

The *ATM* Splice-ome: Charting the multifaceted splicing landscape of *ATM*



Ambry Genetics
A TEMPUS COMPANY

Laura Hudish¹, Terra Brannan¹, Kristin Kohler¹

¹. Ambry Genetics, Aliso Viejo, CA, USA

Email: lhudish@ambrygen.com



BACKGROUND

Pathogenic variants in the DNA damage response gene *ATM* are associated with hereditary cancer susceptibility¹ and ataxia-telangiectasia². While many pathogenic variants directly result in protein truncation, disruption of splicing represents an additional and underrecognized mechanism of pathogenicity. **Splice-altering variants may arise not only at canonical splice sites but also from exonic substitutions that alter splicing regulatory elements or intronic variants that activate pseudoexonization.** Because exonic splice variants can be misinterpreted as missense changes and deep intronic variants may escape routine annotation, the landscape of *ATM* variants that disrupt splicing remains incompletely characterized in clinical diagnostic cohorts.

AIM: To determine the types of splicing variants detected in *ATM* and the mechanisms and types of splice disruption

1. Marabelli M. et al. "Penetrance of ATM Gene Mutations in Breast Cancer: A Meta-Analysis of Different Measures of Risk". *Genet. Epidemiol.* 2016;40:425–431. doi: 10.1002/GEPI.21971.
2. K. Savitsky et al., "A single ataxia telangiectasia gene with a product similar to PI-3 kinase," *Science*, vol. 268, pp. 1749–1753, 1995, doi: 10.1126/science.7792600

METHODS

Variant Identification: Likely pathogenic (LP) and pathogenic (P) single nucleotide variants with a demonstrated splice impact were analyzed. Variants identified during routine clinical diagnostic testing included canonical intronic variants (+/- 1 to 2), first/last exonic nucleotide variants, exonic missense and synonymous variants, and non-canonical intronic variants including intronic (+/- 3 to +/- 5), intronic OOR (+/- 6 to +/- 30) and deep intronic (+/- >30). Internal RNA studies³ were performed for every variant and only those with strong RNA data supporting pathogenicity were included. SpliceAI⁴ predictions with scores >0.2 across a 500bp flanking region around each variant were analyzed and compared to the observed events.

Statistical analyses: Exon-level variant enrichment was tested using exact Poisson rate tests against a uniform global density baseline (0.028 variants/NT), with Benjamini-Hochberg FDR correction across all 61 coding exons (CDS).

3. Horton, Carolyn et al. "Mutational and splicing landscape in a cohort of 43,000 patients tested for hereditary cancer," *NPJ genomic medicine* vol. 7, 1 49. 25 Aug. 2022, doi:10.1038/s41525-022-00323-y
4. Jaganathan, Kishore et al. "Predicting Splicing from Primary Sequence with Deep Learning," *Cell* vol. 176, 3 (2019): 535-548.e24. doi:10.1016/j.cell.2018.12.015

TAKE HOME POINTS

- **Diagnostic Gap:** Over 30% of splice-altering LP/P *ATM* variants are non-canonical and likely missed without RNA analysis.
- **SpliceAI Utility:** SpliceAI demonstrates high predictive accuracy, successfully identifying the vast majority of these non-canonical events.
- **Variant Hotspots:** Coding exons 20 and 52 exhibit the highest variant densities (13.16 and 6.29 var/100bp), with CDS52 containing the most exonic missense splice variants.
- **Functional Impact:** Full exon skipping (ESF) is the dominant splicing consequence, representing 55.3% of all observed events.

