

Evolutionary Analysis of PD-L1 as a Small-Molecule Drug Target

Zvikomborero Mangwanda

christianmangwanda@gmail.com

Clarkson University <https://orcid.org/0009-0002-3293-8151>

Research Article

Keywords:

Posted Date: October 13th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11341956/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Evolutionary Analysis of PD-L1 as a Small-Molecule Drug Target

Christian Mangwanda

Coulter School of Engineering & Applied Sciences, Department of Mathematics
Clarkson University

Abstract

Small molecules that block PD-L1 bind a pocket that exists only when two PD-L1 molecules pair up, and oral compounds acting this way are in phase 2 trials. I asked whether a tumour could escape such a compound by mutating that pocket. Working from crystal structures and a codon alignment of 44 vertebrate orthologs of *CD274* spanning about 429 million years, I find the drug pocket and the PD-1 binding interface are the same surface: all 10 drug-contacting residues lie within 5.0 Å of PD-1, against 1.9 expected by chance (hypergeometric $p = 8.7 \times 10^{-9}$), and the result holds at every distance cutoff I tested. Those residues are under purifying selection. *CD274* evolves at $d_N/d_S = 0.236$, no pocket residue shows positive selection, and the pocket evolves more slowly than accessibility-matched residues elsewhere in the protein: median $d_N/d_S = 0.077$ against a permutation null median of 0.287, roughly 3.8-fold lower, $p = 0.0028$. That comparison stays below 0.05 across every robustness variant I ran, including every single-residue deletion. The overlap is what makes this useful: the residues a compound would have to lose are the residues PD-1 binds, so escape by pocket mutation is constrained by the very function the tumour expresses PD-L1 to perform.

1 Introduction

PD-1 is a receptor on T cells that essentially shuts down their attack mechanism when it is engaged. PD-L1 is a protein that engages this receptor and causes the T cells to switch off their attack mechanism. So, many tumours express PD-L1 a lot, which effectively puts the switch under their control, and they shut down the T cells that would otherwise be able to destroy them. Antibodies against PD-1 and PD-L1 are now standard treatment in several cancers around the world (Chen and Mellman, 2017; Zou et al., 2016).

Due to the fact that antibodies are quite large proteins that, in this case, are given via infusion, they cannot reach every compartment a tumour has occupied. Given this scenario, there has been long-standing interest in small molecules that can block the same type of contact and can be taken orally (Guzik et al., 2019).

BMS-202 and BMS-8 were the ones that opened this route. They bind to PD-L1 molecules and lock them together into a dimer, and this buries the surface of PD-L1 that would otherwise have been engaged by PD-1 (Zak et al., 2016).

Compounds of this class are now in patients. INCB086550 blocks the PD-L1/PD-1 interaction by inducing the same dimerization and internalisation (Koblish et al., 2022), and INCB099280 and ABR003 are in phase 2 studies in solid tumours (Incyte Corporation, 2023; University Health Network, 2025).

The question that I had was if a tumour can change the amino acids the compound touches, so that the compound stops working for that particular case. The question I had was whether the drug pocket and the

PD-1 binding interface are the same surface. If they are, then a tumour cannot look for an escape mutation among the amino acids PD-1 leaves free, because there would be none. The question I had is evolutionary: whether these amino acids have been held at fixed points across vertebrate evolution, which would be evidence that changing them is costly. As another angle to look at it, I asked whether human germline and tumour somatic variation offered any insights.

2 Methods

2.1 Structures

The pocket is the PD-L1 amino acids contacted by BMS-202 in 5J89 or BMS-8 in 5J8O. I took the union across both structures and all chains (Zak et al., 2016): F19, T20, I54, Y56, Q66, V76, M115, A121, Y123 and K124. All 10 contacts were profiled with PLIP (Adasme et al., 2021) and cross-checked against a 4.5 Å heavy-atom criterion. Due to the fact that BMS-202 and BMS-8 are different molecules, I took the union of the two. Both structures are registered one-to-one with UniProt Q9NZQ7.

