

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 Uncer-
5 tain Significance (VUS), limiting their clinical utility
6 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 capac-
20 ity. A total of 349 VUS were flagged as high-risk by
21 both models (ESM-2 log-likelihood ratio ≤ -4.0 and
22 AlphaMissense pathogenicity > 0.564), while 438
23 were concordantly classified as low-risk. We highlight
24 five variants in the DNA-binding domain (L257R,
25 V157D, R248P, C176R, R280I) with extreme concor-
26 dant scores (ESM-2 LLR ≤ -12.5 ; AlphaMissense
27 ≥ 0.99) and structurally validated damage mecha-
28 nisms including loss of DNA contact, zinc coordi-
29 nation abolishment, and hydrophobic core disruption.

30 **Conclusions.** Ensemble AI scoring can systemat-
31 ically reclassify TP53 VUS at scale. The 349 high-
32 confidence pathogenic variants identified here war-
33 rant prioritized functional validation and may inform
34 germline testing guidelines in Li–Fraumeni syndrome
35 and somatic profiling in clinical oncology.

36 **Keywords:** TP53, variants of uncertain significance,
37 ESM-2, AlphaMissense, pathogenicity prediction,

38 precision oncology, protein language model

1 Introduction

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

55 The clinical interpretation of TP53 missense vari-
56 ants has been greatly advanced by large-scale reposi-
57 tories such as ClinVar [Landrum et al., 2018] and the
58 IARC TP53 Database [Bouaoun et al., 2016]. How-
59 ever, a substantial fraction of observed variants—over
60 1,200 unique missense substitutions in ClinVar at the
61 time of this analysis—remain classified as Variants
62 of Uncertain Significance (VUS). The VUS designa-
63 tion creates a “diagnostic grey zone” for clinicians:
64 these variants cannot be used to guide treatment deci-
65 sions, genetic counselling, or cascade family testing,
66 even when the variant resides in a functionally critical
67 domain [Richards et al., 2015].

68 The bottleneck in VUS resolution is fundamen-
69 tally one of evidence accumulation. Under the
70 ACMG/AMP framework [Richards et al., 2015], re-
71 classification requires convergent evidence from pop-
72 ulation frequency data, *in silico* prediction, functional
73 assays, co-segregation studies, and *de novo* occur-

74 rence. For rare variants observed only once or twice 123
75 in clinical databases, such evidence may never accrue
76 through observation alone. Large-scale functional 124
77 assays, such as the saturation mutagenesis screen of 125
78 TP53 by Giacomelli et al. [2018], have made impor- 126
79 tant contributions but remain labour-intensive and 127
80 costly. 128

81 Recent advances in protein language models 129
82 (pLMs) and structure-based pathogenicity predic- 130
83 tors offer a scalable complement to experimental 131
84 approaches. Meta’s ESM-2 [Lin et al., 2023] is a 132
85 transformer-based protein language model trained on 133
86 250 million sequences from UniRef, capable of zero- 134
87 shot variant effect prediction through log-likelihood 135
88 ratios. DeepMind’s AlphaMissense [Cheng et al., 136
89 2023] combines AlphaFold2 structural features with 137
90 sequence context to classify all possible human mis- 138
91 sense variants, achieving state-of-the-art performance 139
92 on ClinVar benchmarks. 140

93 Critically, ESM-2 and AlphaMissense derive their 138
94 predictions from orthogonal information sources: 139
95 ESM-2 operates purely from evolutionary sequence 140
96 conservation patterns, whereas AlphaMissense inte- 141
97 grates structural features from AlphaFold2 with popu- 142
98 lation frequency priors. An ensemble approach lever- 143
99 aging both models therefore provides complementary 144
100 evidence that can strengthen confidence in pathogenic- 145
101 ity assignments beyond what either model achieves 146
102 alone. 147

103 In this study, we systematically apply ESM-2 and 148
104 AlphaMissense to the full landscape of 7,467 possi-
105 ble TP53 single-amino-acid substitutions and cross-
106 reference the results against 1,211 ClinVar VUS. We 149
107 identify 349 variants with concordant high-risk scores, 150
108 highlight five candidates with extreme pathogenic- 151
109 ity signals and structurally validated damage mecha- 152
110 nisms, and discuss the implications of these findings 153
111 for clinical variant reclassification in oncology. 154
155

112 2 Materials and Methods

113 2.1 Variant Ascertainment

114 TP53 missense variants classified as “Uncertain Sig- 161
115 nificance” were retrieved from NCBI ClinVar [Lan- 162
116 drum et al., 2018] via the Entrez E-utilities API. Vari- 163
117 ants were filtered to retain only single-nucleotide mis-
118 sense substitutions mapped to the canonical TP53 pro- 164
119 tein isoform (RefSeq NP_000537.3; UniProt P04637; 165
120 393 amino acids). After deduplication by protein-
121 level change in HGVS notation, 1,211 unique VUS 166
122 were retained. 167

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 ~ 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.

