

Ensemble AI Screening of 7,467 TP53 Variants: Redefining Clinical Pathogenicity for Variants of Uncertain Significance

Mahad Asif

9th Grade Computer Science Student

1 Abstract

2 **Background.** TP53 is the most frequently mutated
3 gene in human cancer, yet over 1,200 missense vari-
4 ants in ClinVar remain classified as Variants of Un-
5 certain Significance (VUS), limiting their clinical
6 utility in precision oncology.

7 **Methods.** We applied two orthogo-
8 nal deep-learning models—Meta’s ESM-2
9 (esm2_t33_650M_UR50D; 650 M parameters)
10 and DeepMind’s AlphaMissense—to score the com-
11 plete landscape of 7,467 possible single-amino-acid
12 substitutions in TP53 (UniProt P04637). Of these,
13 1,199 ClinVar VUS were matched to AlphaMissense
14 predictions and cross-referenced. Structural valida-
15 tion was performed against PDB 1TUP (Cho et al.,
16 1994).

17 **Results.** The two models showed strong anti-
18 correlation (Pearson $r = -0.706$; Spearman $\rho =$
19 -0.714), supporting complementary predictive ca-
20 pacity. A total of 349 VUS were flagged as high-risk
21 by both models (ESM-2 log-likelihood ratio ≤ -4.0
22 and AlphaMissense pathogenicity > 0.564), while
23 438 were concordantly classified as low-risk. We
24 highlight five variants in the DNA-binding domain
25 (L257R, V157D, R248P, C176R, R280I) with ex-
26 treme concordant scores (ESM-2 LLR ≤ -12.5 ; Al-
27 phaMissense ≥ 0.99) and structurally validated dam-
28 age mechanisms including loss of DNA contact, zinc
29 coordination abolishment, and hydrophobic core dis-
30 ruption.

31 **Conclusions.** Ensemble AI scoring can system-
32 atically reclassify TP53 VUS at scale. The 349
33 high-confidence pathogenic variants identified here
34 warrant prioritized functional validation and may in-
35 form germline testing guidelines in Li–Fraumeni syn-
36 drome and somatic profiling in clinical oncology.

37 **Keywords:** TP53, variants of uncertain significance,

38 ESM-2, AlphaMissense, pathogenicity prediction,
39 precision oncology, protein language model

1 Introduction

41 The *TP53* gene encodes the tumour protein p53,
42 a transcription factor widely characterised as the
43 “Guardian of the Genome” for its central role in
44 maintaining genomic integrity [Lane, 1992]. In re-
45 sponse to genotoxic stress, p53 activates cell-cycle ar-
46 rest, DNA repair, senescence, and apoptosis through
47 sequence-specific DNA binding at target promoters
48 [Vogelstein et al., 2000, Levine & Oren, 2009]. Loss-
49 of-function mutations in TP53 are the single most
50 common genetic alteration across all human cancers,
51 observed in over 50% of solid tumours [Kandoth
52 et al., 2013], and germline TP53 mutations are the
53 molecular basis of Li–Fraumeni syndrome, a hered-
54 itary cancer predisposition disorder conferring near-
55 complete lifetime cancer penetrance [Malkin et al.,
56 1990].

57 The clinical interpretation of TP53 missense vari-
58 ants has been greatly advanced by large-scale reposi-
59 tories such as ClinVar [Landrum et al., 2018] and the
60 IARC TP53 Database [Bouaoun et al., 2016]. How-
61 ever, a substantial fraction of observed variants—
62 over 1,200 unique missense substitutions in ClinVar
63 at the time of this analysis—remain classified as Vari-
64 ants of Uncertain Significance (VUS). The VUS des-
65 ignation creates a “diagnostic grey zone” for clin-
66 icians: these variants cannot be used to guide treat-
67 ment decisions, genetic counselling, or cascade fam-
68 ily testing, even when the variant resides in a func-
69 tionally critical domain [Richards et al., 2015].

70 The bottleneck in VUS resolution is fundamen-
71 tally one of evidence accumulation. Under the
72 ACMG/AMP framework [Richards et al., 2015], re-
classification requires convergent evidence from pop-

ulation frequency data, *in silico* prediction, functional assays, co-segregation studies, and *de novo* occurrence. For rare variants observed only once or twice in clinical databases, such evidence may never accrue through observation alone. Large-scale functional assays, such as the saturation mutagenesis screen of TP53 by Giacomelli et al. [2018], have made important contributions but remain labour-intensive and costly.

Recent advances in protein language models (pLMs) and structure-based pathogenicity predictors offer a scalable complement to experimental approaches. Meta’s ESM-2 [Lin et al., 2023] is a transformer-based protein language model trained on 250 million sequences from UniRef, capable of zero-shot variant effect prediction through log-likelihood ratios. DeepMind’s AlphaMissense [Cheng et al., 2023] combines AlphaFold2 structural features with sequence context to classify all possible human missense variants, achieving state-of-the-art performance on ClinVar benchmarks.

Critically, ESM-2 and AlphaMissense derive their predictions from orthogonal information sources: ESM-2 operates purely from evolutionary sequence conservation patterns, whereas AlphaMissense integrates structural features from AlphaFold2 with population frequency priors. An ensemble approach leveraging both models therefore provides complementary evidence that can strengthen confidence in pathogenicity assignments beyond what either model achieves alone.

In this study, we systematically apply ESM-2 and AlphaMissense to the full landscape of 7,467 possible TP53 single-amino-acid substitutions and cross-reference the results against 1,211 ClinVar VUS. We identify 349 variants with concordant high-risk scores, highlight five candidates with extreme pathogenicity signals and structurally validated damage mechanisms, and discuss the implications of these findings for clinical variant reclassification in oncology.

