anti-Tau / MAPT — V4 Target Opportunity Report

Whitespace & Trajectory — public edition

Target: tau / **MAPT** — microtubule-associated protein tau (**HGNC 6893**, gene `Q14865307`; UniProt **P10636**, protein `Q419507`) **Date:** 2026-07-31 · **Substrate:** 48 term-sliced discovery calls across US/EP/WO/CN → 1,736 distinct rows → 1,513 gate-admitted → 934 with an ingest entrypoint (**579 without**) → 903 ingests completed → **783 families → 267 with a machine-read claim set → 223 on-pathway** after a per-family claim-scope read (**44 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 presents the **whitespace and trajectory** reads that accompany the article. The underlying claim-strength scoring, per-family score tables, knowledge graph and raw inputs are part of the internal analyst deliverable and are **not reproduced here**; that strength layer is relative-ordering, calibration-pending work and is deliberately not published as a ranking.*
>
> **Read-me (the guardrails, once):** This is **candidate generation and a whitespace/trajectory signal — NOT a freedom-to-operate (FTO) opinion.** "Whitespace" here 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. **Sequence identity is NOT a lower bound and is not scope-bearing on this target.** The upstream alignment-length defect was fixed on 2026-08-23 and the edges were re-harvested against the floored service (50-residue minimum, `align_len` now persisted on every edge). Flooring did not rescue the layer, it clarified it: of the 138 distinct hit publications, **135 are not tau-related**, only **30 of 2,391 edges (1.3%)** reach a tau-related publication, and **no edge links two on-pathway tau families**. The strongest "neighbours" are humanised anti-TL1A, anti-TNF-α and anti-SIRP-α antibodies — two humanised antibodies share framework regardless of what they bind, so this measures **shared scaffold, not tau scope**. No sequence-identity figure in this report is read as an overlap measurement, and none is an input to any cell verdict (**II-C**). The claim-strength layer that supports this analysis internally is **uncalibrated** — no Gold Set exists — so it is used only for *relative* ordering, **never as a valuation and never to grade whitespace**, and it is **not published here as a ranking**. Two of its six axes are gated by how completely a specification could be machine-read, which makes them **lower bounds** and makes any cross-modality comparison on them inadmissible; they are excluded from every judgement in this document. **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 here**. **Absence-first: `absent_checked` ≠ 0, `deferred` ≠ 0, and a NULL headline is not a zero. The family denominator for this target is UNMEASURED** — 14 of 48 discovery slices saturated at the retrieval cap — so every count is a **lower bound** and **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 no reportable whitespace on the mechanism × direction grid. Zero cells of twenty-four. That is the finding, it survived three adversarial reviews, and it was strengthened rather than softened by every correction those reviews forced.

That sentence is worth more than a manufactured opportunity, so it is worth being precise about what it does and does not mean.

	first pass	after Stage F
cells reportable as whitespace	**0** (asserted)	**0** (measured)
coherent cells in the grid	17	**24** — 2 struck, 9 added
cells the grid raw-labels `OPEN`	3	**7**
cells that survive a dedicated zero-audit as whitespace	not audited	**0 of 7**
live-granted families	73	**109**
"unreachable" discovery rows offered as evidence of open room	579	**223** (61% were already in the corpus)
cells cross-examined one by one	9	**every coherent cell with no broad live claim (18)**

Read the direction of the corrections. The coherence mask was widened *permissively*, from 17 cells to 24 — nine more cells were made **eligible** to be called whitespace — and none of them was. The legal layer was rebuilt and *raised* live grants from 73 to 109. The two PARTIAL cells that the first pass had never cross-examined were examined, and one **flipped to COVERED**. The seven raw-`OPEN` cells were each interrogated and **all seven resolved to something other than whitespace**: eleven of the eighteen audited cells are **UNMEASURED** (an instrument limit, not open room), three are **occupied through a secondary claim direction**, three are **naming-convention artefacts**, and one is **measured, occupied, and simply not broadly fenced**.

