

Ensemble AI Screening of 1,211 TP53 Variants of Uncertain Significance via ESM-2 and AlphaMissense

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 variants
4 in ClinVar remain classified as Variants of Uncertain
5 Significance (VUS), limiting their clinical utility
6 in precision oncology.

7 **Methods.** We applied two orthogonal
8 deep-learning models—Meta’s ESM-2
9 (esm2_t33_650M_UR50D; 650 M parameters)
10 and DeepMind’s AlphaMissense—to screen 1,211
11 TP53 ClinVar VUS. Each variant was scored by
12 ESM-2 and cross-referenced against AlphaMissense
13 predictions (UniProt P04637), yielding 1,199
14 successfully matched variants. Structural validation
15 was performed against PDB 1TUP (Cho et al., 1994).

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

29 **Conclusions.** Ensemble AI scoring can systematically reclassify TP53 VUS at scale. The 349 high-confidence pathogenic variants identified here warrant prioritized functional validation and may inform germline testing guidelines in Li–Fraumeni syndrome and somatic profiling in clinical oncology.

35 **Keywords:** TP53, variants of uncertain significance,
36 ESM-2, AlphaMissense, pathogenicity prediction,
37 precision oncology, protein language model

38 1 Introduction

39 The *TP53* gene encodes the tumour protein p53, a transcription factor widely characterised as the “Guardian
40 of the Genome” for its central role in maintaining genomic integrity [Lane, 1992]. In response to genotoxic stress, p53 activates cell-cycle arrest, DNA repair, senescence, and apoptosis through sequence-specific DNA binding at target promoters [Vogelstein et al., 2000, Levine & Oren, 2009]. Loss-of-function mutations in TP53 are the single most common genetic alteration across all human cancers, observed in over 50% of solid tumours [Kandoth et al., 2013], and germline TP53 mutations are the molecular basis of Li–Fraumeni syndrome, a hereditary cancer predisposition disorder conferring near-complete lifetime cancer penetrance [Malkin et al., 1990].

54 The clinical interpretation of TP53 missense variants has been greatly advanced by large-scale repositories such as ClinVar [Landrum et al., 2018] and the IARC TP53 Database [Bouaoun et al., 2016]. However, a substantial fraction of observed variants—over 1,200 unique missense substitutions in ClinVar at the time of this analysis—remain classified as Variants of Uncertain Significance (VUS). The VUS designation creates a “diagnostic grey zone” for clinicians: these variants cannot be used to guide treatment decisions, genetic counselling, or cascade family testing, even when the variant resides in a functionally critical domain [Richards et al., 2015].

67 The bottleneck in VUS resolution is fundamentally one of evidence accumulation. Under the ACMG/AMP framework [Richards et al., 2015], reclassification requires convergent evidence from population 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 1,211 TP53 ClinVar VUS, cross-referencing both models’ predictions to identify high-confidence pathogenic candidates. 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 ~ 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

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

2.5 Structural Validation

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

Two complementary structural analysis approaches were employed:

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

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

2.6 Computational Environment

Primary analysis was performed on a consumer laptop running Windows. ESM-2 inference utilised local GPU acceleration via PyTorch with CUDA. AlphaMissense data were obtained from Google Cloud Storage / Zenodo. All analysis scripts were imple-

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 (LLR ≤ -4.0), 232 (19.3%) as likely damaging, 282 (23.5%) as possibly damaging, and 305 (25.4%) as likely neutral.

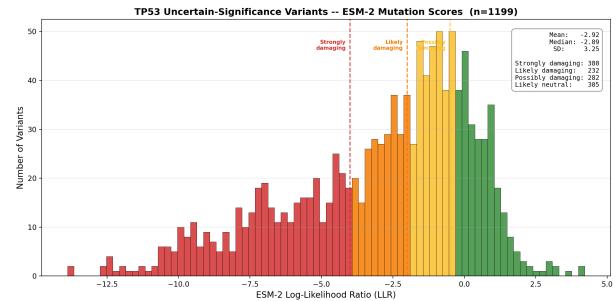


Figure 1: Distribution of ESM-2 log-likelihood ratio (LLR) scores across 1,199 TP53 VUS. Dashed vertical lines indicate classification thresholds. Variants with LLR ≤ -4.0 are designated strongly damaging.