Relative solvent accessibility was computed with DSSP (Kabsch and Sander, 1983) on chain A of 5C3T, the unbound PD-L1 monomer, and further normalised on the empirical maxima of Tien et al. (2013). The native window is residues 18 to 134, giving $n = 117$, with the expression tag at 135 to 143 excluded. Accessibility enters only as a continuous matching covariate, never as a grouping: across the window the relationship between accessibility and rate is a slope rather than a step (Spearman $\rho = 0.42$), which is why

I matched on it rather than binning (Franzosa and Xia, 2009; Ramsey et al., 2011).

Instead of just taking a published crystal list or amino acid list, I derived the PD-1 interface from coordinates in 4ZQK, chain A (which is PD-L1) and chain B (which is PD-1) (Zak et al., 2015). I measured the minimum heavy-atom distance from any chain A residue to any chain B atom. An amino acid is in the interface if the distance is at or below 5.0 Å. The denominator in this case was the 115 residues that were actually present in the chain A model. This is due to the fact that any residue that is missing from the structure has no chance of having any distance to measure and could not have been counted. I tested enrichment with the hypergeometric distribution and swept the cutoff so the one free parameter is visible.

2.2 Sequences

I used 44 vertebrate *CD274* orthologs. All of the sequences were annotated as PD-L1 with no PD-L2 flags. Of these 44, 17 were primates, 13 were other mammals, 6 were birds, 4 were turtles, and 4 were fish, spanning about 429 million years of divergence, the crown Osteichthyes date (Broughton et al., 2013; Kumar et al., 2022). I corrected 16 of 44 species labels, reassigning them from NCBI taxonomy and confirming them by a BLAST search and again by clade monophyly in the tree.

I took coding sequences by parsing the `coded_by` qualifier of each protein record and fetching exactly that span. All 44 were able to be translated back to their original protein. I aligned these proteins with the help of MAFFT (Katoh and Standley, 2013) and then threaded nucleotides through that alignment, three at a time, giving 44 sequences over 535 codon columns.

Human PD-L1 has 290 amino acids, and the remaining columns exist only to hold insertions in other lineages. Therefore a column with no human residue cannot inform a claim about the human residue, so I restricted the alignment to just the 290 measurable columns before fitting any selection model.

The tree was estimated using IQ-TREE (Nguyen et al., 2015) under GTR+F+G4, selected by ModelFinder (Kalyaanamoorthy et al., 2017), with 1,000 bootstrap replicates and seed 20260101, rooted on the fish outgroup. I estimated it on the full 535-column alignment rather than the reference-mappable subset, because the tree from the subset fails primate monophyly. The estimated tree recovers primates, birds, turtles, Mammalia and Sauropsida as monophyletic groups (Figure 1).

2.3 Selection

I made use of FEL (Kosakovsky Pond and Frost, 2005) in HyPhy 2.5.101(MP), which estimated the site-specific synonymous and non-synonymous rates across the 290 codons, further corrected using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) across those 290 sites. Gene-level selection was as-