RESULTS

ATM Coding Exons 20 and 52 Represent Major Spliceogenic Hotspots

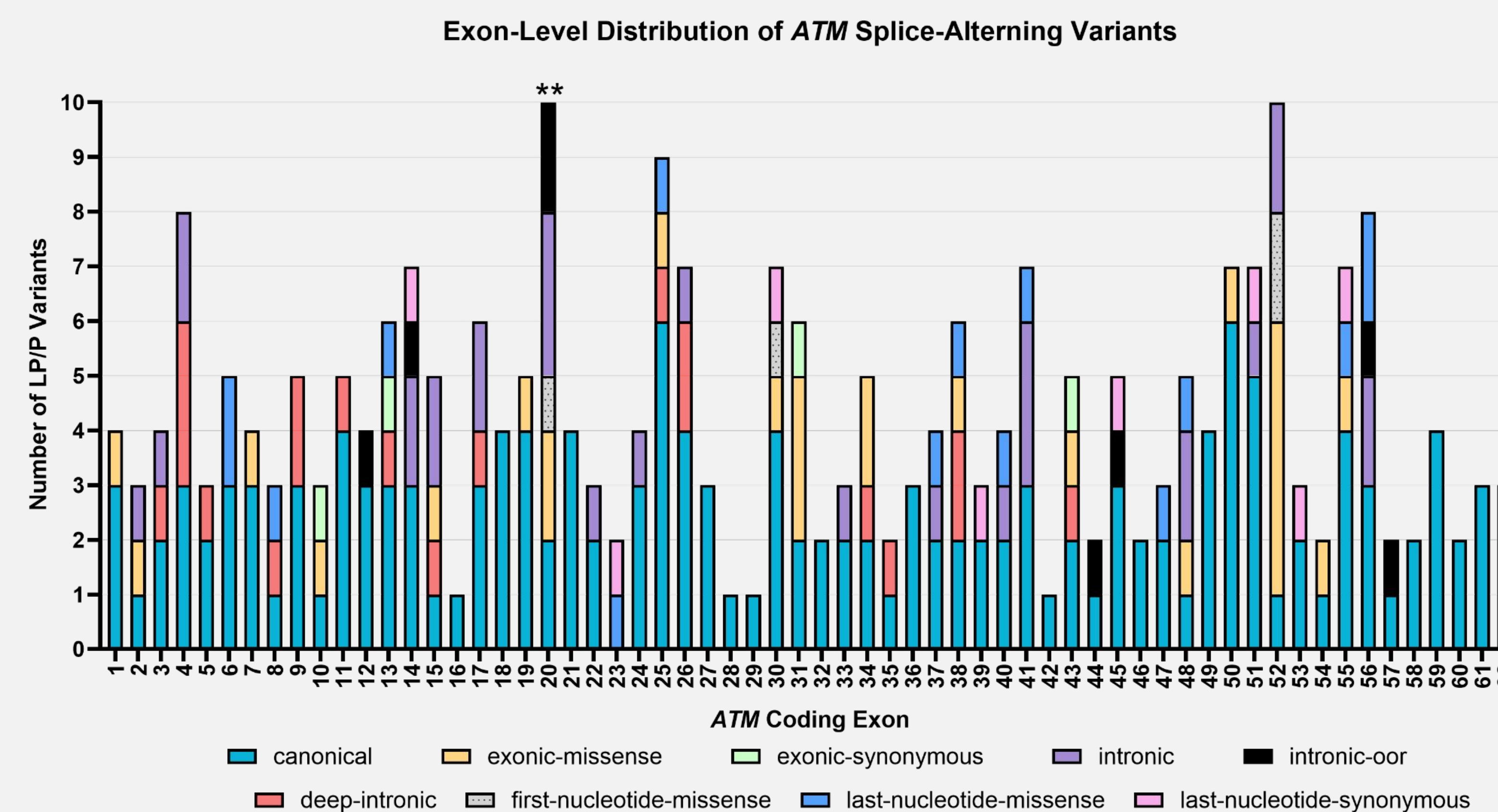
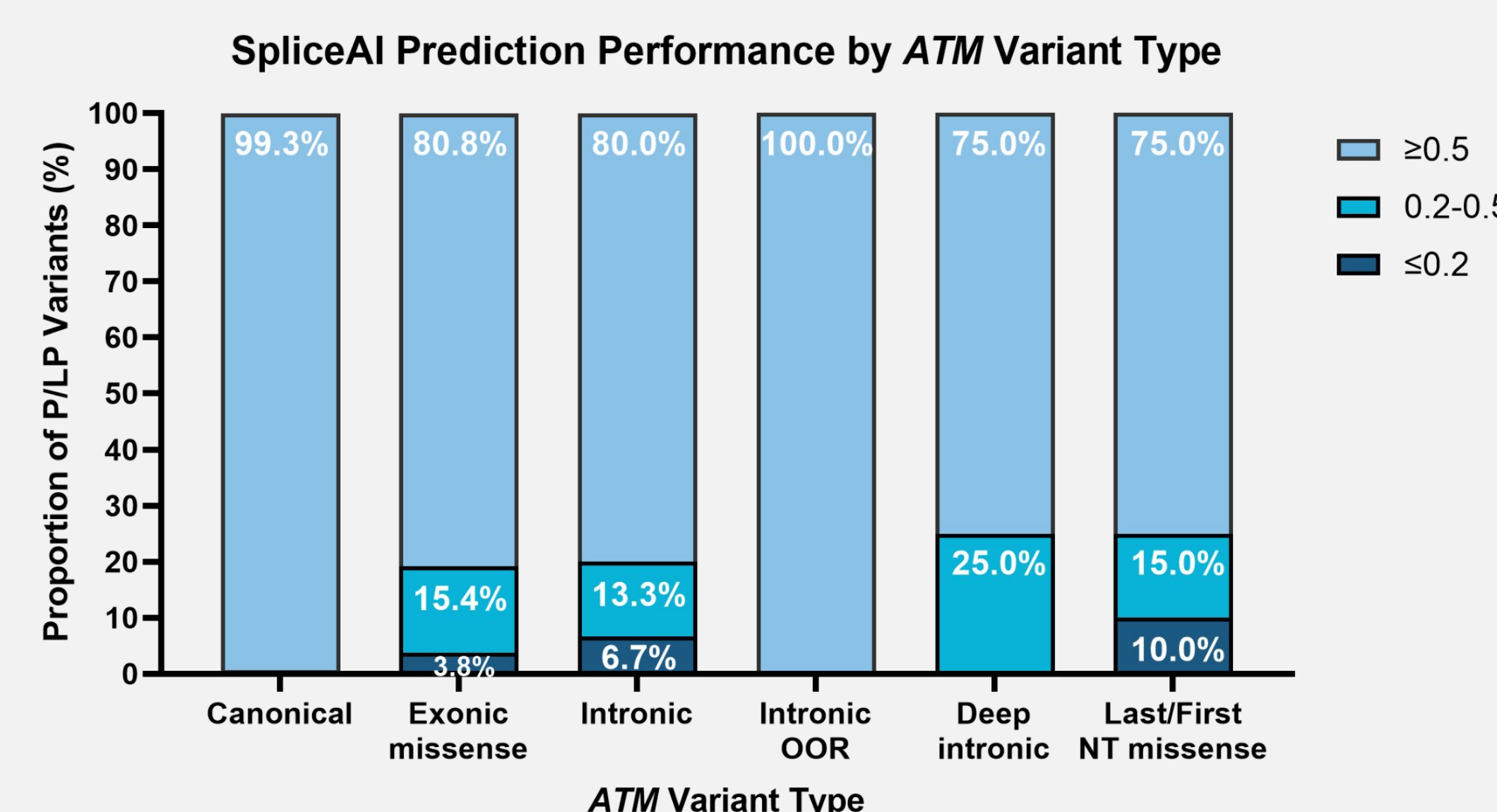