2.4 Cross-Referencing and Concordance Analysis

ESM-2 and AlphaMissense scores were merged on the protein-level variant identifier (e.g., “L257R”). Of

168 the 1,211 ESM-2-scored VUS, 1,199 (99.0%) were
 169 successfully matched to an AlphaMissense prediction.
 170 Concordance between the two models was assessed
 171 using Pearson and Spearman correlation coefficients.
 172 Variants were classified into four quadrants based on
 173 dual thresholds: high-risk concordant (ESM-2 LLR
 174 ≤ -4.0 and AlphaMissense > 0.564), low-risk con-
 175 cordant (ESM-2 LLR > -0.5 and AlphaMissense
 176 < 0.340), and two discordant categories.

177 2.5 Structural Validation

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

186 Two complementary structural analysis approaches
 187 were employed:

188 *Computational geometry analysis.* Spatial context
 189 was computed with BioPython’s NeighborSearch
 190 module [Cock et al., 2009], using a contact distance
 191 threshold of 4.0 Å. For each variant site, we identified
 192 neighbouring protein residues within the contact shell,
 193 DNA atoms within contact distance, and zinc ion prox-
 194 imity. Results were visualised as three-dimensional
 195 scatter plots showing the 12 Å radius structural envi-
 196 ronment around each mutation site (Supplementary
 197 Fig. S1).

198 *Publication-quality structural rendering.* High-
 199 resolution renders were generated using PyMOL
 200 (open-source version 3.1.0) [The PyMOL Molecular
 201 Graphics System, 2015] in headless mode. For each
 202 variant, the wild-type residue was displayed alongside
 203 a computationally modelled mutant rotamer (applied
 204 via PyMOL’s mutagenesis wizard) to visualise steric
 205 and chemical differences. All renders were ray-traced
 206 at 2,400 × 2,400 pixels with antialiasing level 4. An
 207 initial set of single-residue PyMOL renders was also
 208 generated from a scripted .pml pipeline (Supplemen-
 209 tary Fig. S2).

210 2.6 Computational Environment

211 Primary analysis was performed on a consumer lap-
 212 top running Windows. ESM-2 inference utilised lo-
 213 cal GPU acceleration via PyTorch with CUDA. Al-
 214 phaMissense data were obtained from Google Cloud
 215 Storage / Zenodo. All analysis scripts were imple-
 216

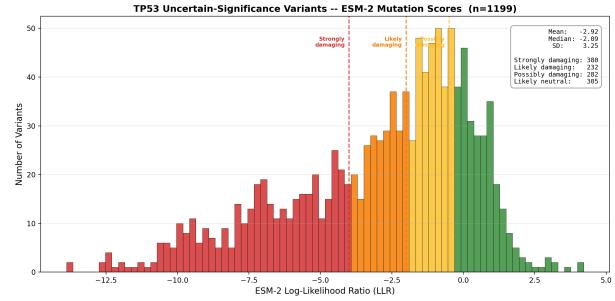


Figure 1: Distribution of ESM-2 log-likelihood ratio (LLR) scores across 1,199 TP53 VUS. Dashed vertical lines indicate classification thresholds. Variants with LLR ≤ -4.0 are designated strongly damaging.

mented in Python 3, using BioPython [Cock et al., 2009], PyTorch, HuggingFace Transformers, NumPy, and Matplotlib. Structural rendering employed PyMOL 3.1.0 via the micromamba package 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.

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.

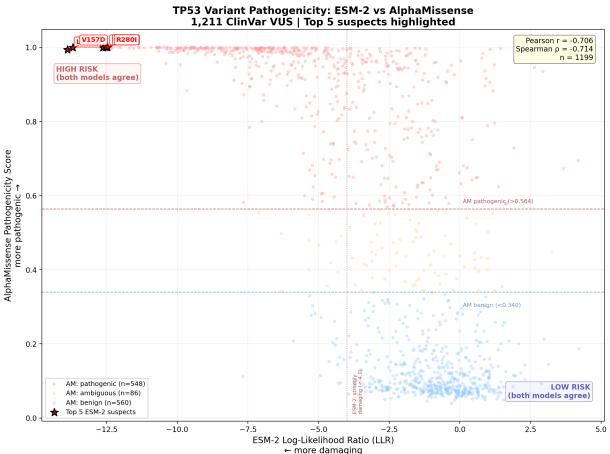


Figure 2: Scatter plot of ESM-2 LLR versus AlphaMissense pathogenicity score for 1,199 matched TP53 VUS. The strong anti-correlation (Pearson $r = -0.706$; Spearman $\rho = -0.714$) supports complementary predictive capacity between the sequence-based and structure-based models.