Every clade recovered from sequence alone

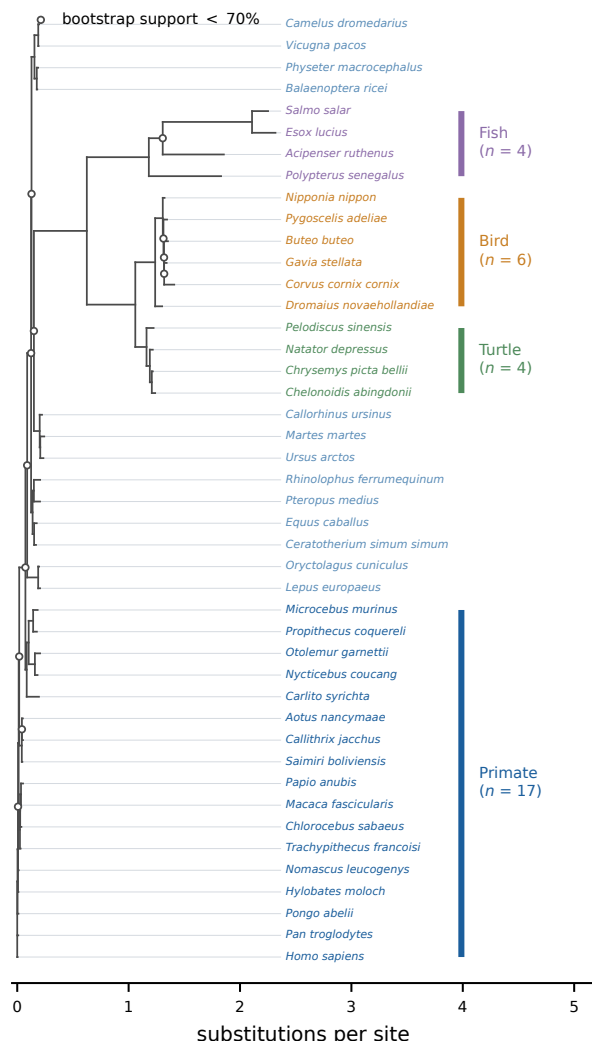


Figure 1: The phylogeny independently validates the corrected species labels. Maximum-likelihood tree of 44 vertebrate *CD274* orthologs, estimated from the 535-column codon alignment with no taxonomic input. Primates, birds, turtles, Mammalia and Sauropsida are all recovered as monophyletic, which a remaining mislabelled taxon would break. Brackets mark groups forming a single clade here; other-mammal tips carry colour but no bracket, because that group excludes the nested primates and so is paraphyletic by construction rather than a clade. Open circles mark the 12 internal nodes below 70 percent bootstrap support, all within shallow radiations. Branch lengths are drawn to scale, so the primate radiation is compressed relative to the deep fish branches; thin leaders connect each tip to its label.

sessed by the Nei-Gojobori method (Nei and Gojobori, 1986) over all pairwise comparisons among the 44 taxa.

I report the per-site d_N/d_S . Normalising by the synonymous rate matters because that rate varies from site to site in real alignments (Kosakovsky Pond and Muse, 2005), and ignoring this variation inflates false positives in selection analyses (Wisotsky et al., 2020). The ratio is undefined at the 2 window sites where the synonymous rate is zero, and those are dropped from the comparison pool; all 10 pocket residues have a non-zero synonymous rate.

For each drug-contacting residue I collected every non-pocket residue within ± 10 accessibility points. I then drew one at random, took the median of the 10 draws, and repeated this 50,000 times with seed 20260605. This test was one-tailed because the hypothesis is directional. The mean comparison pool is 27.1 residues, and the comparison is never pocket against whole protein, but 10 like-for-like comparisons at matched accessibility combined into one statistic.

The per-site d_N/d_S is heavily skewed where the synonymous rate approaches zero, so the mean of the permutation medians is unstable across random seeds while the median is stable to three decimal places. Therefore I quoted the median rather than the mean.

2.4 Variation

I made use of the UniProt variation API for Q9NZQ7 (Consortium, 2023), which gave me germline missense variants. Population frequencies were reported there, and a gene-level cross-check against gnomAD (Kar-czewski et al., 2020) was also conducted there.

I then took the somatic mutations in *CD274* from cBioPortal (Cerami et al., 2012) across 8 cohorts: seven TCGA PanCancer Atlas studies in the tumour types where PD-L1 blockade is used, plus the MSK-IMPACT targeted panel. Table 1 gives each cohort with its sequenced-sample denominator, 14,330 samples in total. I kept all coding mutation classes, and the counts are mutation rows rather than distinct samples.

3 Results

3.1 Overlap

Of the 115 PD-L1 residues modelled in chain A of 4ZQK, 22 lie within 5.0 Å of PD-1. All 10 drug-contacting residues are among them, against 1.9 expected for a randomly chosen set of ten (hypergeometric $p = 8.7 \times 10^{-9}$). Distances run from 2.68 Å to 4.57 Å, mean 3.37 Å, with the closest contacts at hydrogen-bonding distance. The pocket is one physical surface, not two adjacent patches (panel c of Figure 2, and Figure 3).