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.

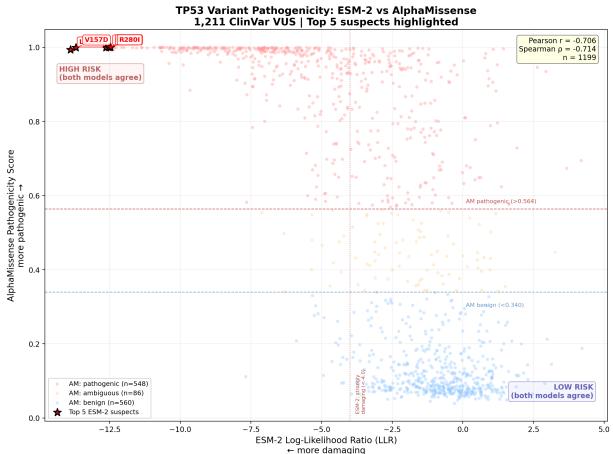


Figure 2: Scatter plot of ESM-2 LLR versus AlphaMissense pathogenicity score for 1,199 matched TP53 VUS. The strong anti-correlation (Pearson $r = -0.706$; Spearman $\rho = -0.714$) supports complementary predictive capacity between the sequence-based and structure-based models.

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).

3.3 Domain-Level Enrichment

246 Variant pathogenicity scores were non-uniformly dis-
247 tributed across TP53 functional domains (Fig. 3). The
248 DNA-binding domain (residues 102–292) showed sig-
249 nificantly more negative mean ESM-2 LLR scores
250 than the transactivation domain (residues 1–92) or 280
252 the tetramerization domain (residues 325–356), con- 281
253 sistent with stronger evolutionary constraint on the
254 DNA-binding interface. All five top-ranked variants 282
255 mapped to the DNA-binding domain.

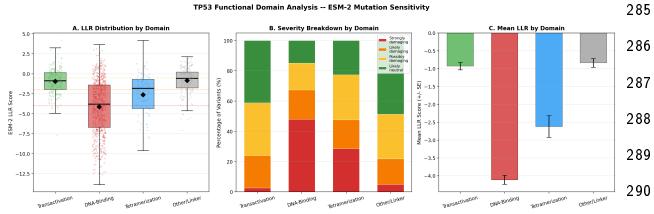


Figure 3: ESM-2 LLR score distribution by TP53 functional domain. The DNA-binding domain (residues 102–292) harbours the most severely scored variants, consistent with its critical role in tumour suppression.

3.4.4 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).

293
294
295
296
297

298
299
300
301

256 **3.4 Top Five Candidate Pathogenic Variants**

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

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

302 3.4.5 R280I (p.Arg280Ile): Loss of DNA Major 337 4.2 Structural Basis of Predicted Pathogenic- 303 Groove Contact 338 ity

304 Arginine 280 forms a direct hydrogen bond with a 339
305 guanine base in the DNA major groove at a distance 340
306 of approximately 2.8 Å. This contact is essential for 341
307 sequence-specific DNA recognition. Substitution to 342
308 isoleucine, a hydrophobic residue with no hydrogen- 343
309 bonding capacity, abolishes this interaction entirely 344
310 (LLR = −12.47; AM = 0.9996) (Fig. 4e). 345
346
347
348
349
350
351

311 4 Discussion

312 4.1 Concordant AI Prediction as Evidence 352 313 for Reclassification 353

314 The strong anti-correlation between ESM-2 and Al- 355
315 phaMissense predictions ($r = -0.706$; $\rho = -0.714$) 356
316 is noteworthy because the two models were trained on 357
317 fundamentally different data representations. ESM-2 358
318 is a pure sequence model that learns evolutionary con- 359
319 straints from 250 million protein sequences without 360
320 any explicit structural information [Lin et al., 2023]. 361
321 AlphaMissense, by contrast, incorporates AlphaFold2- 362
322 derived structural features alongside sequence context 363
323 and population frequency data [Cheng et al., 2023]. 364
324 The convergence of these orthogonal approaches on 365
325 the same set of high-risk variants provides a form of 366
326 computational triangulation analogous to the conver- 367
327 gent evidence required by the ACMG/AMP frame- 368
328 work [Richards et al., 2015]. 369