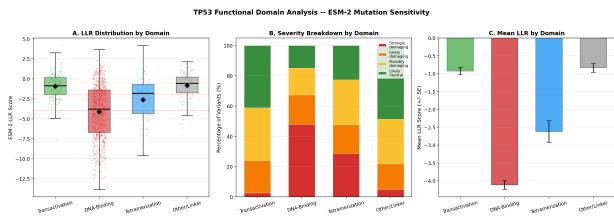


Figure 3: ESM-2 LLR score distribution by TP53 functional domain. The DNA-binding domain (residues 102–292) harbours the most severely scored variants, consistent with its critical role in tumour suppression.

3.3 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.

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).

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 cysteine provides a ligand to zinc at a distance of approximately 2.3 Å. Arginine cannot coordinate zinc, and its introduction is predicted to collapse the local loop scaffold. AlphaMissense assigned the highest score in our dataset (0.9999) to this variant (Fig. 4d).

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

303 **3.4.5 R280I (p.Arg280Ile): Loss of DNA Major 338**
 304 **Groove Contact** 339

305 Arginine 280 forms a direct hydrogen bond with a 340
 306 guanine base in the DNA major groove at a distance 341
 307 of approximately 2.8 Å. This contact is essential for 342
 308 sequence-specific DNA recognition. Substitution to 343
 309 isoleucine, a hydrophobic residue with no hydrogen- 344
 310 bonding capacity, abolishes this interaction entirely 345
 311 (LLR = −12.47; AM = 0.9996) (Fig. 4e).

4.2 Structural Basis of Predicted Pathogenicity

340 The five top-ranked variants illustrate three distinct 341
 342 molecular mechanisms of p53 loss of function, each 343
 344 corroborated by structural analysis of PDB 1TUP and 345
 346 the three-dimensional spatial context computed via 347
 348 BioPython (Supplementary Fig. S1): 349

349 **1. Direct DNA contact loss (R248P, R280I):** These 350
 351 residues form hydrogen bonds with DNA bases 352
 353 or backbone atoms in the p53 response element. 354
 355 R248 is a known mutational hotspot in cancer 356
 357 (R248W, R248Q are among the six most common 358
 359 TP53 mutations) [Bouaoun et al., 2016], 360
 361 but the R248P substitution—which introduces 362
 363 a conformationally rigid proline—has not been 364
 365 previously classified as pathogenic in ClinVar despite 366
 367 affecting the same critical contact residue.

368 **2. Zinc coordination abolition (C176R):** The 369
 370 structural Zn²⁺ ion is essential for the folding 371
 372 and stability of the L2–L3 loop region that forms 373
 374 part of the DNA-binding surface [Cho et al., 375
 376 1994, Bullock et al., 2000]. Loss of even one 377
 378 zinc ligand is expected to destabilise the entire 379
 380 loop scaffold.

381 **3. Hydrophobic core disruption (L257R, 382
 383 V157D):** Introduction of charged residues 384
 385 into the buried β-sandwich core is a well- 386
 387 established mechanism of p53 thermodynamic 388
 389 destabilisation [Bullock et al., 2000]. Such 390
 391 mutations reduce the melting temperature of the 392
 393 DNA-binding domain and accelerate unfolding 394
 395 at physiological temperature.

396 The structural consistency between the AI predictions 397
 398 and the known three-dimensional architecture 399
 400 of the p53–DNA complex provides a mechanistic 401
 402 rationale for the extreme scores observed, and supports 403
 404 the biological plausibility of the reclassification. Two 405
 406 complementary structural visualisation approaches— 407
 408 BioPython-based three-dimensional context plots 409
 410 (Supplementary Fig. S1) and PyMOL ray-traced ren- 411
 412 ders at multiple levels of detail (Fig. 4; Supplementary

312 **4 Discussion**

313 **4.1 Concordant AI Prediction as Evidence 353**
 314 **for Reclassification** 354

315 The strong anti-correlation between ESM-2 and Al-356
 316 phaMissense predictions ($r = -0.706$; $\rho = -0.714$) 357
 317 is noteworthy because the two models were trained on 358
 318 fundamentally different data representations. ESM-2 359
 319 is a pure sequence model that learns evolutionary con- 360
 320 straints from 250 million protein sequences without 361
 321 any explicit structural information [Lin et al., 2023]. 362
 322 AlphaMissense, by contrast, incorporates AlphaFold2- 363
 323 derived structural features alongside sequence context 364
 324 and population frequency data [Cheng et al., 2023]. 365
 325 The convergence of these orthogonal approaches on 366
 326 the same set of high-risk variants provides a form of 367
 327 computational triangulation analogous to the conver- 368
 328 gent evidence required by the ACMG/AMP frame- 369
 329 work [Richards et al., 2015]. 370