The distance convention does not carry the result (Table 2). At 4.0 Å the interface contains 17 residues and 8 of the pocket ($p = 1.6 \times 10^{-6}$); at 5.5 Å it contains 24 and all ten ($p = 2.6 \times 10^{-8}$). Every cutoff

Table 1: The 8 cBioPortal cohorts contributing somatic mutations in *CD274*, with the sequenced-sample denominator for each. Counts are mutation rows, not distinct samples; one MSK-IMPACT sample carries two. *CD274* is rarely mutated in coding sequence in every cohort, which is why the depletion test is underpowered.

Cohort	Sequenced	Mutated	Mutations
blca*	410	5	5
coadread*	534	3	3
hnsc*	515	2	2
luad*	566	4	4
lusc*	484	1	1
msk_impact_2017	10945	30	31
skcm*	440	5	5
stad*	436	4	4
Total	14,330	—	55

cBioPortal study id, shared suffix `_tcga_pan_can_atlas_2018` omitted.

Table 2: Overlap between the BMS pocket and the PD-1 interface is significant at every distance cutoff tested. Interface membership is minimum heavy-atom distance to PD-1 in 4ZQK; the denominator is the 115 residues modelled in chain A. p from the hypergeometric distribution.

Cutoff (Å)	Interface	Pocket	Expected	p
4.0	17	8/10	1.48	1.6×10^{-6}
4.5	18	9/10	1.57	6.4×10^{-8}
5.0	22	10/10	1.91	8.7×10^{-9}
5.5	24	10/10	2.09	2.6×10^{-8}

5.0 Å row is the definition adopted throughout.

gives a p -value at least six orders of magnitude below conventional thresholds.

3.2 Selection

Every drug-contacting amino acid has a non-synonymous rate below its synonymous rate, and none shows any signal of positive selection (Table 3, Figure 2b). At the gene level *CD274* is strongly constrained, with $d_N/d_S = 0.236$, while the mean d_N was 0.280 and the mean d_S was 1.188. That ratio pools 946 pairwise comparisons that share a phylogeny, so I ran it again on 19 branch-disjoint pairs, which share no branch of the tree. The result holds there as well ($Z = 4.62$, $p = 1.9 \times 10^{-6}$).

Therefore, as a group the pocket evolves more slowly than accessibility-matched residues everywhere in the protein. The observed pocket median d_N/d_S is 0.077 against a permutation null median of 0.287, which is a 3.8-fold lower rate ($p = 0.0028$, one-tailed, 50,000 draws). Amino acid by amino acid, 6 of 10 were slower than their own comparison pool (Figure 2a).

4 residues are individually significant after correction across the 290 analysed sites at $q < 0.1$: A121, T20, M115 and Y123. In the rest of the native window 42 of 107 residues clear the same correction, 39.3 percent against the pocket’s 40 percent.