Every one of those corrections ran *against* the thesis, and the null still stands. That is the strongest available evidence that this null is a property of the landscape rather than of the instrument.

What it does not mean. It does not mean anti-tau is closed — 172 of 223 families have **no compiled genus at all**, so the structural fence is unmeasured for 77% of the corpus. It does not mean the 24-cell grid is the only grid: the finer **mechanism × modality** grid has **170 cells, of which 138 are UNMEASURED and un-audited**, and that is where a next run should look. And it emphatically does not mean anyone is free to operate anywhere.

PART I — ACT ON THIS

1. Actionable summary — what a reader should actually do

There are no ranked build-here opportunities to offer from the direction grid, because there are none in the data. What the analysis *does* support are four concrete actions, ranked by how much of the decision they can carry.

Action 1 — Buy measurement, in eleven named places, before buying freedom. Eleven of the twenty-four coherent cells are **UNMEASURED**, and they are unmeasured for three distinguishable and individually addressable reasons. This is the highest-value action in the report because it converts an instrument limit into an answerable question, and because the eleven cells are named rather than gestured at.

the reason a cell is UNMEASURED	cells	the fix
no discovery slice ever targeted this cell's claim type — the retrieval never asked the question	`MAPT_mRNA|detect` · `MAPT_splicing|detect` · `MAPT_transcription|detect` · `tau_translation|detect` · `tau_aggregation|redirect_modulate`	five new discovery slices; cheapest fix in the report
ingested families matching the cell exist but their claims could not be read (KC absent)	`tau_clearance|detect` · `MAPT_transcription|lower_inhibit`	ex-US claim translation, or re-ingest for the repairable subset
direction-appropriate estates exist that were never read at all	`tau_microtubule|detect` · `tau_translation|lower_inhibit` · `tau_PTM|lower_inhibit` · `tau_microtubule|enhance_stabilise`	resolve the 223 genuinely-unread no-entrypoint rows (§III-A)

Action 2 — Treat nine named families as the fences to clear, not the field to avoid. These are the only broad, live-granted, *measured* genus claims in the corpus (§3). Nine of 223 is a small number, and the reason it is small is mostly that the genus layer is missing rather than that the fences are few — but the nine that exist are the ones a chemistry programme has to design around, and each is named with its recall, its precision, its node and its owner.

Action 3 — Build on the trajectory, not on the grid. The single sharpest structural signal in this analysis is a **crossover inside the current era**: intracellular MAPT-lowering filings overtook extracellular antibody filings for the first time in 2022–2025 (**21 against 15**, a ratio moving 0.55 → 1.40), and the `MAPT_mRNA` node's era vector is **0 → 5 → 9 → 19**. The internal claim-architecture read independently ranks `MAPT_mRNA` the strongest of the ten mechanism nodes, on 33 families. Two different instruments — filing dates and claim structure — point at the same node. **The lane is crowded and getting more so, which is a reason to move early or differently within it, not a reason to call it open.**

Action 4 — Do not read the `tau_measurement` lane as a backwater. It was banded last for ingest, it lost 141 families to absence, and the first draft of the KC analysis wrongly put it at the bottom of the modality table on an n of 9. On the authoritative modality axis it is a **27-family lane sitting mid-table** on the internal strength read rather than at the bottom, and `tau_measurement|detect` is a **COVERED** cell with 37 families and 20 live grants. It is a real, occupied, reasonably well-drafted lane.

And one non-action, stated because the temptation is obvious. The seven cells the grid raw-labels `OPEN` are **not** a build list. All seven were individually resolved to UNMEASURED, occupied-via-secondary, or artefact. **Publishing them as opportunities would be the single most misleading thing this report could do**, and the audit that dismantled them is in §2.3.