2 Materials and Methods

2.1 Variant Ascertainment

TP53 missense variants classified as “Uncertain Significance” were retrieved from NCBI ClinVar [Lan- drum et al., 2018] via the Entrez E-utilities API. Variants were filtered to retain only single-nucleotide missense substitutions mapped to the canonical TP53 protein isoform (RefSeq NP_000537.3; UniProt

P04637; 393 amino acids). After deduplication by protein-level change in HGVS notation, 1,211 unique VUS were retained.

2.2 ESM-2 Variant Effect Prediction

Variant effect scores were computed using Meta’s ESM-2 protein language model (esm2_t33_650M_UR50D; 650 million parameters) [Lin et al., 2023], accessed via the HuggingFace Transformers library. For each variant, the wild-type TP53 sequence was passed through the model, and the log-likelihood ratio (LLR) was calculated as:

$$\text{LLR} = \log P(x_{\text{mut}} | \mathbf{x}_{\setminus i}) - \log P(x_{\text{wt}} | \mathbf{x}_{\setminus i}) \quad (1)$$

where $P(x | \mathbf{x}_{\setminus i})$ denotes the model’s predicted probability for amino acid x at position i , conditioned on the full sequence context. Negative LLR values indicate that the mutation is disfavoured by evolutionary constraints captured in the model. Scores were stratified into four tiers: strongly damaging ($\text{LLR} \leq -4.0$), likely damaging ($-4.0 < \text{LLR} \leq -2.0$), possibly damaging ($-2.0 < \text{LLR} \leq -0.5$), and likely neutral ($\text{LLR} > -0.5$).

All 1,211 VUS were scored in batch mode. Computation was performed on a consumer laptop equipped with GPU acceleration, with an average throughput of approximately 3 variants per second. Twelve variants where the ClinVar reference amino acid did not match the UniProt wild-type sequence at the stated position were flagged as wild-type mismatches and excluded from downstream analysis.

2.3 AlphaMissense Pathogenicity Scores

Precomputed AlphaMissense pathogenicity scores [Cheng et al., 2023] for all possible single-amino-acid substitutions in the human proteome were downloaded from the Zenodo repository (record 10813168; file size \sim 1.12 GB compressed). The dataset was stream-filtered for TP53 (UniProt P04637), yielding 7,467 scored substitutions spanning all 393 residue positions. AlphaMissense scores range from 0 to 1, with the recommended classification thresholds of pathogenic (> 0.564), ambiguous (0.340–0.564), and benign (< 0.340) as defined by Cheng et al. [2023].

Of the 7,467 TP53 substitutions scored by AlphaMissense, 3,417 (45.8%) were classified as pathogenic, 763 (10.2%) as ambiguous, and 3,287 (44.0%) as benign.

168 **2.4 Cross-Referencing and Concordance**
169 **Analysis**

170 ESM-2 and AlphaMissense scores were merged on
171 the protein-level variant identifier (e.g., “L257R”).
172 Of the 1,211 ESM-2-scored VUS, 1,199 (99.0%)
173 were successfully matched to an AlphaMissense pre-
174 diction. Concordance between the two models was
175 assessed using Pearson and Spearman correlation co-
176 efficients. Variants were classified into four quad-
177 rants based on dual thresholds: high-risk concor-
178 dant (ESM-2 LLR ≤ -4.0 and AlphaMissense >
179 0.564), low-risk concordant (ESM-2 LLR > -0.5
180 and AlphaMissense < 0.340), and two discordant
181 categories.

182 **2.5 Structural Validation**

183 Structural context for the top-ranked variants was as-
184 sessed using the crystal structure of the p53 core
185 domain–DNA complex (PDB: 1TUP; 2.2 Å resolu-
186 tion) [Cho et al., 1994]. This structure captures
187 the p53 DNA-binding domain (residues ~94–312)
188 of chain B bound sequence-specifically to a 21-bp
189 DNA duplex (chains E and F), with a structural Zn²⁺
190 ion coordinated by Cys176, His179, Cys238, and
191 Cys242.

192 Two complementary structural analysis ap-
193 proaches were employed:

194 *Computational geometry analysis.* Spatial context
195 was computed with BioPython’s NeighborSearch
196 module [Cock et al., 2009], using a contact dis-
197 tance threshold of 4.0 Å. For each variant site, we
198 identified neighbouring protein residues within the
199 contact shell, DNA atoms within contact distance,
200 and zinc ion proximity. Results were visualised as
201 three-dimensional scatter plots showing the 12 Å ra-
202 dius structural environment around each mutation
203 site (Supplementary Fig. S1).

204 *Publication-quality structural rendering.* High-
205 resolution renders were generated using PyMOL
206 (open-source version 3.1.0) [The PyMOL Molecular
207 Graphics System, 2015] in headless mode. For each
208 variant, the wild-type residue was displayed along-
209 side a computationally modelled mutant rotamer (ap-
210 plied via PyMOL’s mutagenesis wizard) to visualise
211 steric and chemical differences. All renders were
212 ray-traced at 2,400 × 2,400 pixels with antialias-
213 ing level 4. An initial set of single-residue PyMOL
214 renders was also generated from a scripted .pml
215 pipeline (Supplementary Fig. S2).

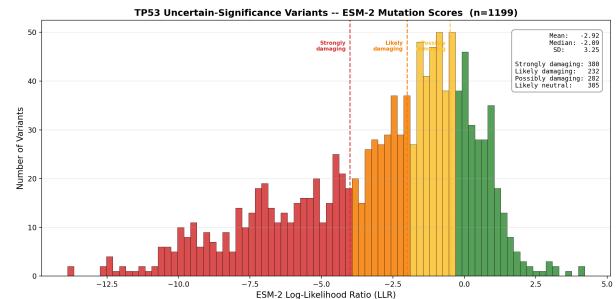


Figure 1: Distribution of ESM-2 log-likelihood ratio (LLR) scores across 1,199 TP53 VUS. Dashed vertical lines indicate classification thresholds. Variants with LLR ≤ -4.0 are designated strongly damaging.