Every robustness variant I ran stays below 0.05 (Ta-

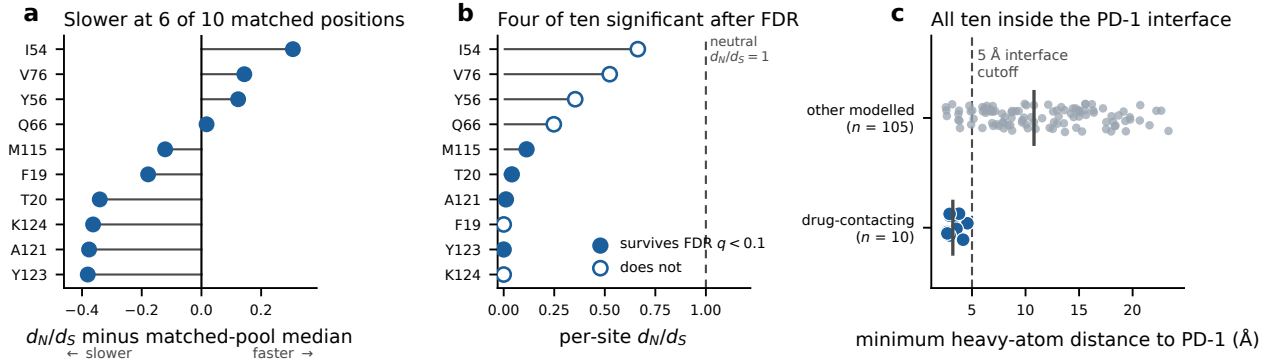


Figure 2: The drug pocket is evolutionarily constrained and coincides with the PD-1 interface. (a) Paired difference for each of the 10 drug-contacting codons: its per-site d_N/d_S minus the median of non-pocket residues within ± 10 accessibility points, which is the comparison the group-level test makes. Negative values are slower than matched comparables; 6 of 10 are. (b) Per-site d_N/d_S per drug-contacting codon, $n = 10$ codons. Every residue sits well left of neutrality. Filled markers survive Benjamini-Hochberg correction across the 290 analysed sites at $q < 0.1$ (4 of 10); open markers do not. (c) Minimum heavy-atom distance to PD-1 in 4ZQK, for the 10 drug-contacting residues and the other modelled residues of chain A ($n = 115$ modelled residues in total). Vertical ticks are group medians; the dashed line is the 5.0 Å interface definition.

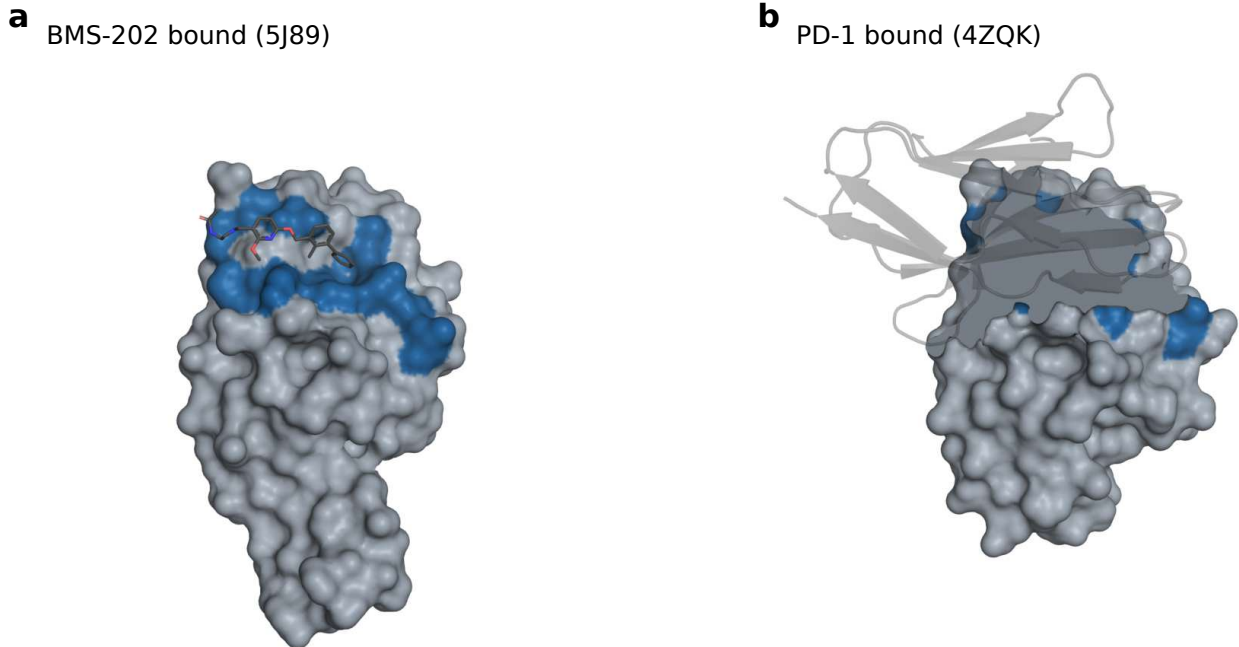


Figure 3: The drug pocket and the PD-1 interface are the same surface. (a) PD-L1 surface with the 10 drug-contacting residues coloured, with BMS-202 bound (5J89). (b) The same surface and colouring against PD-1 (4ZQK). All 10 drug-contacting residues lie within 4.57 Å of PD-1, and 22 of the 115 residues modelled in chain A are within 5.0 Å.