---

2. Whitespace map — the coherent grid

2.1 The coordinate system, and the mask that came before any verdict

The grid is **10 mechanism nodes × 4 claim directions = 40 cells**. A verdict was assigned only after deciding which cells are *chemically and biologically coherent* — a cell like "reduce the amount of the microtubule that tau binds" has no referent, and calling it "open" would be an artefact of the axes rather than a finding about the landscape.

24 of 40 cells are coherent; 16 are not. The mask itself was rebuilt under adversarial review, and rebuilt in the permissive direction: **2 cells struck** (`MAPT_transcription|redirect_modulate`, `tau_microtubule|redirect_modulate` — both referent-free), **9 cells added**, and 14 exclusions confirmed. The nine added cells are the honest ones to name, because each was a cell the first pass had **excluded from being able to be whitespace**: `MAPT_transcription|detect`, `MAPT_mRNA|redirect_modulate`, `MAPT_mRNA|detect`, `MAPT_splicing|detect`, `tau_translation|detect`, `tau_aggregation|redirect_modulate`, `tau_aggregation|enhance_stabilise`, `tau_clearance|detect`, `tau_microtubule|detect`.

Nine cells were made eligible to be whitespace and none of them was whitespace. The mask revision is recorded in `data/ontology.json` with its reasoning and a self-check assertion, so the widening is auditable rather than asserted.

2.2 The grid, with verdicts

Verdict rules, applied mechanically: **COVERED** = a broad live-granted claim **or** ≥ 4 live grants · **PARTIAL** = coverage present, no broad live-granted claim, < 4 live grants · **THIN** = ≤ 2 families and no broad live-granted claim · **OPEN** = 0 on-pathway families claim the cell.

cell	verdict	families	live grants	broad live	max recall (live, measured)
`extracellular_tau|lower_inhibit`	**COVERED**	51	25	1	0.837
`tau_aggregation|lower_inhibit`	**COVERED**	42	23	2	0.630
`tau_measurement|detect`	**COVERED**	37	20	2	0.747
`MAPT_mRNA|lower_inhibit`	**COVERED**	33	10	1	0.739
`tau_PTM|lower_inhibit`	**COVERED**	26	10	0	—
`tau_clearance|lower_inhibit`	**COVERED**	18	10	2	0.920
`tau_PTM|redirect_modulate`	**COVERED**	4	4	1	0.611
`MAPT_splicing|redirect_modulate`	PARTIAL	5	2	0	—
`MAPT_mRNA|redirect_modulate`	OCCUPIED VIA SECONDARY	0	0	0	—
`MAPT_splicing|lower_inhibit`	OCCUPIED VIA SECONDARY	0	0	0	—
`tau_aggregation|enhance_stabilise`	OCCUPIED VIA SECONDARY	0	0	0	—
`tau_microtubule|enhance_stabilise`	THIN	2	2	0	—
`MAPT_transcription|lower_inhibit`	THIN	1	0	0	—
`tau_translation|lower_inhibit`	THIN	1	1	0	—
`tau_PTM|detect`	THIN	1	0	0	—
`tau_aggregation|detect`	THIN	1	1	0	—
`extracellular_tau|detect`	THIN	1	1	0	—
`MAPT_transcription|detect` · `MAPT_mRNA|detect` · `MAPT_splicing|detect` ·	*raw* OPEN	0	0	0	—

cell	verdict	families	live grants	broad live	max recall (live, measured)
`tau_translation|detect` · `tau_aggregation|redirect_modulate` · `tau_clearance|detect` · `tau_microtubule|detect`					

One verdict changed under review and it changed in the direction that closes room. `tau_PTM|redirect_modulate` was PARTIAL in the first pass; cross-examined, **all four of its families are live-granted**, which trips the COVERED rule. It had never been examined because the first pass only cross-examined OPEN and THIN cells — a selection that could only ever *find* whitespace, never lose it.