330 The 349 variants flagged as high-risk by both mod-371
 331 els represent 29.1% of the matched ClinVar VUS. Un-372
 332 der current clinical guidelines, none of these variants 373
 333 are actionable. Our analysis suggests that a substan-374
 334 tial subset—particularly those with extreme scores 375
 335 such as the five candidates highlighted here—carry 376
 336 sufficient *in silico* evidence to support provisional 377
 337 reclassification, pending functional confirmation. 378

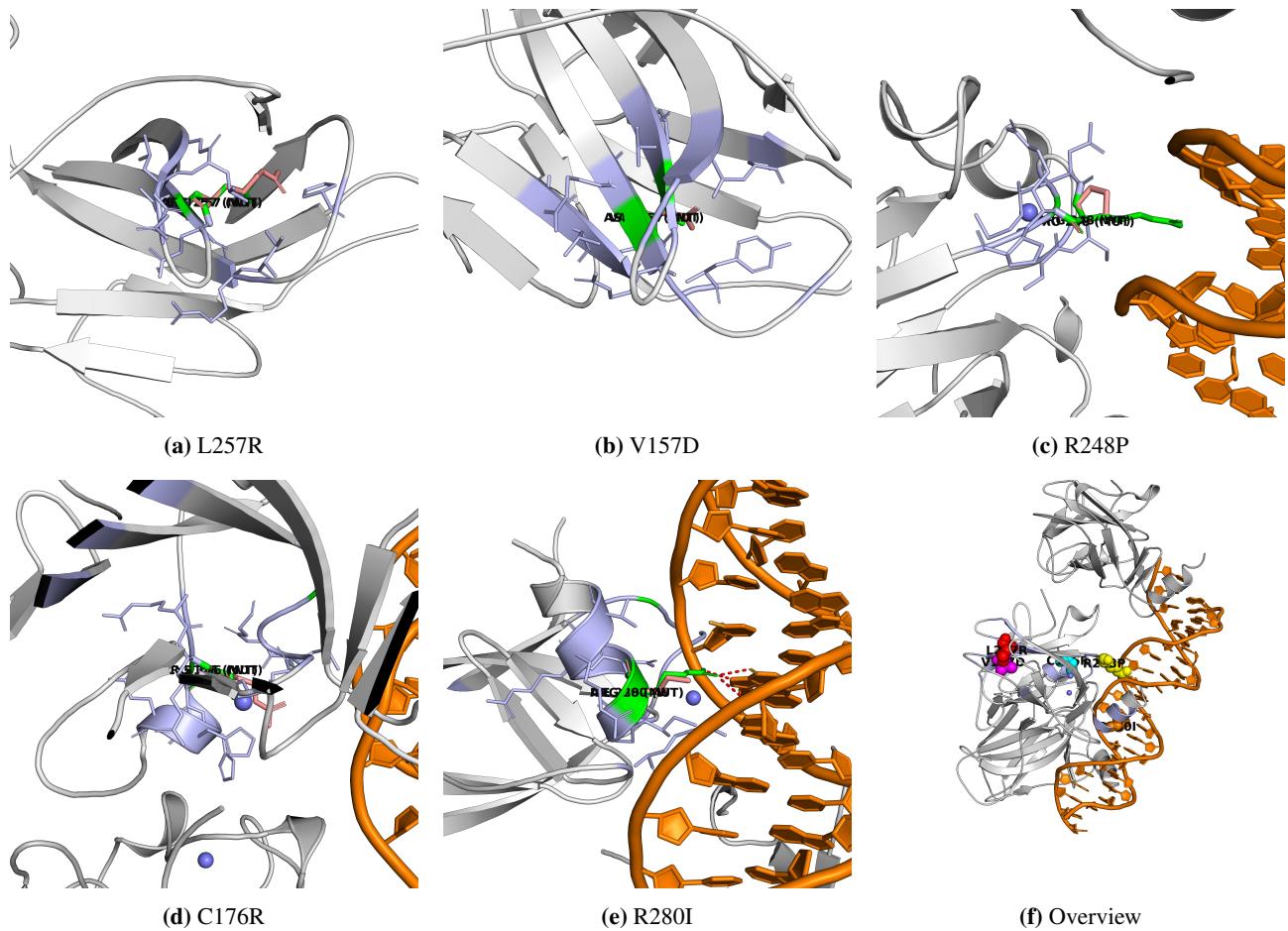


Figure 4: Publication-quality structural renders of the top five TP53 VUS (PDB 1TUP, chain B). Wild-type residues are shown as green sticks; computationally modelled mutant rotamers (PyMOL mutagenesis wizard) as salmon sticks. Neighbouring residues within 4.0 Å are shown as light-blue sticks. DNA polar 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.