Table 3: The ten drug-contacting residues of the BMS pocket. Rates and q -values are from FEL on the 290 reference-mappable codons of a 44-taxon codon alignment, corrected across those 290 sites; PD-1 distances are minimum heavy-atom distances in 4ZQK chain A versus chain B. Relative solvent accessibility is from DSSP on the apo monomer 5C3T chain A on the Tien scale. $n = 10$ codons. d_N and d_S are the component rates the ratio is built from; Source is the structure or structures in which the contact is seen.

Residue	RSA (%)	d_N/d_S	d_N	d_S	p	q	PD-1 (Å)	Source
K124	30.9	0.000	0.000	0.375	0.0329	0.1110	2.91	both
Y123 [†]	44.1	0.000	0.000	0.501	0.0177	0.0705	2.70	both
F19	13.3	0.000	0.000	0.239	0.0717	0.1856	2.92	5J89
A121 [†]	52.7	0.011	0.052	4.695	0.0000	0.0000	2.68	both
T20 [†]	44.8	0.040	0.109	2.721	0.0000	0.0004	4.57	5J89
M115 [†]	21.9	0.113	0.202	1.787	0.0064	0.0377	3.79	both
Q66	20.4	0.248	0.678	2.734	0.0468	0.1358	2.91	both
Y56	21.3	0.354	0.468	1.323	0.2040	0.3819	3.55	both
V76	44.8	0.524	0.534	1.019	0.3516	0.5665	4.17	5J8O
I54	32.5	0.663	0.719	1.083	0.5935	0.7788	3.50	both

[†] survives Benjamini–Hochberg correction at $q < 0.1$ (4 of 10). All ten have $d_N/d_S < 1$; none shows positive selection.

Table 4: Sensitivity of the group-level comparison. Per-site d_N/d_S stays below 0.05 across every variant tested: both extremes of the matching band, every single-residue deletion, and the alternative phylogeny rejected on topological grounds.

Variant	d_N/d_S	Note
Primary (± 10 band)	0.0028	adopted
Band ± 5	0.00185	
Band ± 15	0.00295	
Band ± 20	0.00675	
Leave-one-out	0.00095–0.02525	widest tested
Alternative tree	0.0268	all ten significant rejected on topology

ble 4): both extremes of the matching band, each single-residue deletion, and the alternative phylogeny.

3.3 Variation

34 missense variants are recorded at the ten drug-contacting residues, 0 of them common at a 1 percent threshold. Variant density is the comparison with a denominator, and it runs in the direction of the hypothesis: 3.40 distinct missense variants per residue in the pocket against 4.14 elsewhere in the window ($p = 0.062$, one-sided).

Across 8 cohorts covering 14,330 sequenced samples (Table 1), *CD274* carries 55 somatic mutations, 19 in the analysed window and 1 at a drug-contacting residue, namely T20I, against 1.62 expected by chance (hypergeometric $p = 0.49$).

3.4 Q66

Q66 carries the highest-frequency germline variant in the pocket, it does not reach individual significance, and its d_N/d_S is among the highest of the ten. It still sits 2.91 Å from PD-1, so change there is not free. That makes a testable prediction: if escape occurs, Q66 is where to look first, and A121 and T20 are

where constraint is strongest.

4 Discussion

This entire analysis goes to show that the drug pocket and the PD-1 interface are one surface, and that surface has been held fixed across 429 million years of vertebrate evolution. A tumour cannot look for an escape mutation among the amino acids PD-1 leaves free, because the pocket contains none: every drug-contacting residue is an interface residue. A change that removes the compound is therefore a change in the surface PD-1 binds, so the tumour would be giving up the very function it expresses PD-L1 to perform. And every one of those amino acids has been held fixed across that entire span, which is the clearest evidence available that changing them carries a cost.

The overlap rests on two crystal structures and nothing else. It does not depend on the alignment, the tree, the substitution model or the choice of metric, and it holds at every distance cutoff I tested, each one at least six orders of magnitude below conventional thresholds (Table 2). The pocket is one physical surface rather than two adjacent patches (Figure 3).

The evolutionary measurement agrees with the structural one. *CD274* is strongly constrained at the gene level, no pocket residue shows positive selection, and as a group the pocket evolves 3.8-fold more slowly than residues elsewhere in the protein matched on accessibility. That comparison survives both extremes of the matching band, every single-residue deletion and the alternative phylogeny, so it does not rest on any one analytical choice.