216 **2.6 Computational Environment**

217 Primary analysis was performed on a consumer lap-
218 top running Windows. ESM-2 inference utilised lo-
219 cal GPU acceleration via PyTorch with CUDA. Al-
220 phaMissense data were obtained from Google Cloud
221 Storage / Zenodo. All analysis scripts were im-
222 plemented in Python 3, using BioPython [Cock
223 et al., 2009], PyTorch, HuggingFace Transformers,
224 NumPy, and Matplotlib. Structural rendering em-
225 ployed PyMOL 3.1.0 via the micromamba package
226 manager.

3 Results

3.1 ESM-2 Score Distribution across TP53 VUS

ESM-2 log-likelihood ratios for the 1,199 matched VUS ranged from -13.88 to +4.21 (Fig. 1). The distribution was left-skewed, consistent with the DNA-binding domain harbouring a disproportionate fraction of damaging variants. Of the 1,199 scored variants, 380 (31.7%) were classified as strongly damaging ($\text{LLR} \leq -4.0$), 232 (19.3%) as likely damaging, 282 (23.5%) as possibly damaging, and 305 (25.4%) as likely neutral.

3.2 Model Concordance

ESM-2 LLR and AlphaMissense pathogenicity scores showed a strong negative correlation (Pearson $r = -0.706$; Spearman $\rho = -0.714$; Fig. 2), indicating that variants scored as highly damaging by ESM-2 (more negative LLR) were independently scored as highly pathogenic by AlphaMissense (score approaching 1.0). The anti-correlation is expected given the inverse directionality of the two scoring scales.

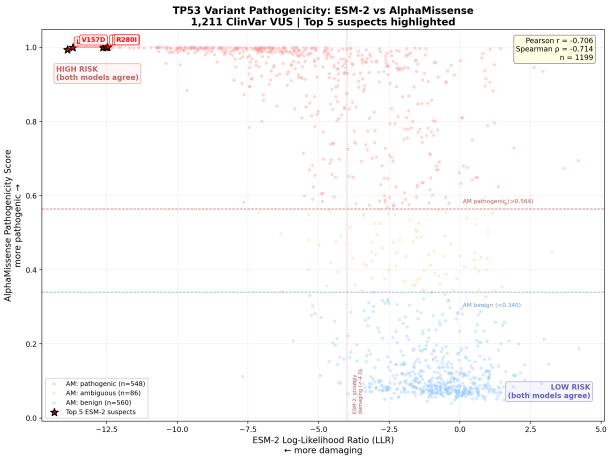


Figure 2: Scatter plot of ESM-2 LLR versus AlphaMissense pathogenicity score for 1,199 matched TP53 VUS. The strong anti-correlation (Pearson $r = -0.706$; Spearman $\rho = -0.714$) supports complementary predictive capacity between the sequence-based and structure-based models.

3.4 Top Five Candidate Pathogenic Variants

Table 1 presents the five VUS with the most extreme concordant pathogenicity scores. All five reside in the DNA-binding domain of p53 and exhibit structurally validated damage mechanisms.

3.4.1 L257R (p.Leu257Arg): Hydrophobic Core Disruption

Leucine 257 is buried within the hydrophobic β -sandwich core of the p53 DNA-binding domain. The substitution to arginine introduces a positively charged, bulky side chain into a tightly packed non-polar environment. This variant received the most extreme ESM-2 LLR of -13.88 and an AlphaMissense score of 0.9938, consistent with severe destabilisation of the protein fold (Fig. 4a).

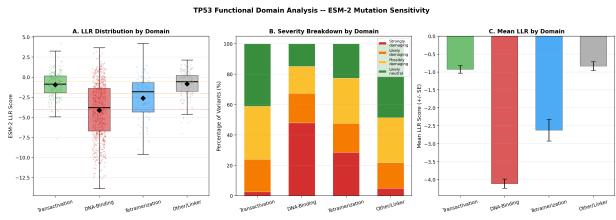


Figure 3: ESM-2 LLR score distribution by TP53 functional domain. The DNA-binding domain (residues 102–292) harbours the most severely scored variants, consistent with its critical role in tumour suppression.

3.4.2 V157D (p.Val157Asp): Hydrophobic Core Disruption

Valine 157 occupies a β -strand within the interior of the same β -sandwich. Substitution to aspartate introduces a negative charge and a shorter side chain, creating both electrostatic repulsion and a packing cavity. The ESM-2 LLR of -13.69 and AlphaMissense score of 0.9992 both indicate near-certain pathogenicity (Fig. 4b).

3.4.3 R248P (p.Arg248Pro): Loss of DNA Minor Groove Contact

Arginine 248 is one of the most frequently mutated residues in human cancer. In the wild-type structure, R248 inserts directly into the DNA minor groove at a distance of approximately 3.4 Å, forming critical hydrogen bonds with the DNA backbone. Proline at this position eliminates all hydrogen-bonding capacity and introduces a rigid kink in the polypeptide backbone. Both models scored this variant at extreme levels (LLR = -12.63 ; AM = 0.9994) (Fig. 4c).

3.4.4 C176R (p.Cys176Arg): Zinc Coordination Abolished

Cysteine 176 is one of four residues (C176, H179, C238, C242) that coordinate the structural Zn²⁺ ion essential for the folding of the L2 and L3 loops of the DNA-binding domain. The thiolate side chain of

Applying dual thresholds, 349 variants (29.1%) were classified as high-risk by both models, and 438 (36.5%) were classified as low-risk by both models. The remaining 412 variants (34.4%) showed discordant classifications, occupying the ambiguous region where additional evidence is needed.