A structural asymmetry worth stating plainly: because genus recall is a **lower bound**, a cell can only ever be **under-called as thinner than it really is**. Every error this grid can make points toward *more* apparent whitespace, not less. Getting zero out of a biased instrument is therefore more informative than getting zero out of a neutral one. **Sequence identity is deliberately excluded from this claim.** It is not an input to any cell verdict — verdicts are computed from family membership, legal status and the presence of a broad live-granted structural claim — and, as **II-C** records, its edges are not dependable in the direction the phrase "lower bound" would imply.

2.3 The zero-audit — why none of the seven `OPEN` cells is an opportunity

Eighteen cells had no broad live claim and were each interrogated individually against the corpus, the absence ledger and the unread pool. **Zero came back as reportable whitespace.**

resolution	cells	what it means
UNMEASURED	**11**	the instrument did not look, or looked and could not read. Not open room.
OCCUPIED VIA SECONDARY DIRECTION	**3**	a family binned in a neighbouring cell has claim language that reaches into this one
CONVENTION ARTEFACT AND CONTESTED	**3**	the cell exists only because of how the direction axis names things
MEASURED, NOT WHITESPACE	**1**	genuinely occupied; simply not broadly fenced

The three occupied-via-secondary cells rest on verbatim claim evidence, not inference. Six families in the corpus claim in two directions at once, and each override is recorded with the language that justifies it — Otsuka's `92263044` (claims 3 and 7, *"reducing the amount of TAU protein"*, in a family binned to splicing), Würzburg `85685250`, Kentucky `90835463`, EPFL `62167111`, Keio `83320469`, MSD `62489273`. Four antibody families were considered for override and **explicitly not overridden**. A secondary direction **never** adds to a cell's family count — it only blocks the cell from being called OPEN, which is the conservative use of the evidence.

One reviewer finding was itself wrong, and correcting it made the null stronger. Review #3 argued that `tau_aggregation|redirect_modulate` was occupied by Acelot's `74503198`. Reading that family's own evidence, its compounds *"reduce or prevent the formation of oligomers"* — that is **lowering**, not redirection. The cell is genuinely empty, and it is empty **and UNMEASURED**, because no discovery slice ever asked for aggregation-redirecting chemistry. The correction removed an occupant and produced no whitespace.

2.4 The figure

anti-Tau / MAPT whitespace & trajectory heatmap

anti-Tau / MAPT — mechanism node × claim direction, 40 cells, of which 24 are coherent. Verdicts: ***7 COVERED****,* ***10 coverage-limited*** *(PARTIAL + THIN),* ***4 fence-unmeasured****,* ***3 occupied via a secondary claim direction****,* ***2 mask artefacts****, 14 not-applicable.* ***No cell is whitespace.*** *Colour encodes the verdict, not a score. The four "fence-unmeasured" cells are cells where the structural genus layer could not be measured at all — they are not thin cells and must not be read as open. Candidate generation, not freedom-to-operate; genus recall is a lower bound. Sequence identity is not an input to any cell verdict in this grid and is not dependable as a lower bound — see II-C/II-E.*

---

3. The actionable chemical core — the nine fences that are real

77% of this corpus has no compiled genus. That is the single most important structural fact in the report, and it has to be stated before any fence is named, because it bounds what the fence table can mean.

structural genus layer	n	of 223
families with **any** compiled genus	51	22.9%
families with **no** genus at all (`no_family_best`)	**172**	**77.1%**
compiles that **failed or did not parse → UNMEASURED**	34	66.7% of the 51
families carrying a recall value	47	—
…of which the recall comes from a **failed** compile	**30**	**63.8%**
families with a **usable** recall	**17**	7.6%
broad AND live-granted AND measured (recall ≥ 0.60, precision ≥ 0.50)	**9**	4.0%

Sixty-four percent of the recall values in this corpus come from compiles that failed. A failed compile is **UNMEASURED — not narrow, not broad** — and it is the guardrail most likely to be violated by a reader skimming a recall column, so it is repeated here.