This matters for the compounds now in patients. The pocket INCB099280 and ABSK043 are built to occupy the pocket PD-1 uses, so the escape route available to a tumour is not a free choice among surface residues: it runs through the receptor interaction the tumour expresses PD-L1 to exploit in the first

place.

What any particular substitution costs each interaction is a separate measurement, and Q66 is where I would start.

4.1 Limitations

No patient has yet been sequenced after progressing on a dimerization-inducing PD-L1 small molecule, so the direct test of this argument is not yet available. And conservation across species measures what a change costs an organism over evolutionary time, rather than what it costs a tumour cell under drug pressure.

5 Reproducibility

Every figure, table and number in this manuscript regenerates from saved inputs with one command, and the inputs are recorded with checksums. Every random procedure is seeded: permutations at 20260605, the co-variation background sample at 7, tree search at 20260101. No measured or computed value is transcribed into the text; each reaches it through a machine-generated macro file, and the build asserts both that a re-run of the analysis reproduces the committed results value for value and that no hand-typed number has crept into the manuscript.

Code and data:

<https://github.com/ChristianMangwanda/pd11-evolutionary-analysis>.

References

- Melissa F. Adasme, Katja L. Linnemann, Sarah N. Bolz, Florian Kaiser, Sebastian Salentin, V. Joachim Haupt, and Michael Schroeder. PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Research*, 49(W1):W530–W534, 2021. doi: 10.1093/nar/gkab294.
- Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1):289–300, 1995. doi: 10.1111/j.2517-6161.1995.tb02031.x.
- Richard E. Broughton, Ricardo Betancur-R, Chen-hong Li, Gloria Arratia, and Guillermo Orti. Multi-locus phylogenetic analysis reveals the pattern and tempo of bony fish evolution. *PLoS Currents*, 5: ecurrents.tol.2ca8041495ffafd0c92756e75247483e, 2013. doi: 10.1371/currents.tol.2ca8041495ffafd0c92756e75247483e.
- Ethan Cerami, Jianjiong Gao, Ugur Dogrusoz, Benjamin E. Gross, Selcuk Onur Sumer, Bulent Arman Aksoy, Anders Jacobsen, Caitlin J. Byrne, Michael L. Heuer, Erik Larsson, Yevgeniy Antipin, Boris Reva, Arthur P. Goldberg, Chris Sander, and Nikolaus Schultz. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery*, 2(5):401–404, 2012. doi: 10.1158/2159-8290.CD-12-0095.
- Daniel S. Chen and Ira Mellman. Elements of cancer immunity and the cancer-immune set point. *Nature*, 541(7637):321–330, 2017. doi: 10.1038/nature21349.
- The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, 51(D1):D523–D531, 2023. doi: 10.1093/nar/gkac1052.
- Eric A. Franzosa and Yu Xia. Structural determinants of protein evolution are context-sensitive at the residue level. *Molecular Biology and Evolution*, 26(10):2387–2395, 2009. doi: 10.1093/molbev/msp146.
- Katarzyna Guzik, Marcin Tomala, Damian Muszak, Marcin Konieczny, Aleksandra Hec, Urszula Blaszkiewicz, Marcin Pustula, Roberto Butera, Alexander Domling, and Tad A. Holak. Development of the inhibitors that target the PD-1/PD-L1 interaction: a brief look at progress on small molecules, peptides and macrocycles. *Molecules*, 24(11):2071, 2019. doi: 10.3390/molecules24112071.
- Incyte Corporation. A phase 2 study evaluating INCB099280 in participants with select solid tumors who are immune checkpoint inhibitor naive. ClinicalTrials.gov identifier NCT05879822, 2023. Phase 2, active not recruiting; accessed 2026.
- Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12):2577–2637, 1983. doi: 10.1002/bip.360221211.
- Subha Kalyaanamoorthy, Bui Quang Minh, Thomas K. F. Wong, Arndt von Haeseler, and Lars S. Jermiin. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6): 587–589, 2017. doi: 10.1038/nmeth.4285.
- Konrad J. Karczewski, Laurent C. Francioli, Grace Tiao, Beryl B. Cummings, Jessica Alfoldi, Qingbo Wang, Ryan L. Collins, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581(7809):434–443, 2020. doi: 10.1038/s41586-020-2308-7.
- Kazutaka Katoh and Daron M. Standley. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4):772–780, 2013. doi: 10.1093/molbev/mst010.
- Holly K. Koblish, Liangxing Wu, Liang-Chuan S. Wang, Phillip C. C. Liu, Richard Wynn, Jonathan Rios-Doria, et al. Characterization of INCB086550: a potent and novel small-molecule PD-L1 inhibitor.