379 Fig. S2)—confirm the spatial relationships described
380 above.

427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
plex; dynamic effects and post-translational modifications are not represented. Third, ESM-2 LLR thresholds (≤ -4.0) and AlphaMissense cutoffs (> 0.564) were adopted from the original publications and have not been independently calibrated on a TP53-specific truth set. Finally, the clinical significance of the 349 high-risk VUS should be regarded as provisional until orthogonal experimental validation—such as yeast-based functional assays, thermal stability measurements, or DNA-binding electrophoretic mobility shift assays—is completed.

381 4.3 Implications for Personalised Oncology

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

384 **Germline testing in Li-Fraumeni syndrome.** Li-
385 Fraumeni syndrome (LFS) is diagnosed by the iden-
386 tification of a pathogenic germline TP53 variant. In-
387 dividuals with LFS face a cumulative cancer risk ex-
388 ceeding 90% by age 60, and benefit from intensive
389 surveillance protocols including annual whole-body
390 MRI [Villani et al., 2016]. When a TP53 variant de-
391 tected on germline panel testing is classified as VUS,
392 the patient and their family members cannot be of-
393 fered definitive risk stratification. Reclassification
394 of high-confidence VUS to likely pathogenic would
395 directly enable cascade testing and early surveillance.

396 **Somatic tumour profiling.** TP53 mutational status
397 is a key biomarker in haematological malignancies,
398 where it predicts resistance to chemoimmunotherapy
399 in chronic lymphocytic leukaemia [Zenz et al., 2010]
400 and adverse prognosis in myelodysplastic syndromes
401 [Bejar et al., 2011]. In solid tumours, TP53 status
402 informs prognosis and, increasingly, therapy selec-
403 tion in the context of synthetic lethality approaches.
404 Resolving VUS enables more precise molecular strat-
405 ification.

406 **Emerging p53-targeted therapies.** A new gen-
407 eration of therapeutics aims to restore or stabilise
408 mutant p53 function. Small molecules such as APR-
409 246 (eprenetapopt) and PC14586 (rezatapopt) have
410 entered clinical trials for tumours harbouring specific
411 TP53 mutations [Chen et al., 2021]. Accurate classi-
412 fication of TP53 variants is a prerequisite for patient
413 selection in these trials. Variants that destabilise the
414 protein fold (L257R, V157D) may respond to fold-
415 stabilising compounds, whereas those that abolish
416 DNA contact (R248P, R280I) may require distinct
417 therapeutic strategies.

418 4.4 Limitations

419 Several limitations should be noted. First, *in silico*
420 predictions, regardless of model concordance, do not
421 constitute functional evidence under ACMG/AMP
422 criteria (PP3) and cannot alone support a pathogenic
423 classification beyond “supporting” evidence strength.
424 Second, our structural analysis is based on a single
425 crystal structure (PDB 1TUP) that captures only one
426 conformational state of the p53 tetramer–DNA com-

431 4.5 Future Directions

432 This work motivates several follow-up investigations.
433 The 349 concordant high-risk variants represent a pri-
434 oritised set for experimental functional validation, po-
435 tentially through high-throughput approaches such as
436 multiplexed assays of variant effect (MAVEs) [Find-
437 lay et al., 2018]. Integration of additional *in silico*
438 tools—including EVE [Frazer et al., 2021], REVEL
439 [Ioannidis et al., 2016], and molecular dynamics
440 simulations—could further refine the confidence tiers.
441 Longitudinal tracking of these variants in ClinVar will
442 reveal whether independent clinical evidence even-
443 tually converges with the AI predictions presented
444 here, providing a natural validation of the ensemble
445 approach.

5 Conclusions

453 We demonstrate that ensemble AI scoring using ESM-
454 2 and AlphaMissense can systematically identify high-
455 confidence pathogenic variants among the 1,211 TP53
456 VUS currently in ClinVar. The 349 concordant high-
457 risk variants—and particularly the five extreme can-
458 didates (L257R, V157D, R248P, C176R, R280I)—
459 exhibit both computational and structural hallmarks
460 of loss of function. These findings support the integra-
461 tion of orthogonal AI models as a scalable component
462 of variant classification pipelines, with direct impli-
463 cations for germline testing, somatic profiling, and
464 patient selection for emerging p53-targeted therapies
465 in precision oncology.

