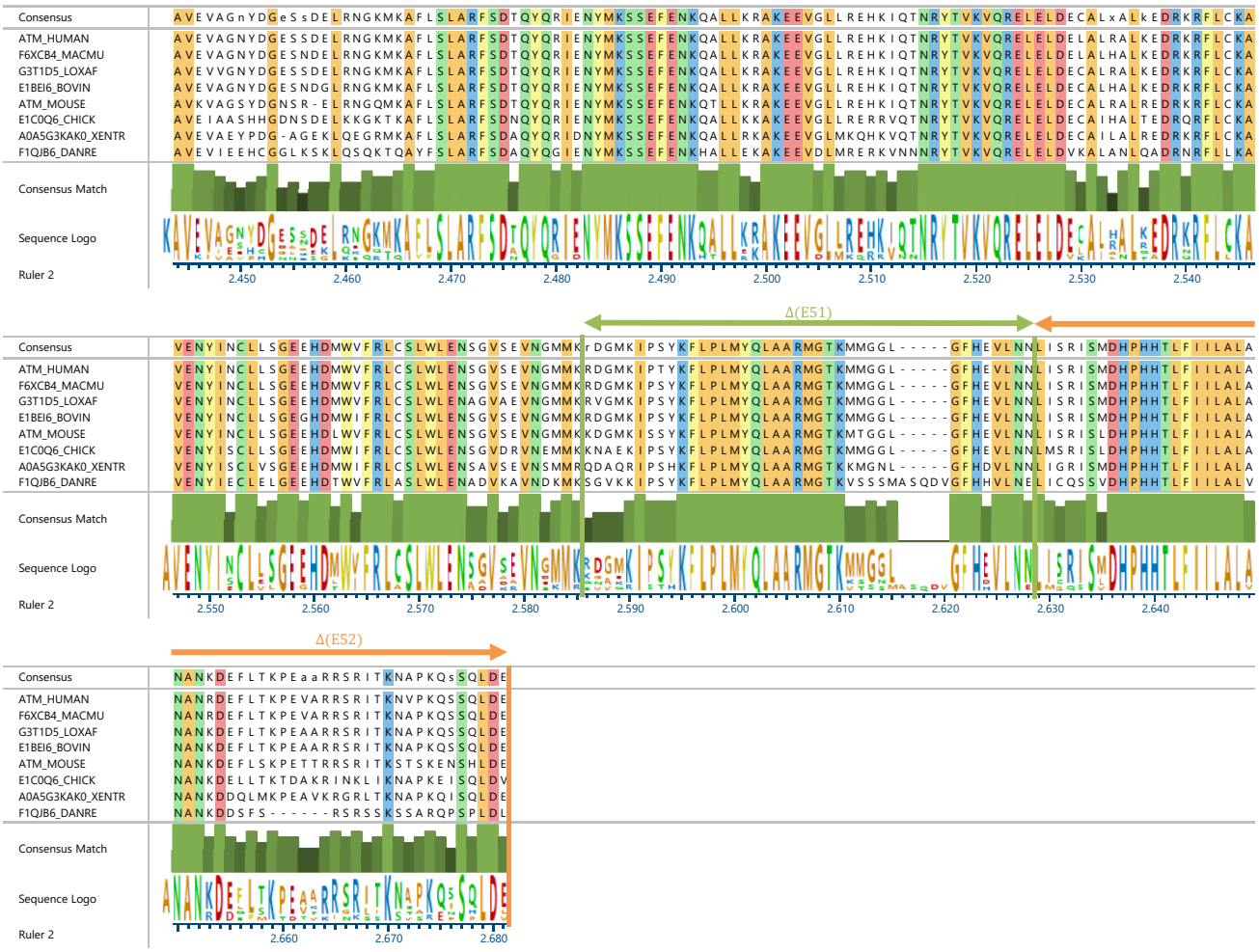
Supplementary Figure S5. Proposed "dosage-sensitive expression model" and tentative integration into the ACMG/AMP classification scheme to assigning a PVS1_O/BP7_O code strength to *ATM* leaky variants. Taken together, the evidence of leakiness in c.1898+2T>G, c.496+5G>A and c.1066-6T>G variants, together with its associated clinical features (top panel), suggests a tentative model in which: (i) Any ATM allele producing ≥30% of full-length transcripts might be benign (i.e. no risk associated), (ii) Any *ATM* alleles producing loss-of-function transcripts (and ≤13% of full-length transcripts) might be overtly pathogenic, and (iii) *ATM* alleles producing 14%-30% of full-length transcripts are in a grey zone, and might be associated with intermediate effects. At any rate, the model has obvious implications for the classification of spliceogenic leaky variants (bottom panel). At present, this is just a tentative model that needs to be confirmed (or refuted) by additional clinical evidences in well-characterized splicegenic leaky variants (note that, even if the model proves true, cut-off percentages as determined here by RT-PCR and capillary electrophoresis in mgATM read-outs are not necessarily applicable to splicing assays based on different methodologies).