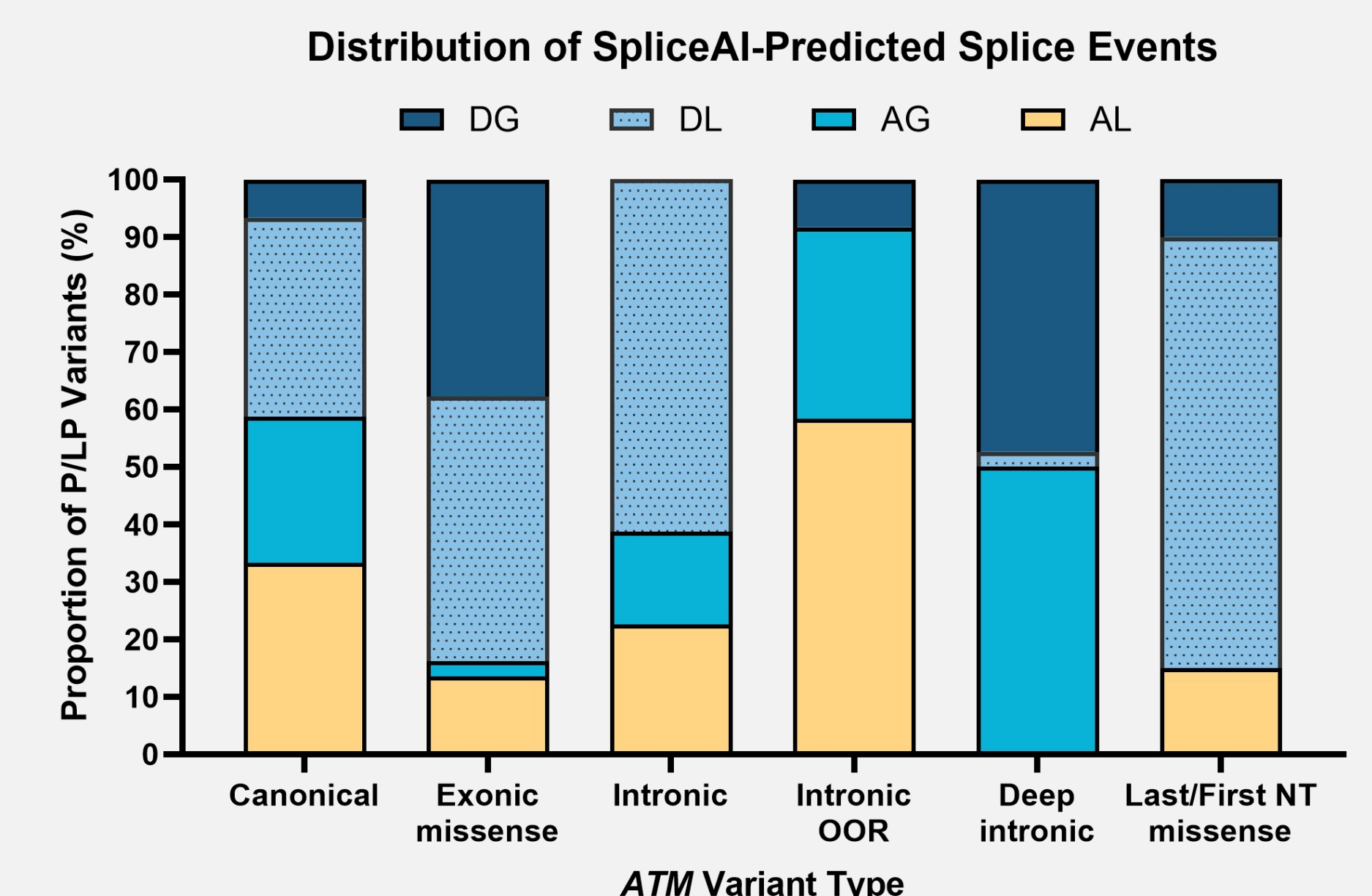
Figure 1. Distribution of splicing variants across the *ATM* gene. A total of 253 eligible *ATM* splicing variants were identified in the patient cohort, consisting of 72.8% VLP and 27.2% P classifications. Canonical variants (+/-1,2bp) represented the most frequent category, accounting for 56% (143) of the total. Eligible variants were mapped to their respective exons. Canonical variants were detected across all exons except coding exon 1 (CDS1). CDS20 and CDS52 exhibited the highest number of splice-impacting variants, with 10 identified in each. **After normalizing for exon length, CDS20 exhibited a statistically significant enrichment of splice-altering variants ($p<0.01$).** Overall, there was a non-uniform distribution across coding exons (χ^2 ; $p<0.0001$), with shorter coding exons displaying significantly higher variant density per base pair ($r_s=-0.554$; $p<0.00001$).