Data Availability

467 All analysis scripts, scored variant data, and structural
468 renders are available in the project repository. ClinVar
469 data were accessed via NCBI Entrez. AlphaMissense
470 proteome-wide predictions are available from Zen-

⁴⁷² odo (record 10813168). The PDB structure 1TUP is
⁴⁷³ available from the RCSB Protein Data Bank.

⁴⁷⁴ **Acknowledgements**

⁴⁷⁵ The author thanks the developers of ESM-2 (Meta AI),
⁴⁷⁶ AlphaMissense (Google DeepMind), ClinVar (NCBI),
⁴⁷⁷ BioPython, and PyMOL for making their tools and
⁴⁷⁸ data freely accessible. The 1TUP crystal structure
⁴⁷⁹ was obtained from the RCSB Protein Data Bank. This
⁴⁸⁰ project was developed as an independent computa-
⁴⁸¹ tional biology research effort.

⁴⁸² **Author Contact Information**

- ⁴⁸³ • **Author:** Mahad Asif
⁴⁸⁴ • **Email:** mahaddevx@gmail.com
⁴⁸⁵ • **GitHub:** github.com/mahaddev-x
⁴⁸⁶ • **Project Repository:** [github.com/mahaddev-](https://github.com/mahaddev-x/TP53-VUS-Predict)
⁴⁸⁷ [x/TP53-VUS-Predict](https://github.com/mahaddev-x/TP53-VUS-Predict)

⁴⁸⁸ All source code, scored variant tables, structural anal-
⁴⁸⁹ ysis outputs, and PyMOL render scripts used in this
⁴⁹⁰ study are publicly available in the project repository
⁴⁹¹ above.