3.3 Domain-Level Enrichment

Variant pathogenicity scores were non-uniformly distributed across TP53 functional domains (Fig. 3). The DNA-binding domain (residues 102–292) showed significantly more negative mean ESM-2 LLR scores than the transactivation domain (residues 1–92) or the tetramerization domain (residues 325–356), consistent with stronger evolutionary constraint on the DNA-binding interface. All five top-ranked variants mapped to the DNA-binding domain.

Table 1: Top five TP53 VUS with extreme concordant pathogenicity scores. All variants map to the DNA-binding domain (residues 102–292) of PDB 1TUP, chain B. ESM-2 LLR values below -4.0 indicate strong damage; AlphaMissense scores above 0.564 indicate pathogenicity.

Variant	HGVS	Residue	ESM-2 LLR	AM Score	ClinVar ID	Structural Mechanism
L257R	p.Leu257Arg	257	-13.88	0.9938	142134	Hydrophobic core disruption
V157D	p.Val157Asp	157	-13.69	0.9992	482231	Hydrophobic core disruption
R248P	p.Arg248Pro	248	-12.63	0.9994	237954	DNA minor groove contact lost
C176R	p.Cys176Arg	176	-12.53	0.9999	376573	Zinc coordination abolished
R280I	p.Arg280Ile	280	-12.47	0.9996	161517	DNA major groove contact lost

309 cysteine provides a ligand to zinc at a distance of ap- 345
 310 proximately 2.3 Å. Arginine cannot coordinate zinc, 346
 311 and its introduction is predicted to collapse the local 347
 312 loop scaffold. AlphaMissense assigned the highest 348
 313 score in our dataset (0.9999) to this variant (Fig. 4d). 349

substantial subset—particularly those with extreme scores such as the five candidates highlighted here—carry sufficient *in silico* evidence to support provisional reclassification, pending functional confirmation.

314 3.4.5 R280I (p.Arg280Ile): Loss of DNA Major 350 315 Groove Contact 351

316 Arginine 280 forms a direct hydrogen bond with a 352
 317 guanine base in the DNA major groove at a distance 353
 318 of approximately 2.8 Å. This contact is essential for 354
 319 sequence-specific DNA recognition. Substitution to 355
 320 isoleucine, a hydrophobic residue with no hydrogen- 356
 321 bonding capacity, abolishes this interaction entirely 357
 322 (LLR = -12.47 ; AM = 0.9996) (Fig. 4e). 358

4.2 Structural Basis of Predicted Pathogenicity

The five top-ranked variants illustrate three distinct molecular mechanisms of p53 loss of function, each corroborated by structural analysis of PDB 1TUP and the three-dimensional spatial context computed via BioPython (Supplementary Fig. S1):

1. **Direct DNA contact loss** (R248P, R280I): These residues form hydrogen bonds with DNA bases or backbone atoms in the p53 response element. R248 is a known mutational hotspot in cancer (R248W, R248Q are among the six most common TP53 mutations) [Bouaoun et al., 2016], but the R248P substitution—which introduces a conformationally rigid proline—has not been previously classified as pathogenic in ClinVar despite affecting the same critical contact residue.

2. **Zinc coordination abolishment** (C176R): The structural Zn²⁺ ion is essential for the folding and stability of the L2–L3 loop region that forms part of the DNA-binding surface [Cho et al., 1994, Bullock et al., 2000]. Loss of even one zinc ligand is expected to destabilise the entire loop scaffold.

3. **Hydrophobic core disruption** (L257R, V157D): Introduction of charged residues into the buried β -sandwich core is a well-established mechanism of p53 thermodynamic destabilisation [Bullock et al., 2000]. Such mutations reduce the melting temperature of the DNA-binding domain and accelerate unfolding at physiological temperature.

The structural consistency between the AI predictions and the known three-dimensional architecture

323 4 Discussion

324 4.1 Concordant AI Prediction as Evidence 363 325 for Reclassification 364

326 The strong anti-correlation between ESM-2 and Al- 366
 327 phaMissense predictions ($r = -0.706$; $\rho = -0.714$) 367
 328 is noteworthy because the two models were trained 368
 329 on fundamentally different data representations. 369
 330 ESM-2 is a pure sequence model that learns evo- 370
 331 lutionary constraints from 250 million protein se- 371
 332 quences without any explicit structural information 372
 333 [Lin et al., 2023]. AlphaMissense, by contrast, incor- 373
 334 porates AlphaFold2-derived structural features along- 374
 335 side sequence context and population frequency data 375
 336 [Cheng et al., 2023]. The convergence of these or- 376
 337 thogonal approaches on the same set of high-risk vari- 377
 338 ants provides a form of computational triangulation 378
 339 analogous to the convergent evidence required by the 379
 340 ACMG/AMP framework [Richards et al., 2015]. 380

341 The 349 variants flagged as high-risk by both mod- 381
 342 els represent 29.1% of the matched ClinVar VUS. 382
 343 Under current clinical guidelines, none of these vari- 383
 344 ants are actionable. Our analysis suggests that a 384