329 The 349 variants flagged as high-risk by both mod- 370
330 els represent 29.1% of the matched ClinVar VUS. Un- 371
331 der current clinical guidelines, none of these variants 372
332 are actionable. Our analysis suggests that a substan- 373
333 tial subset—particularly those with extreme scores 374
334 such as the five candidates highlighted here—carry 375
335 sufficient *in silico* evidence to support provisional 376
336 reclassification, pending functional confirmation. 377

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 abolition** (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 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 renders at multiple levels of detail (Fig. 4; Supplementary

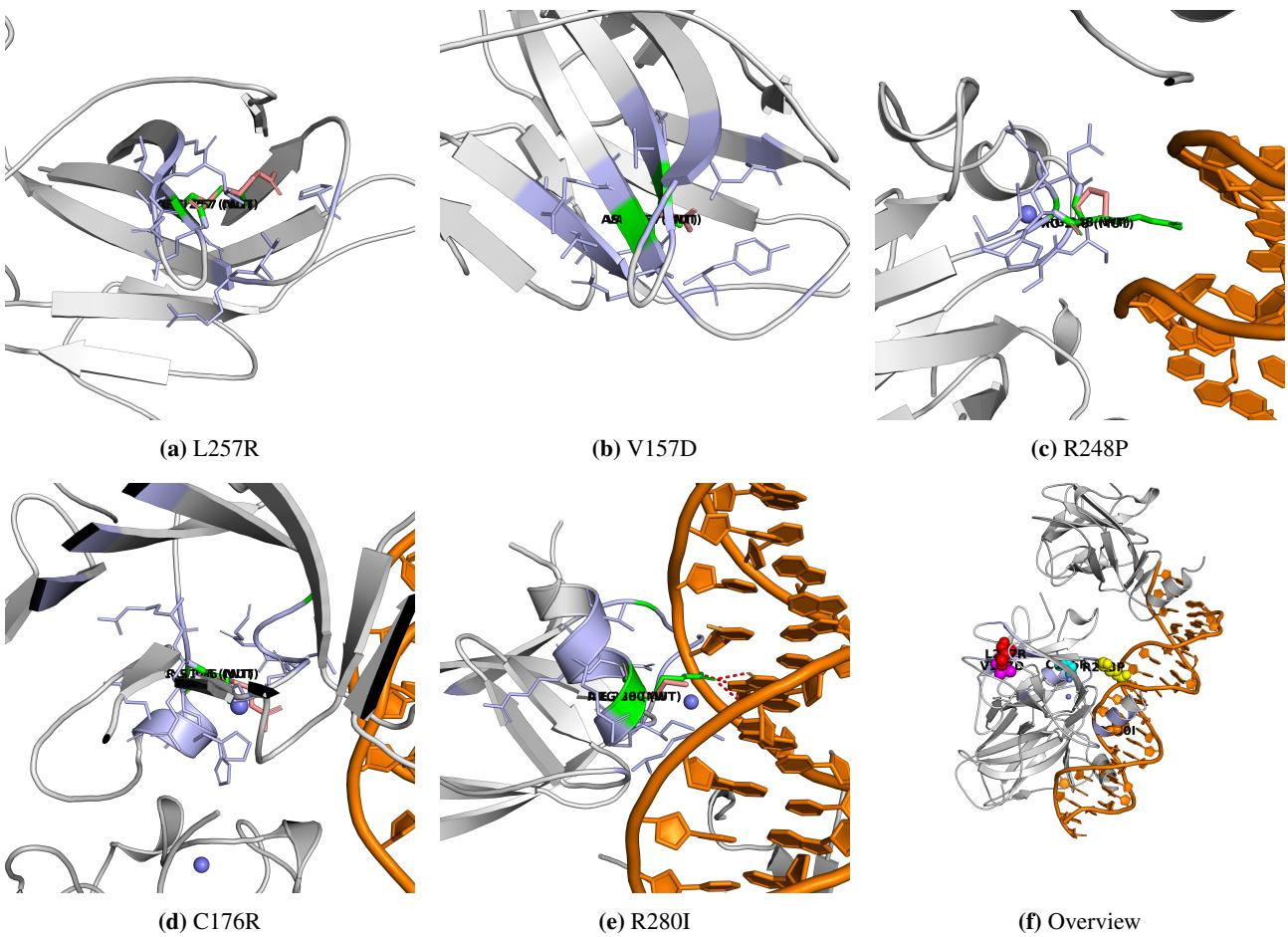


Figure 4: Publication-quality structural renders of the top five TP53 VUS (PDB 1TUP, chain B). Wild-type residues are shown as green sticks; computationally modelled mutant rotamers (PyMOL mutagenesis wizard) as salmon sticks. Neighbouring residues within 4.0 Å are shown as light-blue sticks. DNA polar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

Fig. S2)—confirm the spatial relationships described above. directly enable cascade testing and early surveillance.

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.

Emerging p53-targeted therapies. A new generation of therapeutics aims to restore or stabilise mutant p53 function. Small molecules such as APR-246 (eprenetapopt) and PC14586 (rezatapopt) have entered clinical trials for tumours harbouring specific TP53 mutations [Chen et al., 2021]. Accurate classification of TP53 variants is a prerequisite for patient

412 selection in these trials. Variants that destabilise the 458
413 protein fold (L257R, V157D) may respond to fold- 459
414 stabilising compounds, whereas those that abolish 460
415 DNA contact (R248P, R280I) may require distinct 461
416 therapeutic strategies. 462
463
464
465

417 4.4 Limitations

418 Several limitations should be noted. First, *in silico* 466
419 predictions, regardless of model concordance, do not 467
420 constitute functional evidence under ACMG/AMP 468
421 criteria (PP3) and cannot alone support a pathogenic 469
422 classification beyond “supporting” evidence strength. 470
423 Second, our structural analysis is based on a single 471
424 crystal structure (PDB 1TUP) that captures only one 472
425 conformational state of the p53 tetramer–DNA com- 473
426 plex; dynamic effects and post-translational modifica- 474
