TYK2 — V4 Target Opportunity Report

Whitespace, trajectory and IP-portfolio strength, in one decision-first view

Target: TYK2 — non-receptor tyrosine-protein kinase 2 (**HGNC 12440**, gene Q14914487; UniProt **P29597**) **Date:** 2026-08-01 · **Substrate:** 60 term-sliced discovery calls across US/EP/WO/CN → 383 distinct rows → 201 banded ingest targets → 188 completed (**13 hard failures**) → **279 families** → **228 on-pathway** after a per-family claim-scope read (**51 reclassified off-pathway**) **What this is for:** to let an inventor, R&D strategist or BD&L reader *act* — see where the defensible room is, where the field is actually going, and which live claims stand in the way — without first reading how the analysis was built.

> **Public edition.** This is a trimmed version of an internal analyst deliverable. The uncalibrated claim-architecture (KC) scoring layer, the knowledge graph, the per-family scored data and the internal audit records are **withheld** — not because they are unfavourable but because an uncalibrated per-patent score reads as precision that does not exist, and because the machine-readable estate map is the method rather than the finding. What is published here is the **structural layer**: where claim coverage is present, absent or thin, and where filings are moving. Every guardrail from the internal edition is reproduced below verbatim.
>
> **Read-me (the guardrails, once).** This is **candidate generation and a whitespace/trajectory signal — NOT a freedom-to-operate opinion.** "Whitespace" means the *absence of strong structural claim coverage in this corpus*, which is **necessary but not sufficient** for opportunity. Genus recall and genus-overlap fractions are **lower bounds**; an INDETERMINATE verdict counts as a candidate, never as clearance. The internal edition's claim-architecture scoring layer is **uncalibrated** (no Gold Set) and is **withheld here**; it never graded whitespace in the internal edition either, and no judgement below depends on it. **A failed genus compile is UNMEASURED — not narrow and not broad. A lapsed patent is not public-domain**, "lapsed" is not "expired", and no revival window was computed for any family. **Absence-first: `absent_checked` ≠ 0, `deferred` ≠ 0, and a NULL headline is not a zero.** The family denominator is a **lower bound** — 5 of 46 pan-JAK slices saturated at the retrieval cap — so **no comparative landscape-size claim is made**. Verify any specific family with counsel before relying on it. Full method, reclassifications and limits: **Part III**.

The headline, stated before anything else

This run found one candidate whitespace cell on the mechanism × direction grid, and it is not shown to be open. Everything else that looked like whitespace was an instrument artefact, an unexamined cell, occupied art, or a different protein.

That is a thinner claim than "we found three opportunities", which is what an earlier revision of the whitespace document said. It is thinner because an adversarial review forced it to be, and the corrections all ran in the same direction: **against** the opportunity.

	first pass	after Stage F
cells published as reportable whitespace	**3** (undifferentiated)	**1** genuine coherent zero + 2 instrument artefacts
coherent cells in the grid	23	23 (unchanged)
cells the grid raw-labels `OPEN`	12	12, of which **2 are on the TYK2 protein**
thin/empty candidates	18 (file row count)	**17** (derived from the grid)
candidates actually cross-examined	"every one"	**14 of 17**; 3 never examined
full-loss events	18 over 5 families	**4 distinct losses** over 4 families
breadth slope	"robust"	**not separable** from a coverage collapse (74.1% → 26.7%)

Read the direction of the corrections. Two of the three headline cells turned out to be `ARTEFACT_ZERO` — their entire mechanism row is unpopulated across all 279 corpus families, so the zero measures our binner, not the landscape. The `OPEN` column shrank from 12 usable cells to 2 once cells on downstream STATs, upstream receptors and solid-state process art were excluded. The loss-event count fell by a factor of 4.5× once off-pathway families, duplicate rows and territorial EP cease codes were removed. Nothing was corrected in the direction of more opportunity.

What it does not mean. It does not mean TYK2 is closed: 99 of 119 compiled genera have no structural adjacency at all, and only 112 of 228 families have any recall figure, so the fence is **unmeasured for most of the corpus**. It does not mean the 60-cell grid is the only grid. And it emphatically does not mean anyone is free to operate anywhere.