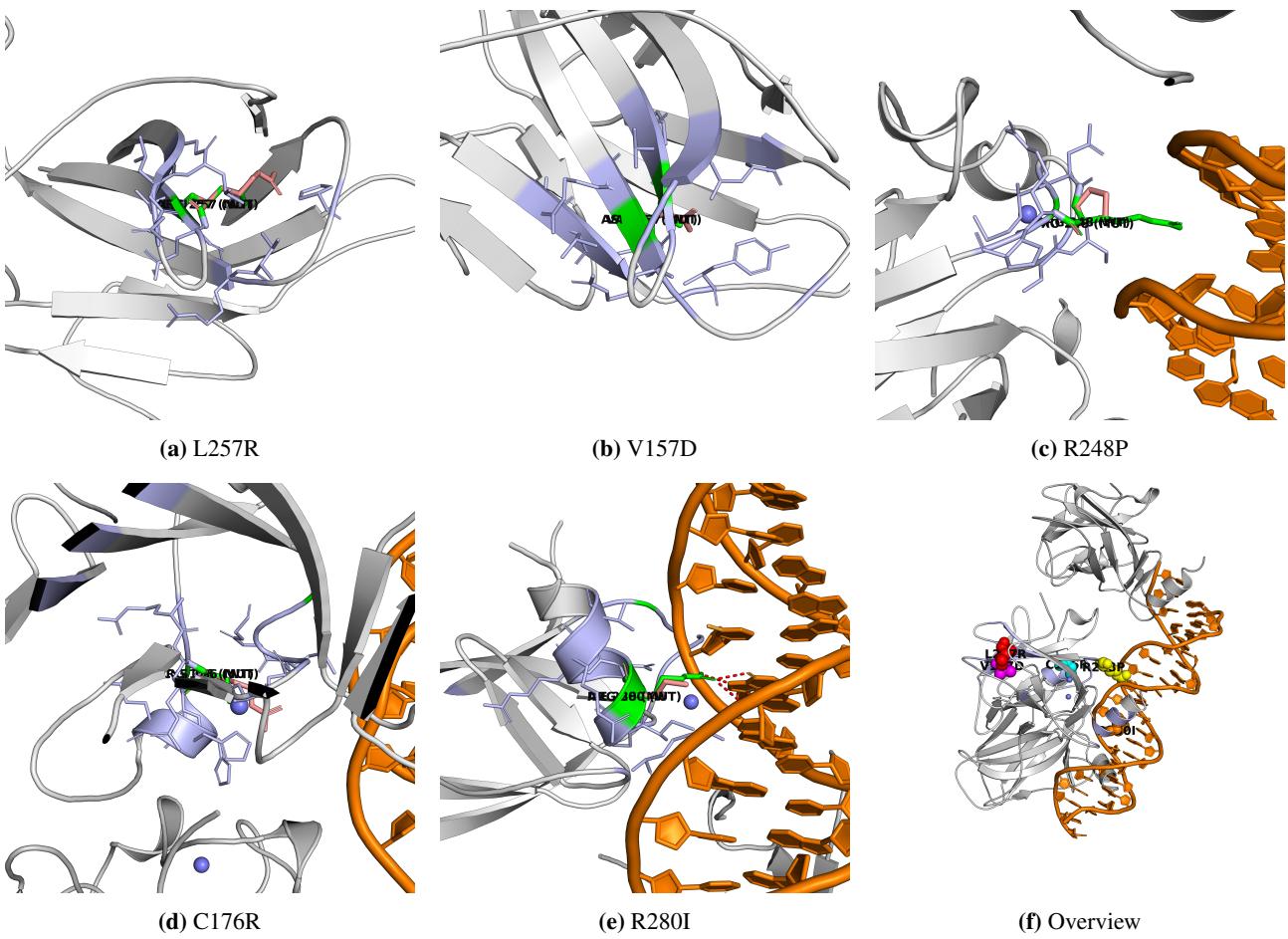


Figure 4: Publication-quality structural renders of the top five TP53 VUS (PDB 1TUP, chain B). Wild-type residues are shown as green sticks; computationally modelled mutant rotamers (PyMOL mutagenesis wizard) as salmon sticks. Neighbouring residues within 4.0 Å are shown as light-blue sticks. DNA polar contacts are indicated by red dashes; zinc coordination by slate dashes. (a) L257R: charged arginine disrupts the hydrophobic β-sandwich core. (b) V157D: aspartate introduces charge and a packing void. (c) R248P: proline eliminates DNA minor groove contact. (d) C176R: arginine abolishes Zn²⁺ coordination. (e) R280I: isoleucine abolishes DNA major groove hydrogen bond. (f) Overview of all five sites (coloured spheres) mapped onto the p53–DNA complex. Protein shown as gray cartoon; DNA as orange cartoon; Zn²⁺ ions as slate spheres. All panels ray-traced at 2,400×2,400 pixels, 300 DPI.

of the p53–DNA complex provides a mechanistic rationale for the extreme scores observed, and supports the biological plausibility of the reclassification. Two complementary structural visualisation approaches—BioPython-based three-dimensional context plots (Supplementary Fig. S1) and PyMOL ray-traced renderers at multiple levels of detail (Fig. 4; Supplementary Fig. S2)—confirm the spatial relationships described above.

4.3 Implications for Personalised Oncology

The clinical impact of resolving TP53 VUS extends across multiple domains of cancer care:

Germline testing in Li–Fraumeni syndrome. Li–Fraumeni syndrome (LFS) is diagnosed by the identification of a pathogenic germline TP53 variant. Individuals with LFS face a cumulative cancer risk

exceeding 90% by age 60, and benefit from intensive surveillance protocols including annual whole-body MRI [Villani et al., 2016]. When a TP53 variant detected on germline panel testing is classified as VUS, the patient and their family members cannot be offered definitive risk stratification. Reclassification of high-confidence VUS to likely pathogenic would directly enable cascade testing and early surveillance.

Somatic tumour profiling. TP53 mutational status is a key biomarker in haematological malignancies, where it predicts resistance to chemoimmunotherapy in chronic lymphocytic leukaemia [Zenz et al., 2010] and adverse prognosis in myelodysplastic syndromes [Bejar et al., 2011]. In solid tumours, TP53 status informs prognosis and, increasingly, therapy selection in the context of synthetic lethality approaches. Resolving VUS enables more precise molecular stratification.

419 **Emerging p53-targeted therapies.** A new gener- 466
420 ation of therapeutics aims to restore or stabilise mu-
421 tant p53 function. Small molecules such as APR- 467
422 246 (eprenetapopt) and PC14586 (rezatapopt) have 468
423 entered clinical trials for tumours harbouring specific 469
424 TP53 mutations [Chen et al., 2021]. Accurate classi- 470
425 fication of TP53 variants is a prerequisite for patient 471
426 selection in these trials. Variants that destabilise the 472
427 protein fold (L257R, V157D) may respond to fold- 473
428 stabilising compounds, whereas those that abolish 474
429 DNA contact (R248P, R280I) may require distinct 475
430 therapeutic strategies.