SpliceAI Accurately Predicts Diverse Non-Canonical Splicing Events

A.



B.



C.

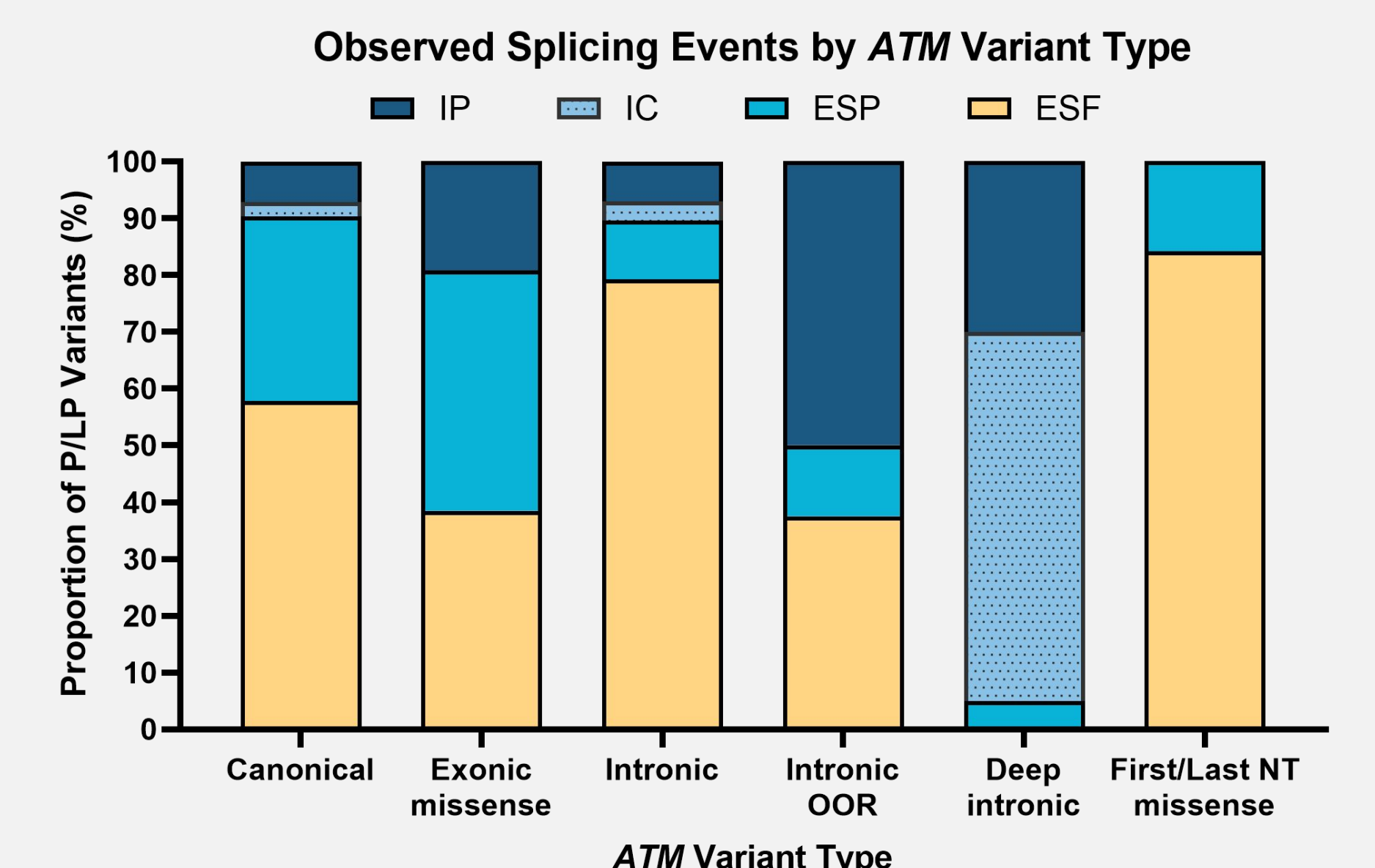


Figure 2. Comparison of Predicted and observed splicing events in *ATM*. (A) SpliceAI successfully predicted splicing events in the majority of variants. **Sensitivity was maintained even at the more stringent 0.5 score cutoff.** (B) Canonical variants have a comparatively uniform distribution across different event types, while exonic missense variants were primarily linked to donor-site impacts, with donor loss (DL) (45.9%) and donor gain (DG) (37.8%). Intronic alterations demonstrated a marked preference for DL (61.3%) and acceptor loss (AL) (22.6%) consequences. (C) Canonical variants show balanced splice-site loss/gain predictions aligned with the observed full (ESF) and partial (ESP) exon skipping. Deep-intronic variants show high predicted acceptor/donor gains that result in cryptic intron inclusion (IC) or partial intron retention (IP). Exonic missense variants primarily impact donor sites (DL/DG), correlating with predominant full and partial exon skipping (ESF/ESP). **Collectively, the high degree of concordance between predicted mechanisms and functional outcomes indicates that SpliceAI effectively captures the complex and multifaceted splicing landscape of the *ATM* gene.**

ATM Splice-ome: Putting the Pieces Together