427 tions are not represented. Third, ESM-2 LLR thresh- 475
428 olds (≤ -4.0) and AlphaMissense cutoffs (> 0.564) 476
429 were adopted from the original publications and have 477
430 not been independently calibrated on a TP53-specific 478
431 truth set. Finally, the clinical significance of the 349 479
432 high-risk VUS should be regarded as provisional until 480
433 orthogonal experimental validation—such as yeast- 481
434 based functional assays, thermal stability measure- 482
435 ments, or DNA-binding electrophoretic mobility shift 483
436 assays—is completed. 484

437 4.5 Future Directions

438 This work motivates several follow-up investigations. 481
439 The 349 concordant high-risk variants represent a pri- 482
440 oritised set for experimental functional validation, po- 483
441 tentially through high-throughput approaches such as 484
442 multiplexed assays of variant effect (MAVEs) [Find- 485
443 lay et al., 2018]. Integration of additional *in silico* 486
444 tools—including EVE [Frazer et al., 2021], REVEL 487
445 [Ioannidis et al., 2016], and molecular dynamics 488
446 simulations—could further refine the confidence tiers. 489
447 Longitudinal tracking of these variants in ClinVar will 490
448 reveal whether independent clinical evidence even- 491
449 tually converges with the AI predictions presented 492
450 here, providing a natural validation of the ensemble 493
451 approach. 494

452 5 Conclusions

453 We demonstrate that ensemble AI scoring using ESM- 495
454 2 and AlphaMissense can systematically identify high- 496
455 confidence pathogenic variants among the 1,211 TP53 497
456 VUS currently in ClinVar. The 349 concordant high- 498
457 risk variants—and particularly the five extreme can- 499

dicates (L257R, V157D, R248P, C176R, R280I)— 460
461 exhibit both computational and structural hallmarks 462
463 of loss of function. These findings support the integra- 464
465 tion of orthogonal AI models as a scalable component 466
467 of variant classification pipelines, with direct impli- 468
469 cations for germline testing, somatic profiling, and 470
471 patient selection for emerging p53-targeted therapies 472
473 in precision oncology. 474

Data Availability

All analysis scripts, scored variant data, and structural renders are available in the project repository. ClinVar data were accessed via NCBI Entrez. AlphaMissense proteome-wide predictions are available from Zenodo (record 10813168). The PDB structure 1TUP is available from the RCSB Protein Data Bank.