5 Conclusions

We demonstrate that ensemble AI scoring using ESM-2 and AlphaMissense can systematically identify high-confidence pathogenic variants among the 1,211 TP53 VUS currently in ClinVar. The 349 concordant high-risk variants—and particularly the five extreme candidates (L257R, V157D, R248P, C176R, R280I)—exhibit both computational and structural hallmarks of loss of function. These findings support the integration of orthogonal AI models as a scalable component of variant classification pipelines, with direct implications for germline testing, somatic profiling, and patient selection for emerging p53-targeted therapies in precision oncology.

431 4.4 Limitations

432 Several limitations should be noted. First, *in silico*
433 predictions, regardless of model concordance, do not
434 constitute functional evidence under ACMG/AMP
435 criteria (PP3) and cannot alone support a pathogenic
436 classification beyond “supporting” evidence strength.
437 Second, our structural analysis is based on a single
438 crystal structure (PDB 1TUP) that captures only one
439 conformational state of the p53 tetramer–DNA com-
440 plex; dynamic effects and post-translational modi-
441 fications are not represented. Third, ESM-2 LLR
442 thresholds (≤ -4.0) and AlphaMissense cutoffs ($>$
443 0.564) were adopted from the original publications
444 and have not been independently calibrated on a
445 TP53-specific truth set. Finally, the clinical sig-
446 nificance of the 349 high-risk VUS should be re-
447 garded as provisional until orthogonal experimen-
448 tal validation—such as yeast-based functional assays,
449 thermal stability measurements, or DNA-binding
450 electrophoretic mobility shift assays—is completed.

451 4.5 Future Directions

452 This work motivates several follow-up investigations.
453 The 349 concordant high-risk variants represent a
454 prioritised set for experimental functional validation,
455 potentially through high-throughput approaches such
456 as multiplexed assays of variant effect (MAVEs)
457 [Findlay et al., 2018]. Integration of additional *in*
458 *silico* tools—including EVE [Frazer et al., 2021],
459 REVEL [Ioannidis et al., 2016], and molecular dy-
460 namics simulations—could further refine the confi-
461 dence tiers. Longitudinal tracking of these variants in
462 ClinVar will reveal whether independent clinical ev-
463 idence eventually converges with the AI predictions
464 presented here, providing a natural validation of the
465 ensemble approach.

480 **Data Availability**

481 All analysis scripts, scored variant data, and struc-
482 tural renders are available in the project repository.
483 ClinVar data were accessed via NCBI Entrez. Al-
484 phaMissense proteome-wide predictions are avail-
485 able from Zenodo (record 10813168). The PDB
486 structure 1TUP is available from the RCSB Protein
487 Data Bank.

488 **Acknowledgements**

489 The author thanks the developers of ESM-2 (Meta
490 AI), AlphaMissense (Google DeepMind), ClinVar
491 (NCBI), BioPython, and PyMOL for making their
492 tools and data freely accessible. The 1TUP crystal
493 structure was obtained from the RCSB Protein Data
494 Bank. This project was developed as an independent
495 computational biology research effort.

496 **Author Contact Information**

- 497 • **Author:** Mahad Asif
498 • **Email:** mahaddevx@gmail.com
499 • **GitHub:** github.com/mahaddevx
500 • **Project Repository:**
501 github.com/mahaddevx/TP53-AI-Screening

502 All source code, scored variant tables, structural anal-
503 ysis outputs, and PyMOL render scripts used in this
504 study are publicly available in the project repository
505 above.