The nine broad, live-granted, measured fences

recall / precision	family	node	modality	owner
0.920 / 0.985	`58488491`	tau_clearance	small-molecule enzyme	NY State Psychiatric Inst / Univ. Colorado
0.857 / 0.890	`81276404`	tau_clearance	degrader	Univ. CAS
0.837 / 0.935	`66247701`	extracellular_tau	vaccine / immunogen	United Neuroscience / UNS IP
0.739 / 0.985	`52346860`	MAPT_mRNA	antisense oligonucleotide	**Biogen**
0.689 / 0.990	`41478644`	tau_measurement	imaging tracer	WisTa / TauRx
0.636 / 0.975	`10771951`	tau_measurement	screening assay	Univ. Aberdeen
0.630 / 0.915	`44858713`	tau_aggregation	small molecule	Maccioni
0.618 / 0.845	`43567101`	tau_aggregation	small molecule	Alassia / Chaltin
0.611 / 1.000	`49551168`	tau_PTM	small-molecule enzyme	**Merck Sharp & Dohme**

A precision floor of 0.50 is applied, and applying it matters. Without it, thirteen families would qualify as "broad live" instead of nine. A high recall on a low-precision compile means the genus sweeps in non-members — the fence is **sloppy, not wide** — and letting those four families into a COVERED verdict would have hidden coverage-limited cells behind them. The four excluded families are named, each with its actual reason, because two different reasons are in play and conflating them is exactly the error the guardrail exists to prevent: `45994015` (0.884 / 0.900) is excluded because its compile is **`parse_failed` — UNMEASURED**, while `88695491` (0.747 / **0.36**), `38669650` (0.628 / **0.44**) and `72921263` (0.611 / **0.42**) are excluded on **precision**.

Where to build, on the structural axis, in one sentence: the two `tau_clearance` fences are the broadest in the corpus and sit on a node with only 18 families, the `MAPT_mRNA` fence is a single Biogen antisense family on the node the trajectory says is accelerating fastest, and **`tau_PTM|lower_inhibit` has 26 families and 10 live grants but not one broad measured fence** — which makes it the most interesting node in this table, and also one of the eleven UNMEASURED cells, for the same reason.

The finer grid, and the 138 cells nobody has looked at

The mechanism × **modality** grid is **170 cells**: **9 COVERED, 7 PARTIAL, 16 THIN, and 138 `UNMEASURED_UNAUDITED`**. The 138 are not whitespace and are not claimed as whitespace — **they are cells that received no dedicated audit at all**, and they are the honest answer to "where should the next run look".

The occupied lanes, for orientation: `tau_aggregation|small_molecule_aggregation` 28 families / 16 live · `extracellular_tau|antibody_extracellular` 28 / 17 · `tau_measurement|immunoassay_reagent` 25 / 14 · `tau_PTM|small_molecule_enzyme` 15 / 6 · `MAPT_mRNA|antisense_oligo` 14 / 5 · `MAPT_mRNA|sirna_rnai` 13 / 3 (PARTIAL) · `extracellular_tau|vaccine_immunogen` 12 / 7 · `tau_clearance|degrader` 8 / 4.

`MAPT_mRNA|sirna_rnai` is the lane to watch: 13 families, only 3 live-granted, verdict **PARTIAL** — the most active modality on the fastest-growing node, and mostly still pending. That is a lane in formation, not a lane with room.