492 Supplementary Figures

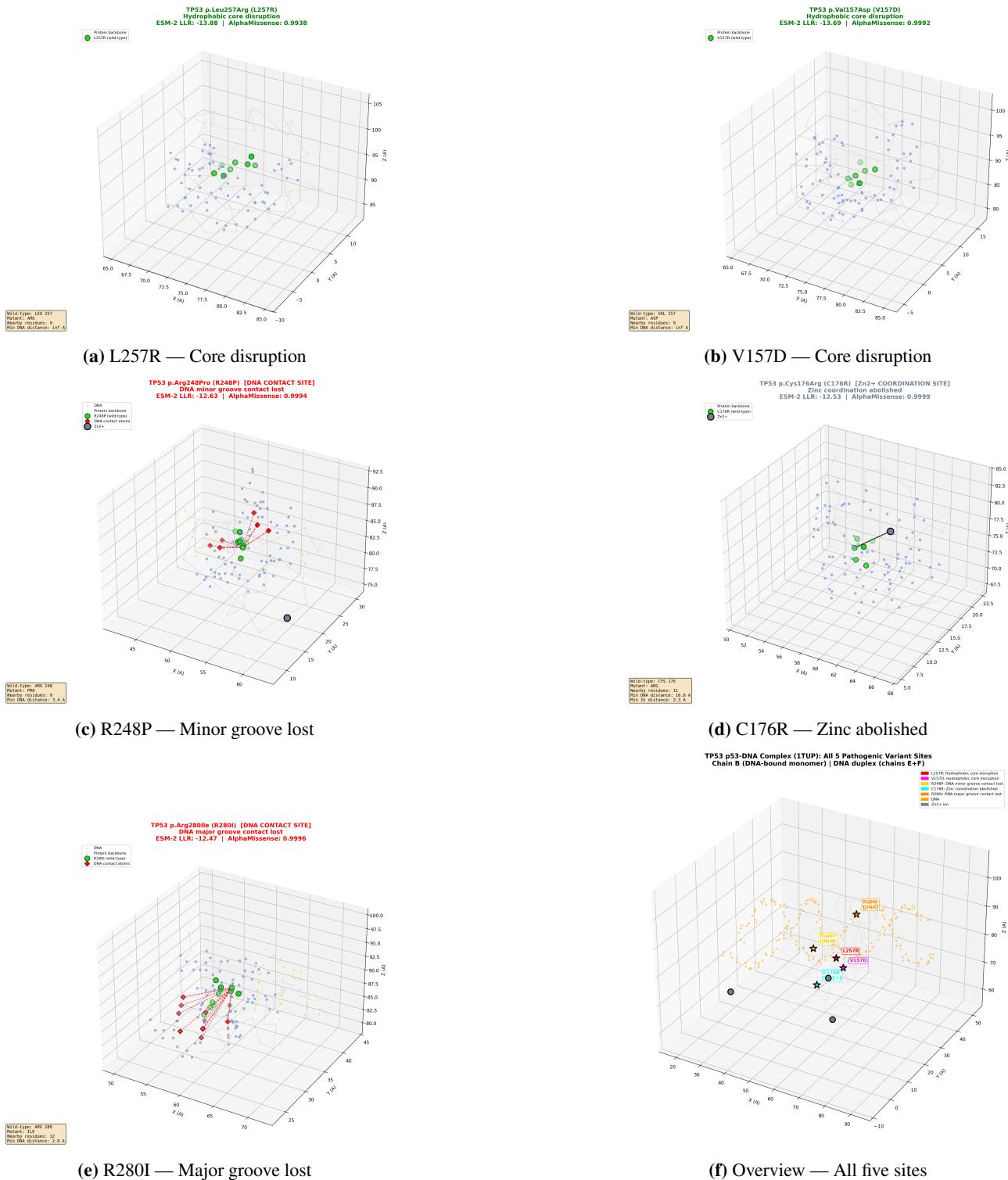


Figure S1: Supplementary Figure S1: Three-dimensional structural context of the top five TP53 VUS (BioPython analysis). Each panel shows a 12 Å radius view centred on the variant residue (green) in PDB 1TUP chain B. Orange: DNA; light gray: C α backbone; cornflower blue: neighbours within 4.0 Å; red 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.