Acknowledgements

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

Author Contact Information

- **Author:** Mahad Asif
- **Email:** mahaddevx@gmail.com
- **GitHub:** github.com/mahaddev-x
- **Project Repository:** github.com/mahaddev-x/TP53-VUS-Predict

All source code, scored variant tables, structural analysis outputs, and PyMOL render scripts used in this study are publicly available in the project repository above.

References

- Bejar, R., Stevenson, K., Abdel-Wahab, O., et al. Clinical effect of point mutations in myelodysplastic syndromes. *N. Engl. J. Med.*, 364(26):2496–2506, 2011.
- Bouaoun, L., Sonber, D., Ardin, M., et al. TP53 variations in human cancers: new lessons from the IARC TP53 Database and genomics data. *Hum. Mutat.*, 37(9):865–876, 2016.

- 500 Bullock, A. N., Henckel, J. & Fersht, A. R. Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene*, 19(10):1245–1256, 2000.
- 505 Chen, S., Wu, J.-L., Liang, Y., et al. Small molecule therapeutics for TP53-mutant cancers. *Trends Pharmacol. Sci.*, 42(12):1049–1062, 2021.
- 508 Cheng, J., Novati, G., Pan, J., et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*, 381(6664):eadg7492, 2023.
- 512 Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. Crystal structure of a p53 tumor suppressor–DNA complex: understanding tumorigenic mutations. *Science*, 265(5170):346–355, 1994.
- 516 Cock, P. J. A., Antao, T., Chang, J. T., et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 2009.
- 520 Findlay, G. M., Daza, R. M., Martin, B., et al. Accurate classification of BRCA1 variants with saturation genome editing. *Nature*, 562(7726):217–222, 2018.
- 524 Frazer, J., Notin, P., Dias, M., et al. Disease variant prediction with deep generative models of evolutionary 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 in human cancer. *Nat. Genet.*, 50(10):1381–1387, 2018.
- 531 Ioannidis, N. M., Rothstein, J. H., Pejaver, V., et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am. J. Hum. Genet.*, 99(4):877–885, 2016.
- 535 Kandoth, C., McLellan, M. D., Vandin, F., et al. Mutational landscape and significance across 12 major cancer types. *Nature*, 502(7471):333–339, 2013.
- 538 Landrum, M. J., Lee, J. M., Benson, M., et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.*, 46(D1):D1062–D1067, 2018.
- 542 Lane, D. P. p53, guardian of the genome. *Nature*, 358(6381):15–16, 1992.
- 544 Levine, A. J. & Oren, M. The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer*, 9(10):749–758, 2009.
- 549 Lin, Z., Akin, H., Rao, R., et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.
- 551 Malkin, D., Li, F. P., Strong, L. C., et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, 250(4985):1233–1238, 1990.
- 553 The PyMOL Molecular Graphics System, Version 3.1, Schrödinger, LLC.
- 555 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.*, 17(5):405–424, 2015.
- 558 Villani, A., Shore, A., Wasserman, J. D., et al. Biochemical and imaging surveillance in germline TP53 mutation carriers with Li–Fraumeni syndrome: 11 year follow-up of a prospective observational study. *Lancet Oncol.*, 17(9):1295–1305, 2016.
- 562 Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature*, 408(6810):307–310, 2000.
- 564 Zenz, T., Eichhorst, B., Busch, R., et al. TP53 mutation and survival in chronic lymphocytic leukemia. *J. Clin. Oncol.*, 28(29):4473–4479, 2010.