---

4. Trajectory — where this target is going

Dates are **earliest member filings**. `get_family.earliest_priority_date` returned **null for all 223 families**, so priority is unavailable and every era assignment is a **late bound** on the family's true age. Span: **1995-01-19 → 2025-09-03**.

era	families
1995–2009 foundational	27
2010–2016 antibody era	59
2017–2021 pivot	66
2022–2025 genetic-medicine wave	71

4.1 The crossover — the sharpest signal in the analysis

era	intracellular MAPT-lowering	extracellular antibody node	ratio intra : extra
1995–2009	0	0	—
2010–2016	8	17	0.47
2017–2021	11	20	0.55

era	intracellular MAPT-lowering	extracellular antibody node	ratio intra : extra
2022–2025	**21**	**15**	**1.40**

The inversion happens inside the current era, not between eras. For twelve years the field filed roughly two extracellular antibody families for every intracellular-lowering one; in the most recent window that reversed. The `MAPT_mRNA` node's own vector is **0 → 5 → 9 → 19**, and `MAPT_splicing` is flat at 2 → 1 → 2. This is the structural signature of a field that spent a decade betting on clearing tau from outside the cell and has now moved to preventing its production inside it.

Two independent instruments agree. The internal claim-architecture read makes `MAPT_mRNA` the strongest of the ten mechanism nodes, on 33 families, and the three modality lanes with the deepest claim hierarchies are all nucleic-acid or vector platforms. Filing dates and claim structure are different measurements of different things, and they point the same way.

4.2 The other three trajectory signals

China-affiliated filings, by era: 0 → 1 → 2 → 14. Fourteen of the last era's 71 families. This flag is **assignee-string inference and PROVISIONAL** — no assignee's national origin was independently verified, and a co-assignment with a non-Chinese party still trips it — so it is directional only.

Actor type does not show a platform-company takeover. By era, `other_company_or_individual` runs 16 / 30 / 30 / 30 while academic institutions run 6 / 14 / 12 / **21** and established sponsors 5 / 15 / 24 / 20. The most recent era's growth is **academic**, not corporate.

Indication claiming is overwhelmingly absent, and that is a finding. 154 of 223 families (69.1%) claim no indication at all — they are platform or composition-of-matter filings. 60 claim a tauopathy umbrella, **7 name Alzheimer's**, 2 name FTD. Separately, Wikidata's disease associations for MAPT (`P2293`) return **five diseases and Alzheimer's is not among them** — a grounding gap worth knowing about when reading any tau-to-AD claim, in a patent or anywhere else.

4.3 The antibody-rescue cluster — a 37-day window

Four families, one filer, one construct pattern, thirty-seven days:

filed	family	construct
2024-08-10	`98698437`	single-chain **gosuranemab**–transferrin fusion
2024-08-26	`98830058`	single-chain **semorinemab**–transferrin fusion
2024-08-31	`98899994`	single-chain **tilavonemab**–transferrin fusion
2024-09-16	`99059480`	single-chain **zagotenemab**–transferrin fusion

Consecutive gaps of 16, 5 and 16 days. All four are **pending; none is live-granted.** All four take a clinically-failed anti-tau antibody and re-file it as a transferrin fusion — a blood-brain-barrier-delivery bet on antibodies whose failures are usually read as target-engagement failures.

Attribution discipline: the originator mapping (Biogen/BMS, Roche/Genentech, AbbVie, Lilly) and the clinical outcomes are **external facts, not corpus findings**. **No originator family for any of the four appears in this basis.** Semorinemab originated at AC Immune (in this corpus as `72921263`) and was licensed to Genentech, so "Roche/Genentech" is a **licensee, not an originator** — the kind of distinction an assignee-string rollup will get wrong if nobody checks it.

PART II — THE STRUCTURAL EVIDENCE

II-A. The scope-coordinate system

Four axes, read per family from the claims by an LLM claim-scope pass over six validated shards, then used as the whitespace layer's input: **10 mechanism nodes** (`MAPT_transcription`, `MAPT_splicing`, `MAPT_mRNA`, `tau_translation`, `tau_PTM`, `tau_aggregation`, `tau_microtubule`, `tau_clearance`, `extracellular_tau`, `tau_measurement`), **4 directions** (`lower_inhibit` 172 · `detect` 40 · `redirect_modulate` 9 · `enhance_stabilise` 2), **17 modalities** (20 families `unresolved` — a real bin, not a residual), and **indication**.