PART I — ACT ON THIS

1. Actionable summary — what a reader should do

There is one candidate cell, and four actions. The actions are ranked by how much of a decision they can carry.

Action 1 — treat `JH2_pseudokinase|activator_agonist` as the only mechanism candidate, and fund the measurement before the chemistry. It is the one coherent zero in a row the corpus does populate. A TYK2 *activator* at the JH2 pseudokinase pocket is the mirror image of the entire field. The supporting evidence is genuinely thin: of 519 pan-JAK fence titles, **0 name TYK2**, and the single vocabulary match was on the word *"activator"* — **zero titles contain "agonist"**. That is weak evidence of absence, which is not evidence of

opportunity. The cheap next step is not a screening campaign; it is ingesting the 765 fence rows so their genera compile.

Action 2 — do not build on the two `ARTEFACT_ZERO` cells. `JH2_JH1_interface|allosteric_inhibitor`, `FERM_SH2_receptor_box|allosteric_inhibitor` both have `row_unpopulated_corpus_wide: true`. They tell you the binner has no vocabulary for `JH2_JH1_interface` or `FERM_SH2_receptor_box`; they tell you **nothing** about whether the art exists. Fix the ontology, re-bin, then re-ask. Acting on them now would be substituting an instrument limit for a market gap.

Action 3 — assume the two crowded lanes are genuinely crowded. `JH1_catalytic|orthosteric_inhibitor` (n=92, 26 live-granted) and `JH2_pseudokinase|allosteric_inhibitor` (n=89, 24 live-granted) hold 79.7% of all 227 coherently-binned families. 25 distinct families hold a genus at or above the 0.5 breadth threshold **and are live and granted** — that is the practical fence, and because recall is a lower bound it understates them. Freedom to operate in either lane is a question for counsel, not for this report.

Action 4 — re-run the mask before trusting any empty cell. The coherence mask is authored, and it is **already known to be wrong in one place**: family 80628904 (UNIV OXFORD INNOVATION LTD) is binned into a cell the mask calls incoherent. The error direction that *manufactures* whitespace — wrongly calling a pairing coherent — has only been bounded, not audited: 10 of 12 `OPEN` cells are not on the TYK2 protein at all.

2. The whitespace map

cell (mechanism node | direction)	n	verdict	live granted	legal unreadable	broad+ live genera	median recall
`JH1_catalytic	orthosteric_inhibitor`	92	COVERED	26	30	10
`JH2_pseudokinase	allosteric_inhibitor`	89	COVERED	24	31	13
`solid_state_process	allosteric_inhibitor`	30	COVERED	3	10	1
`TYK2_biomarker_dx	diagnostic`	6	FENCE_UNMEASURED	0	5	UNMEASURED
`solid_state_process	orthosteric_inhibitor`	6	COVERED	2	3	1
`TYK2_protein_level	degrader`	3	PARTIAL	0	0	0
`TYK2_transcript	expression_knockdown`	1	THIN	0	0	UNMEASURED

`JH2_pseudokinase|activator_agonist`, `JH2_JH1_interface|allosteric_inhibitor`, `FERM_SH2_receptor_box|allosteric_inhibitor`. Those are what survived cross-examination of 14 probed zero cells: 4 proved externally occupied, 4 were never searched, and 3 sit downstream of a thin cell — so 3 of 14 is the reportable remainder, not a count of empty space. Their grid verdicts differ, and the difference decides what you may do with them: `JH2_pseudokinase|activator_agonist` (OPEN); `JH2_JH1_interface|allosteric_inhibitor` (ARTEFACT_ZERO); `FERM_SH2_receptor_box|allosteric_inhibitor` (ARTEFACT_ZERO).

Of 60 cells, 23 are coherent (38.3%); the other 37 pair a mechanism node with a direction that cannot be filed against it, so **their emptiness measures the ontology, not the landscape**. 7 coherent cells carry art; 16 are empty. Attrition on apparent whitespace was 78.6% of the probed set.