574 **Supplementary Figures**

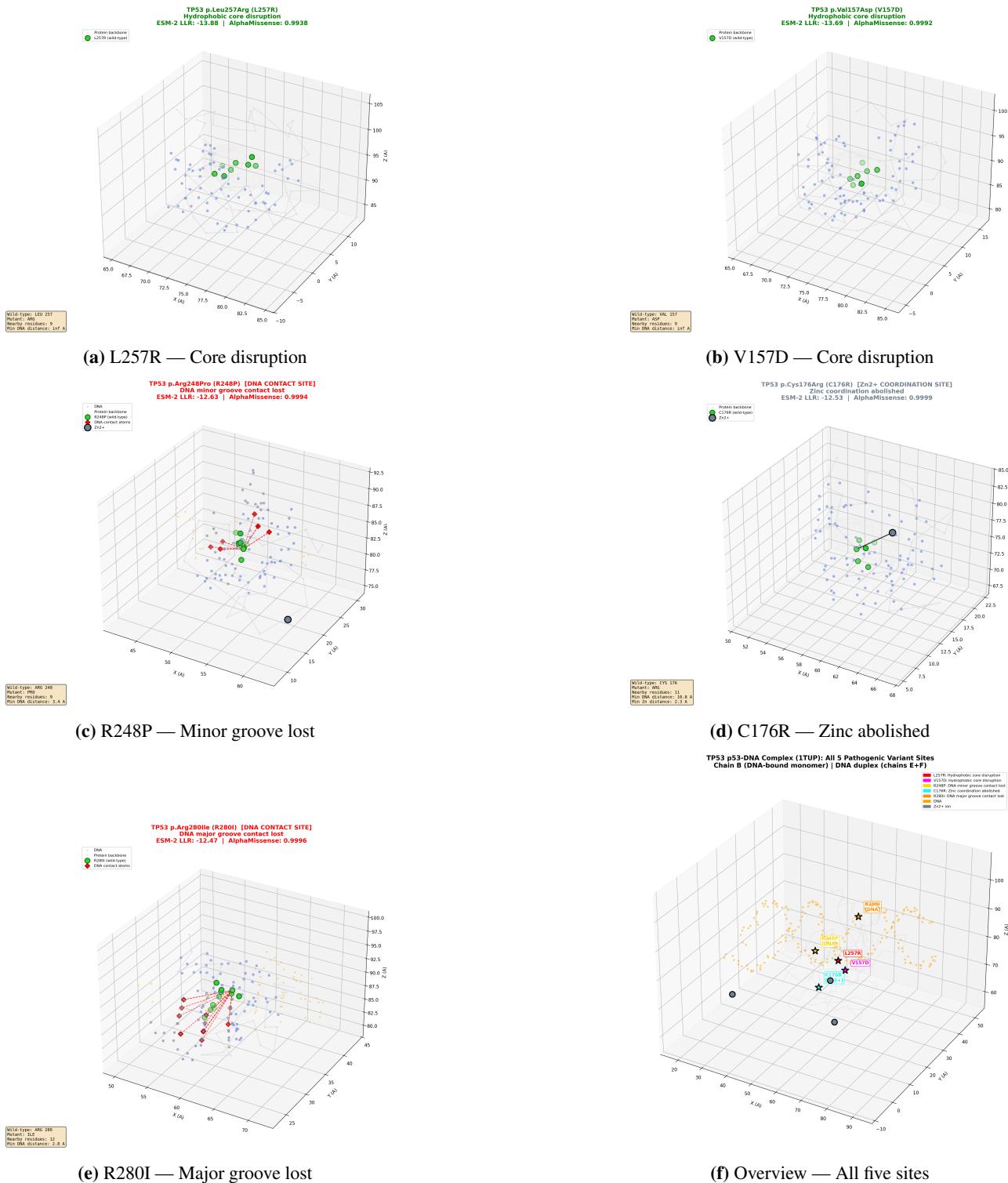


Figure S1: Supplementary Figure S1: Three-dimensional structural context of the top five TP53 VUS (BioPython analysis). Each panel shows a 12 Å radius view centred on the variant residue (green) in PDB 1TUP chain B. Orange: DNA; light gray: C α backbone; cornflower blue: neighbours within 4.0 Å; red diamonds: DNA contacts; slate: Zn²⁺. (a) L257R buried, no DNA contact. (b) V157D buried in β -sandwich. (c) R248P DNA minor groove contact. (d) C176R Zn²⁺ coordination. (e) R280I DNA major groove contact. (f) Overview of all five positions on chain B.