Per-family bin confidence on the 223-family basis: **high 100 / medium 98 / low 25**. A quarter of the basis is medium-or-low confidence on its own coordinates, and the grid inherits that.

One known inconsistency is left in place rather than quietly fixed. Family `43662090` is binned on-pathway although its own `scope_evidence` reads *"No tau limitation in the claim"*. It is disclosed here because a reader auditing the binning will find it, and finding it disclosed is better than finding it hidden.

II-B. Legal layer — and why "lapsed" is not "free"

family verdict	n	of 223
granted_active (live)	**109**	48.9%
pending	86	38.6%
withdrawn / abandoned	17	7.6%
lapsed	6	2.7%
expired	5	2.2%

91 families carry at least one adverse event — and 45 of them are still live-granted through a sibling member. An adverse event on one member is not the death of a family, and reading the 91 as 91 openings would be wrong by a factor of two.

The adverse-event taxonomy was calibrated, because the raw strings mislead. `LAPSED IN A CONTRACTING STATE` appears **2,648 times** and is a *territorial* EP-state event, not a family death — it is excluded from decay entirely. `OPPOSITION FILED` (34 occurrences) is a validity challenge and is tracked separately. Eleven application-loss strings (175 occurrences) were added to the taxonomy. Adverse events do not partition: 39 families carry one kind, 23 two, 24 three, 5 four.

Guardrails on this table, all load-bearing. "Lapsed" is **not** "expired" and the two are kept apart. **No revival window was computed for any family** — so nothing here says whether any of the 6 lapsed families could be revived. **The first pass reported 73 live grants** by inferring grant from the USPTO application status string `"Patented Case"`; the corrected figure from the card's `legal_status.verdict` is **109**, and the earlier number appears nowhere in this report except as a superseded value.

By era, the adverse-event share runs **81.5% → 55.9% → 39.4% → 14.1%** — which is an **exposure gradient, not a quality gradient**. A 1997 filing has had thirty years to lapse; 54 of the last era's 71 families are **still pending and cannot have lapsed at all**. Restricting to ever-granted families, survival runs **55% / 98% / 100% / 100%**, and only that first figure carries any signal, on a base of 22.

II-C. Structural transfer — sequence and genus adjacency

Sequence-identity edges exist for **46 of 223** families; genus-overlap edges for **10**. `sequences_claimed_n` rolls up to 0 across the basis, but the underlying `claimed` flag is **unpopulated (null)** on every sequence

record checked — **an absent measurement, not a measured absence** — so the claimed-sequence layer cannot support an overlap read here at all.

A second and independent reason this layer cannot be read — added 2026-08-18. The identity figures returned by the upstream aligner (`seq-local/v1`) carried **no minimum alignment-length floor**. On a 12-publication, 161-hit audit of this target's own edges, **63.4% of hits rested on alignments shorter than 50 residues**, 34 were shorter than 30, and **31 of the 37 hits reported at 100% identity aligned over fewer than 50 residues** — the shortest over **7**. A perfect match across seven residues is noise, not scope adjacency. The stored graph compounded this: `align_len` was **not written to the edge**, so a shipped edge could not be re-qualified without re-querying the source. Audit: `align_len_audit_2026-08-18.json`.

The alignment-length defect has since been fixed upstream — and the conclusion above did not change. Updated 2026-08-23. A 50-residue floor now applies at production and read time. We re-harvested this target's sequence edges against the floored service rather than filtering the old ones in place, because the old slice was **biased and not merely noisy**: it was page one of an `identity_pct DESC` ordering taken with no floor, so short perfect matches had crowded the long legitimate ones out before truncation. Re-asking returned **2,391** edges at `min_align_len` 50, against **303** legitimate ones in the old slice — the defect had been suppressing real signal, not just adding noise.