Supplementary Figure S9-A. Alignment and amino acid conservation of deleted in-frame sequences corresponding to the anomalous *ATM* transcripts Δ(E5) and Δ(E7_9). The deleted in-frame regions are marked in Green and purple stripes and arrows. Protein sequences were aligned using the Align tool of the Uniprot database (https://www.uniprot.org/align/).The alignment file was visualized with MegAlign Pro version 15.0.0 of DNASTAR's Lasergene software.The conserved residues are highlighted (nonpolar amino acids in orange, polar amino acids in green, polar basic amino acids in blue and polar acidic amino acids in red) and the color intensity of green bars above the protein sequence indicates the degree of conservation of each amino acid. Organisms: Human (Homo sapiens); Rhesus Macaque (Macaca mulatta: MACMU);African elephant (Loxodonta africana: LOXAF); Bovine (Bos taurus: BOVIN); Mouse (Mus musculus: MOUSE); Chicken (Gallus gallus: CHICK);African Clawed Frog (Xenopus tropicalis: XENTR); Zebrafish (Danio rerio: DANRE).

Supplementary Figure S9-B. Alignment and amino acid conservation of deleted in-frame sequences corresponding to the anomalous *ATM* transcripts Δ(E11), Δ(E12), Δ(E15), Δ(E15_16), Δ(E16p3) and Δ(E16). The deleted in-frame regions are marked in blue, orange, green, red and purple stripes and arrows. Protein sequences were aligned using the Align tool of the Uniprot database (https://www.uniprot.org/align/).The alignment file was visualized with MegAlign Pro version 15.0.0 of DNASTAR's Lasergene software.The conserved residues are highlighted (nonpolar amino acids in orange, polar amino acids in green, polar basic amino acids in blue and polar acidic amino acids in red) and the color intensity of green bars above the protein sequence indicates the degree of conservation of each amino acid. Organisms: Human (Homo sapiens); Rhesus Macaque (Macaca mulatta: MACMU);African elephant (Loxodonta africana: LOXAF); Bovine (Bos taurus: BOVIN); Mouse (Mus musculus: MOUSE); Chicken (Gallus gallus: CHICK);African Clawed Frog (Xenopus tropicalis: XENTR); Zebrafish (Danio rerio: DANRE).



Supplementary Figure S9-C. Alignment and amino acid conservation of deleted in-frame sequences corresponding to the anomalous *ATM* transcripts Δ(E25_26), Δ(E25p159), Δ(E26q120) and Δ(E28_29). The deleted in-frame regions are marked in blue, orange, red and purple stripes and arrows. Protein sequences were aligned using the Align tool of the Uniprot database (https://www.uniprot.org/align/).The alignment file was visualized with MegAlign Pro version 15.0.0 of DNASTAR's Lasergene software.The conserved residues are highlighted (nonpolar amino acids in orange, polar amino acids in green, polar basic amino acids in blue and polar acidic amino acids in red) and the color intensity of green bars above the protein sequence indicates the degree of conservation of each amino acid. Organisms: Human (Homo sapiens); Rhesus Macaque (Macaca mulatta: MACMU);African elephant (Loxodonta africana: LOXAF); Bovine (Bos taurus: BOVIN); Mouse (Mus musculus: MOUSE); Chicken (Gallus gallus: CHICK);African Clawed Frog (Xenopus tropicalis: XENTR); Zebrafish (Danio rerio: DANRE).



Supplementary Figure S9-D. Alignment and amino acid conservation of deleted in-frame sequences corresponding to the anomalous *ATM* transcripts Δ(E51) and Δ(E52). The deleted in-frame regions are marked in orange and Green stripes and arrows. Protein sequences were aligned using the Align tool of the Uniprot database (https://www.uniprot.org/align/).The alignment file was visualized with MegAlign Pro version 15.0.0 of DNASTAR's Lasergene software.The conserved residues are highlighted (nonpolar amino acids in orange, polar amino acids in green, polar basic amino acids in blue and polar acidic amino acids in red) and the color intensity of green bars above the protein sequence indicates the degree of conservation of each amino acid. Organisms: Human (Homo sapiens); Rhesus Macaque (Macaca mulatta: MACMU);African elephant (Loxodonta africana: LOXAF); Bovine (Bos taurus: BOVIN); Mouse (Mus musculus: MOUSE); Chicken (Gallus gallus: CHICK);African Clawed Frog (Xenopus tropicalis: XENTR); Zebrafish (Danio rerio: DANRE).