- Cancer Discovery*, 12(6):1482–1499, 2022. doi: 10.1158/2159-8290.CD-21-1156.
- Sergei L. Kosakovsky Pond and Simon D. W. Frost. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, 22(5):1208–1222, 2005. doi: 10.1093/molbev/msi105.
- Sergei L. Kosakovsky Pond and Spencer V. Muse. Site-to-site variation of synonymous substitution rates. *Molecular Biology and Evolution*, 22(12):2375–2385, 2005. doi: 10.1093/molbev/msi232.
- Sudhir Kumar, Michael Suleski, Jack M. Craig, Adrienne E. Kasprowitz, Maxwell Sanderford, Michael Li, Glen Stecher, and S. Blair Hedges. TimeTree 5: an expanded resource for species divergence times. *Molecular Biology and Evolution*, 39(8):msac174, 2022. doi: 10.1093/molbev/msac174.
- Masatoshi Nei and Takashi Gojobori. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3(5):418–426, 1986. doi: 10.1093/oxfordjournals.molbev.a040410.
- Lam-Tung Nguyen, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1):268–274, 2015. doi: 10.1093/molbev/msu300.
- Duncan C. Ramsey, Michael P. Scherrer, Tong Zhou, and Claus O. Wilke. The relationship between relative solvent accessibility and evolutionary rate in protein evolution. *Genetics*, 188(2):479–488, 2011. doi: 10.1534/genetics.111.128025.
- Matthew Z. Tien, Austin G. Meyer, Dariya K. Sydykova, Stephanie J. Spielman, and Claus O. Wilke. Maximum allowed solvent accessibilities of residues in proteins. *PLOS ONE*, 8(11):e80635, 2013. doi: 10.1371/journal.pone.0080635.
- University Health Network. A phase II trial of ABSK043, an oral PD-L1 inhibitor, in patients with angiogenic sarcomas. ClinicalTrials.gov identifier NCT07014137, 2025. Phase 2, recruiting; accessed 2026.
- Sadie R. Wisotsky, Sergei L. Kosakovsky Pond, Stephen D. Shank, and Spencer V. Muse. Synonymous site-to-site substitution rate variation dramatically inflates false positive rates of selection analyses: ignore at your own peril. *Molecular Biology and Evolution*, 37(8):2430–2439, 2020. doi: 10.1093/molbev/msaa037.
- Krzysztof M. Zak, Radosław Kitel, Sara Przetoćka, Przemysław Golik, Katarzyna Guzik, Bogdan Musielak, Alexander Domling, Grzegorz Dubin, and Tad A. Holak. Structure of the complex of human programmed death 1, PD-1, and its ligand PD-L1. *Structure*, 23(12):2341–2348, 2015. doi: 10.1016/j.str.2015.09.010.
- Krzysztof M. Zak, Przemysław Grudnik, Katarzyna Guzik, Bartosz J. Zieba, Bogdan Musielak, Alexander Domling, Grzegorz Dubin, and Tad A. Holak. Structural basis for small molecule targeting of the programmed death ligand 1 (PD-L1). *Oncotarget*, 7(21):30323–30335, 2016. doi: 10.18632/oncotarget.8730.
- Weiping Zou, Jedd D. Wolchok, and Lieping Chen. PD-L1 (B7-H1) and PD-1 pathway blockade for cancer therapy: mechanisms, response biomarkers, and combinations. *Science Translational Medicine*, 8(328):328rv4, 2016. doi: 10.1126/scitranslmed.aad7118.