506 **Supplementary Figures**

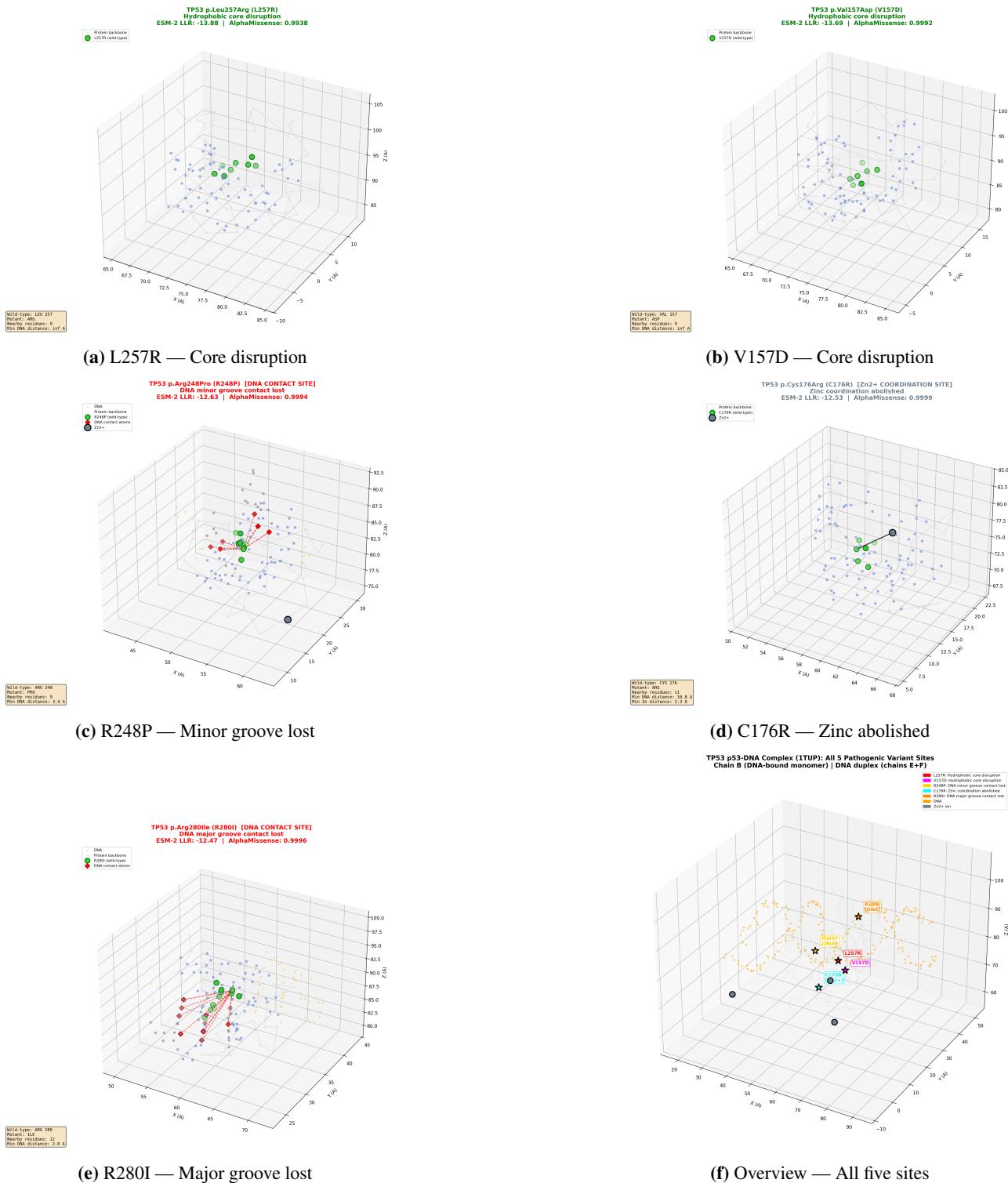


Figure S1: Supplementary Figure S1: Three-dimensional structural context of the top five TP53 VUS (BioPython analysis). Each panel shows a 12 Å radius view centred on the variant residue (green) in PDB 1TUP chain B. Orange: DNA; light gray: C α backbone; cornflower blue: neighbours within 4.0 Å; red diamonds: DNA contacts; slate: Zn $^{2+}$. (a) L257R buried, no DNA contact. (b) V157D buried in β -sandwich. (c) R248P DNA minor groove contact. (d) C176R Zn $^{2+}$ coordination. (e) R280I DNA major groove contact. (f) Overview of all five positions on chain B.