3. Trajectory — where this target is going

era	families	share	median genus recall	genera compiled
pre-2013	16	7.0%	0.288	12
2013-2018	27	11.8%	0.643	20
2019-2022	95	41.7%	0.413	56
2023-2026	90	39.5%	0.279	24
total	**228**			

Two findings, one of them heavily qualified.

The direction mix inverted. Within TYK2-labelled art, allosteric filings now outnumber orthosteric **57 to 31** in the newest era, having started the other way round. That is the deucravacitinib effect: JH2-binding selectivity became the paradigm. **This is scoped to art that names TYK2** — the pan-JAK fence excluded 765 rows dominated by ATP-competitive chemistry, and their era distribution is **UNMEASURED**, so the inversion cannot be tested against them.

Measured claim breadth fell, but the measurement fell faster. Median genus recall peaks in 2013-2018 at 0.643 and ends at 0.279 — 43.4% of peak. But genus-compilation **coverage** collapses over the same eras, 74.1% → 26.7%, so the newest era's breadth rests on a 26.7% sample. The narrowing is **consistent with** a maturing field and **not separable** from declining coverage on this data. Do not use this slope as a forecast.

4. The claim architecture — who holds what

The claim-architecture scoring layer is withheld from this public edition; see the note at the top. What is published is the structural and legal state.

Legal state: 149 of 228 on-pathway families have readable status and **55 are live-granted**; the other 79 are **unreadable, not "not live"**. Loss events resolve to **4 distinct losses** across 4 families — see §II-D for why the raw figure was 18.

PART II — THE EVIDENCE

II-A. Scope coordinates

Rows are 10 mechanism_node values, columns 6 direction values. A 23-pair coherence mask across 10 nodes declares which pairings are meaningful. Thresholds fixed in advance: thin at n ≤ 2, broad at recall ≥ 0.5.

Verdict vocabulary, derived from the grid: ARTEFACT_ZERO, COVERED, FENCE_UNMEASURED, OPEN, PARTIAL, THIN. Of these, 1 is **absent from the published `cell_verdict_rule`** (FENCE_UNMEASURED) — the rule as applied lives in the analysis script, which is a documentation gap.

II-D. Legal layer — and why the raw loss count was unusable

18 raw -> 12 on-pathway -> 10 de-duplicated -> 8 after excluding country-prefixed EP cease codes -> 4 distinct losses. 6 of the raw events belonged to an off-pathway family, 2 were duplicate rows, and 2 were country-prefixed EP cease codes — territorial lapses that the original filter missed because it matched only the literal string `LAPSED IN A CONTRACTING STATE`. Separately 642 territorial EP lapses were excluded by that filter.

A lapsed patent is not public-domain. No revival window was computed.

II-E. The pan-JAK fence

a **lower bound of 765 pan-JAK discovery rows**. Read that as rows, not families: 519 of the 765 carry a distinct title, and the ids are a **mixed namespace** — 721 numeric INPADOC family ids plus 44 PCT application dockets, which are not families at all. The true family count is therefore lower than 765 and is **UNMEASURED**, so no rows-to-families ratio against this 279-family corpus is like-for-like. Titles are not families either, so **no ratio here is like-for-like** and none should be quoted as the fence's size relative to the corpus. It is also a lower bound in the other direction: 5 of the 46 slices across both passes are recorded saturated at the 100-row cap, the first pass (16 slices) recorded no saturation field at all, and 5 legs in pass 2 plus 8 in pass 1 returned an upstream error and were skipped — so the fence is larger than 765 by an unknown margin.

The fence layer is NOT merged into the on-pathway set, and no fence row was ingested for this analysis. It is **not** true, as earlier revisions of this text said, that no fence family appears in any KC or genus number: **7 of the fence rows are already corpus families** (67479598, 78725685, 80247730, 82260225, 85602036, 90573424, 93794614) and **3 of those sit inside the KC basis** (67479598, 80247730, 90573424). They are there because TYK2-anchored discovery found them first — they are TYK2 families that also answer a pan-JAK query — not because the fence layer imported them.