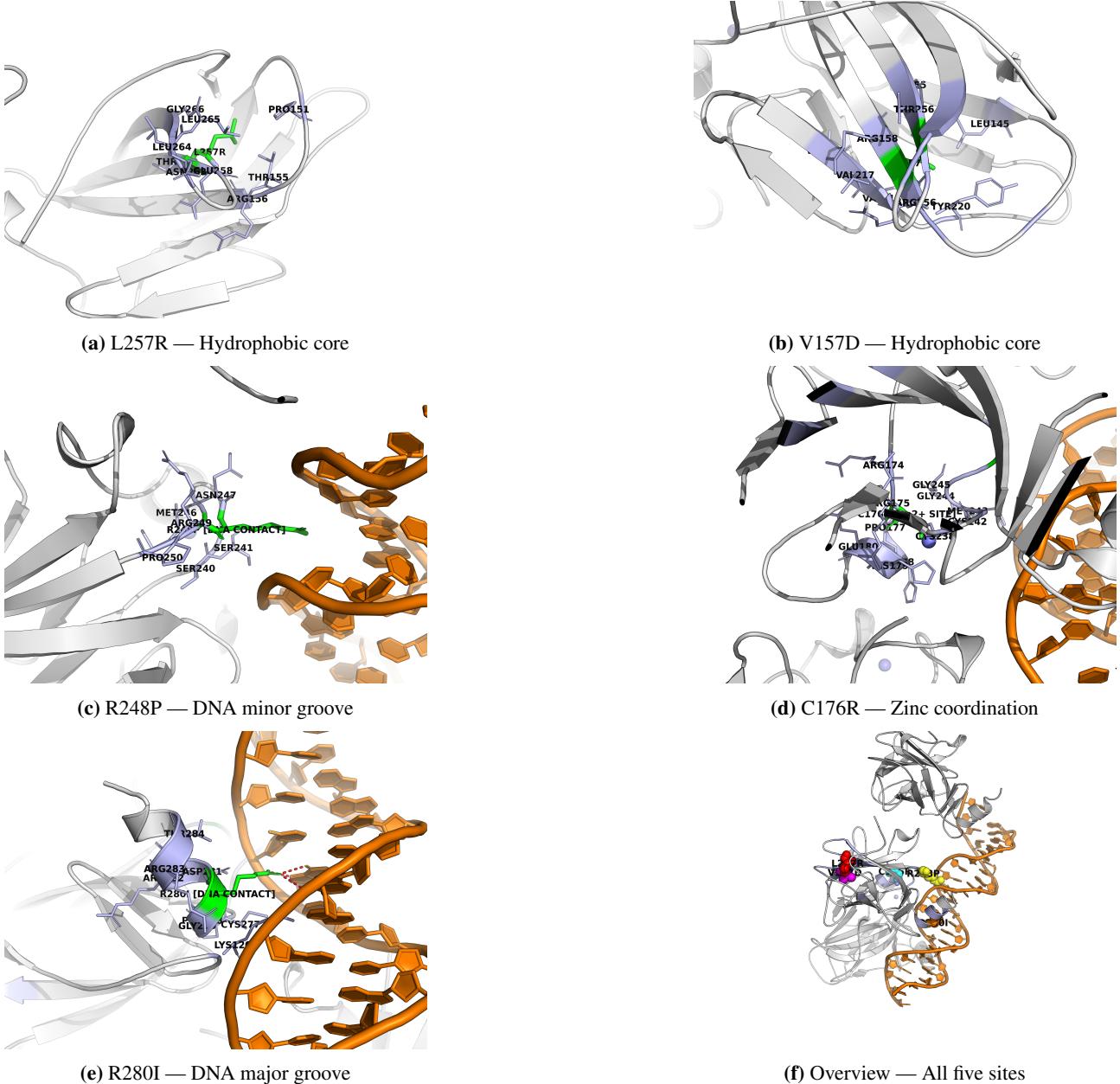


Figure S2: Supplementary Figure S2: Initial PyMOL ray-traced renders of the top five TP53 VUS (single-residue view). Generated via a scripted .pml pipeline from PDB 1TUP. Each panel shows the wild-type residue (green sticks) in context: protein cartoon (gray), DNA (orange), Zn²⁺ (slate spheres). Neighbours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