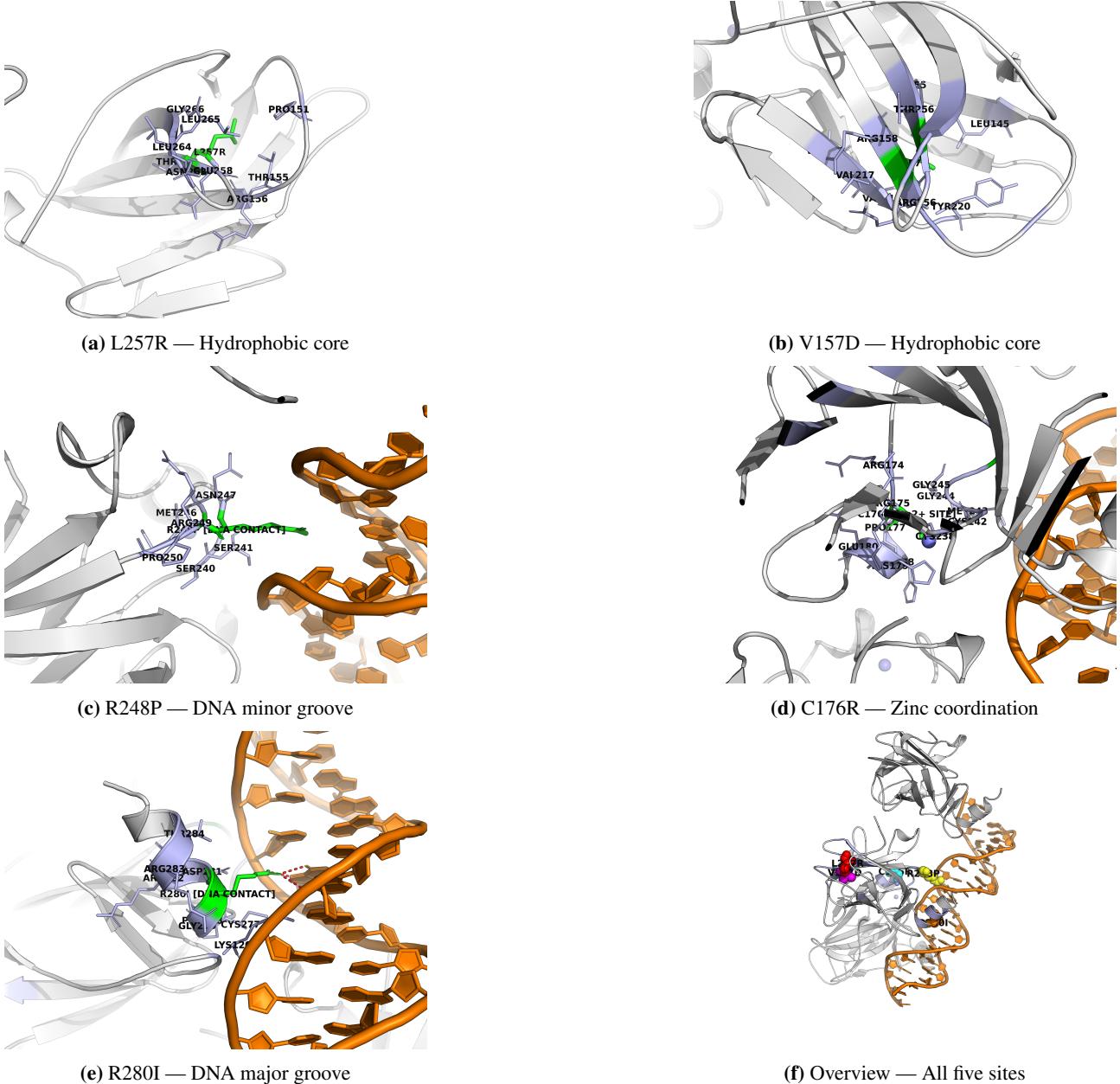


Figure S2: Supplementary Figure S2: Initial PyMOL ray-traced renders of the top five TP53 VUS (single-residue view). Generated via a scripted .pml pipeline from PDB 1TUP. Each panel shows the wild-type residue (green sticks) in context: protein cartoon (gray), DNA (orange), Zn²⁺ (slate spheres). Neighbours 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.

493 References

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

- 499 p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene*, 19(10):1245–1256,
500 2000.
- 501 Chen, S., Wu, J.-L., Liang, Y., et al. Small molecule therapeutics for TP53-mutant cancers. *Trends Pharma-*
502 *col. Sci.*, 42(12):1049–1062, 2021.
- 503 Cheng, J., Novati, G., Pan, J., et al. Accurate proteome-wide missense variant effect prediction with AlphaMis-
504 sense. *Science*, 381(6664):eadg7492, 2023.
- 505 Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. Crystal structure of a p53 tumor suppressor–DNA complex:
506 understanding tumorigenic mutations. *Science*, 265(5170):346–355, 1994.
- 507 Cock, P. J. A., Antao, T., Chang, J. T., et al. Biopython: freely available Python tools for computational molecular
508 biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 2009.
- 509 Findlay, G. M., Daza, R. M., Martin, B., et al. Accurate classification of BRCA1 variants with saturation genome
510 editing. *Nature*, 562(7726):217–222, 2018.
- 511 Frazer, J., Notin, P., Dias, M., et al. Disease variant prediction with deep generative models of evolutionary data.
512 *Nature*, 599(7883):91–95, 2021.
- 513 Giacomelli, A. O., Yang, X., Lintber, R. E., et al. Mutational processes shape the landscape of TP53 mutations
514 in human cancer. *Nat. Genet.*, 50(10):1381–1387, 2018.
- 515 Ioannidis, N. M., Rothstein, J. H., Pejaver, V., et al. REVEL: an ensemble method for predicting the pathogenicity
516 of rare missense variants. *Am. J. Hum. Genet.*, 99(4):877–885, 2016.
- 517 Kandoth, C., McLellan, M. D., Vandin, F., et al. Mutational landscape and significance across 12 major cancer
518 types. *Nature*, 502(7471):333–339, 2013.
- 519 Landrum, M. J., Lee, J. M., Benson, M., et al. ClinVar: improving access to variant interpretations and supporting
520 evidence. *Nucleic Acids Res.*, 46(D1):D1062–D1067, 2018.
- 521 Lane, D. P. p53, guardian of the genome. *Nature*, 358(6381):15–16, 1992.
- 522 Levine, A. J. & Oren, M. The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer*, 9(10):749–758,
523 2009.
- 524 Lin, Z., Akin, H., Rao, R., et al. Evolutionary-scale prediction of atomic-level protein structure with a language
525 model. *Science*, 379(6637):1123–1130, 2023.
- 526 Malkin, D., Li, F. P., Strong, L. C., et al. Germ line p53 mutations in a familial syndrome of breast cancer,
527 sarcomas, and other neoplasms. *Science*, 250(4985):1233–1238, 1990.
- 528 The PyMOL Molecular Graphics System, Version 3.1, Schrödinger, LLC.
- 529 Richards, S., Aziz, N., Bale, S., et al. Standards and guidelines for the interpretation of sequence variants: a joint
530 consensus recommendation of the American College of Medical Genetics and Genomics and the Association
531 for Molecular Pathology. *Genet. Med.*, 17(5):405–424, 2015.
- 532 Villani, A., Shore, A., Wasserman, J. D., et al. Biochemical and imaging surveillance in germline TP53 mutation
533 carriers with Li–Fraumeni syndrome: 11 year follow-up of a prospective observational study. *Lancet Oncol.*,
534 17(9):1295–1305, 2016.
- 535 Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature*, 408(6810):307–310, 2000.
- 536 Zenz, T., Eichhorst, B., Busch, R., et al. TP53 mutation and survival in chronic lymphocytic leukemia.
537 *J. Clin. Oncol.*, 28(29):4473–4479, 2010.