PART III — METHOD, RECLASSIFICATIONS & LIMITS

III-A. How the substrate was built

Discovery ran 60 slices; 17 returned zero rows. Jurisdiction slicing was found **not to partition** the BigQuery index leg — EP/WO/CN slices returned byte-identical row sets — so term slicing carried the de-saturation (pan-JAK: 195 rows jurisdiction-sliced against 708 term-sliced). Ingest was banded; 188 of 201 reached `done`.

III-B. The load-bearing reclassifications

51 of 279 families were reclassified off-pathway; 2 were reversed by audit (72428015, 81601788), including the family that had ranked first on the withheld claim-architecture ordering, which was a TTK inhibitor with **zero** TYK2 recitation. The on-pathway 228 is therefore an **upper bound**.

III-C. Adversarial audits

The KC report went through **seven** reviews (ACCEPT-WITH-REVISIONS → REJECT ×4 → ACCEPT-WITH-REVISIONS → releasable); the whitespace document through **one**, which rejected it with 8 blocking findings. Verifier coverage now stands at **349 independent checks** over 914 computed figures. Audit records ship with the bundle.

The single most useful thing those reviews established: **a check that recomputes from primary data is not independent when the corruption is in that primary data.** Five wrong-but-plausible data mutations passed a 237-check verifier, including one that silently changed *which* cells were called whitespace. Closing them required cross-source invariants, internal consistency properties and explicitly-labelled plausibility bounds.

III-D. Limits — read before relying

13 hard ingest failures (5 `gp_404_family_resolution`, 8 `bigquery_snapshot_lag_no_rows`), of which **8 are individually named in the ledger** and 5 are counted but not itemised there — `US-12655108-B2`, `US-RE47929-B1`, `WO-2026096291-A1`, `WO-2026147992-A1`, `WO-2026156128-A1`

- **KC is uncalibrated** and covers 94 of 279 families — essentially the post-2020 US subset. It is not a valuation and never grades whitespace.
- **228 on-pathway is an upper bound** (35 rest on inference, 166 are low-confidence).
- **The reverse mask audit has not been run** systematically; the mask has one known defect and every cell verdict inherits unquantified mask risk.
- **The fence is title-screened only.** The one candidate cell is **NOT CONTRADICTED**, never "open".
- **Breadth is confounded** by a genus-compilation coverage collapse from 74.1% to 26.7%. The collapse itself is measured; what is **UNMEASURED** is how much of the breadth decline it accounts for, so the two are **not separable** on this data.

- **397 numeric tokens remain hand-authored** across all templates and fragments, and the stale-number detector is **structurally blind below ~10** because every small integer coincides with some computed figure. Small load-bearing numbers are defended by explicit binding checks, not by the detector.

III-E. What the next run should do

1. **Ingest the 765 pan-JAK fence rows** so their genera compile. This is the single highest-value action and it gates the only mechanism candidate.
2. **Run the reverse mask audit** — look for wrongly-*coherent* pairings, the direction that manufactures whitespace.
3. **Fix the ontology for the two ARTEFACT_ZERO rows** and re-bin, so `JH2_JH1_interface` and `FERM_SH2_receptor_box` can be asked about at all.
4. **Probe the 3 unexamined candidates** (`FERM_SH2_receptor_box|activator_agonist`, `JH2_JH1_interface|activator_agonist`, `TYK2_transcript|expression_knockdown`).
5. **Trigger agent-card builds** for the 43 families whose KC absence is a build/timing cause — the repairable subset of the 152 families without KC. The other 109 are ex-US biblio-only rows that **no card build can fix**; they need ex-US claim translation.
6. **Build a Gold Set** (200–300 families) so KC can be calibrated and cross-target comparison becomes legitimate.

What is published, and what is withheld

layer	published here
whitespace grid + zero cross-examination	**yes**
filing trajectory + its confounds	**yes**
legal state (live grants, loss events)	**yes**
claim-architecture scores + per-patent ranking	withheld — uncalibrated
knowledge graph + ontology	withheld — method substrate
per-family scored JSON + scripts	withheld — method substrate
internal audit records + ingest ledger	withheld — internal QA

The companion structural document, `TYK2_V4_Whitespace_and_Trajectory_PUBLIC.md`, carries the grid and trajectory in full.