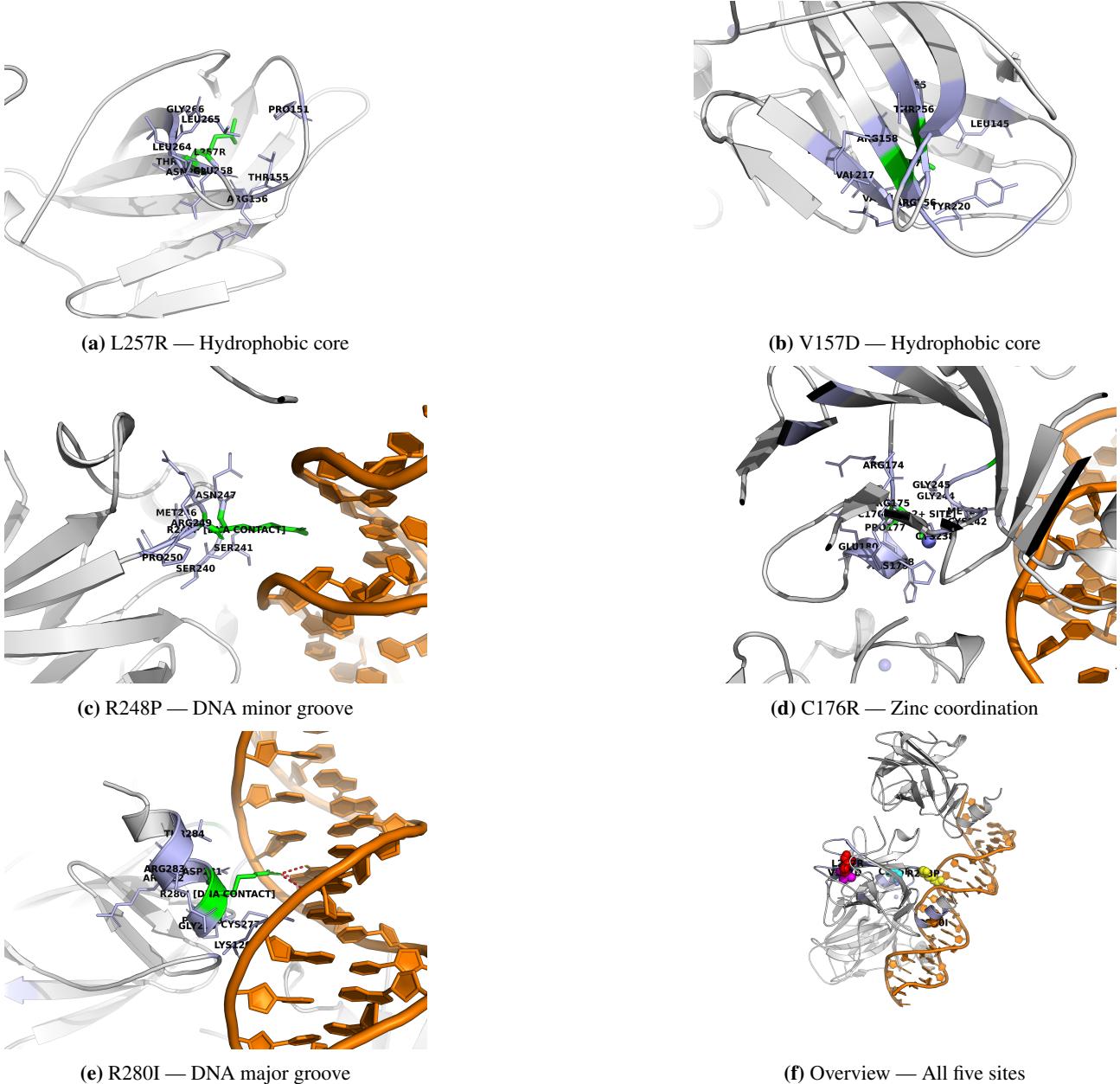


Figure S2: Supplementary Figure S2: Initial PyMOL ray-traced renders of the top five TP53 VUS (single-residue view). Generated via a scripted .pml pipeline from PDB 1TUP. Each panel shows the wild-type residue (green sticks) in context: protein cartoon (gray), DNA (orange), Zn²⁺ (slate spheres). Neighbours within 4.0 Å shown as light-blue sticks; DNA contacts as red dashes; zinc coordination as slate dashes. These renders show only the wild-type residue, in contrast to Fig. 4 which overlays both wild-type and mutant rotamers. All panels ray-traced at 2,400×1,800 pixels, 300 DPI. (a) L257R buried in the hydrophobic core. (b) V157D in the β-sandwich interior. (c) R248P with DNA minor groove contacts visible. (d) C176R with zinc coordination bonds. (e) R280I with DNA major groove contacts. (f) Overview of all five variant positions as coloured spheres on the full p53–DNA complex.

507 References

- 508 Bejar, R., Stevenson, K., Abdel-Wahab, O., et al. Clinical effect of point mutations in myelodysplastic syn-
509 dromes. *N. Engl. J. Med.*, 364(26):2496–2506, 2011.
- 510 Bouaoun, L., Sonber, D., Ardin, M., et al. TP53 variations in human cancers: new lessons from the IARC TP53
511 Database and genomics data. *Hum. Mutat.*, 37(9):865–876, 2016.
- 512 Bullock, A. N., Henckel, J. & Fersht, A. R. Quantitative analysis of residual folding and DNA binding in

- 513 mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene*, 19(10):1245–
514 1256, 2000.
- 515 Chen, S., Wu, J.-L., Liang, Y., et al. Small molecule therapeutics for TP53-mutant cancers. *Trends Pharma-*
516 *col. Sci.*, 42(12):1049–1062, 2021.
- 517 Cheng, J., Novati, G., Pan, J., et al. Accurate proteome-wide missense variant effect prediction with AlphaMis-
518 sense. *Science*, 381(6664):eadg7492, 2023.
- 519 Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. Crystal structure of a p53 tumor suppressor–DNA complex:
520 understanding tumorigenic mutations. *Science*, 265(5170):346–355, 1994.
- 521 Cock, P. J. A., Antao, T., Chang, J. T., et al. Biopython: freely available Python tools for computational
522 molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 2009.
- 523 Findlay, G. M., Daza, R. M., Martin, B., et al. Accurate classification of BRCA1 variants with saturation
524 genome editing. *Nature*, 562(7726):217–222, 2018.
- 525 Frazer, J., Notin, P., Dias, M., et al. Disease variant prediction with deep generative models of evolutionary
526 data. *Nature*, 599(7883):91–95, 2021.
- 527 Giacomelli, A. O., Yang, X., Lintber, R. E., et al. Mutational processes shape the landscape of TP53 mutations
528 in human cancer. *Nat. Genet.*, 50(10):1381–1387, 2018.
- 529 Ioannidis, N. M., Rothstein, J. H., Pejaver, V., et al. REVEL: an ensemble method for predicting the pathogenic-
530 ity of rare missense variants. *Am. J. Hum. Genet.*, 99(4):877–885, 2016.
- 531 Kandoth, C., McLellan, M. D., Vandin, F., et al. Mutational landscape and significance across 12 major cancer
532 types. *Nature*, 502(7471):333–339, 2013.
- 533 Landrum, M. J., Lee, J. M., Benson, M., et al. ClinVar: improving access to variant interpretations and
534 supporting evidence. *Nucleic Acids Res.*, 46(D1):D1062–D1067, 2018.
- 535 Lane, D. P. p53, guardian of the genome. *Nature*, 358(6381):15–16, 1992.
- 536 Levine, A. J. & Oren, M. The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer*, 9(10):749–
537 758, 2009.
- 538 Lin, Z., Akin, H., Rao, R., et al. Evolutionary-scale prediction of atomic-level protein structure with a language
539 model. *Science*, 379(6637):1123–1130, 2023.
- 540 Malkin, D., Li, F. P., Strong, L. C., et al. Germ line p53 mutations in a familial syndrome of breast cancer,
541 sarcomas, and other neoplasms. *Science*, 250(4985):1233–1238, 1990.
- 542 The PyMOL Molecular Graphics System, Version 3.1, Schrödinger, LLC.
- 543 Richards, S., Aziz, N., Bale, S., et al. Standards and guidelines for the interpretation of sequence variants:
544 a joint consensus recommendation of the American College of Medical Genetics and Genomics and the
545 Association for Molecular Pathology. *Genet. Med.*, 17(5):405–424, 2015.
- 546 Villani, A., Shore, A., Wasserman, J. D., et al. Biochemical and imaging surveillance in germline TP53 mu-
547 tation carriers with Li–Fraumeni syndrome: 11 year follow-up of a prospective observational study. *Lancet*
548 *Oncol.*, 17(9):1295–1305, 2016.
- 549 Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature*, 408(6810):307–310, 2000.
- 550 Zenz, T., Eichhorst, B., Busch, R., et al. TP53 mutation and survival in chronic lymphocytic leukemia.
551 *J. Clin. Oncol.*, 28(29):4473–4479, 2010.