Flooring the edges made them legible, and what they legibly say is that they are not about tau. Every one of the **138** distinct hit publications was resolved and its title tested for tau / MAPT / microtubule / tangle / neurofibrillary / Alzheimer's / tauopathy. **Three** are tau-related. **135 are not.** Only **30 of the 2,391 edges (1.3%)** reach a tau-related publication, so **98.7% are off-target**, and **no edge at all** links two on-pathway tau families. The strongest "neighbours" are humanised **anti-TL1A** antibodies (`US-8263743-B2`, `US-8728482-B2`, `US-9416185-B2`), an Amgen **anti-TL1A/anti-TNF-α** bispecific, an **anti-SIRP-α**, and two Cedars-Sinai **IBD** patents. Two humanised antibodies share framework regardless of what they bind: this layer is measuring **shared humanised scaffold, not tau scope**. The tell is that the same anti-TL1A publications also top the hit list for unrelated **TTR** queries.

Accordingly **no sequence-identity figure in this report may be read as an overlap measurement**, and none is used for one. The 2026-08-18 withdrawal stands — on stronger and different grounds than the ones originally given. Audit: `tau_seq_hit_identity_2026-08-23.json`.

Why no structural-overlap figure accompanies this edition. A family-by-family sequence-overlap exhibit was attempted for this target once the upstream floor landed. It cannot be built honestly: there is **no within-corpus adjacency to plot** (zero family pairs), and the cross-corpus adjacency is scaffold artefact. Such a figure would be a picture of antibody framework conservation captioned as tau competition. The absence of the figure is the finding, and it is reported here rather than filled.

This is the layer that would ordinarily carry a combination/transfer analysis, and on this target it cannot — now for three reasons: one about what was never measured (`claimed`), one about what the measurement was worth before the fix, and one about what it is worth after. Saying so is the finding.

II-D. Knowledge graph

A typed knowledge graph underlies the analysis — mechanism nodes, claim directions, modalities, indications, patent families, assignees, filing years and publications, joined by the structural relations this report reads (`acts_on`, `directed_as`, `uses_modality`, `targets_indication`, `lapses_on`, `genus_overlaps`, `sequence_identity`). It is part of the internal deliverable and is **not published**. One property of it is worth stating because it bears on how any tau-to-Alzheimer's claim should be read: the graph's gene→disease edges come from Wikidata's disease associations for MAPT, and **Alzheimer's disease is not among them**.

Three cautions apply to the graph and to the structural adjacency layer above it. Per-family edge caps mean the graph holds a **subset** of the stored structural edges, so nothing in it is a corpus total. The claimed-sequence layer cannot be read at all here — the underlying `claimed` flag is **unpopulated** on every sequence record checked, which is an **absent measurement, not a measured absence**. And the graph's `sequence_identity` edges are **carried but not relied upon**.

The graph's sequence layer was rebuilt 2026-08-23 after the upstream alignment-length floor landed. `align_len` is now persisted on **every** edge (min 50, median 115, max 467), where previously it was written on **none** — so an edge can now be re-qualified without re-querying the source. The rebuild touched only that layer: all other relations are unchanged. The effect was not a tidy-up but a correction of direction. The contaminated graph was **artificially broad — 234 of its 278** hit publications had been surfaced by sub-50-residue matches alone, **31** genuine ones had been crowded out entirely, and **22 of the 46** sequence-bearing families turn out to have **no** edge that clears the floor. Any earlier impression of wide cross-sponsor sequence adjacency on this target was reading the defect.

The layer is nevertheless marked **`FLOORED_BUT_NOT_SCOPE_BEARING`** and `evidentiary: false`, because 98.7% of its edges are off-target scaffold matches (II-C). It is carried for structure and relied upon for nothing.

PART III — METHODOLOGY, RECLASSIFICATIONS & LIMITS

III-A. How the substrate was built

stage	result
discovery	**48 term-sliced calls** across US / EP / WO / CN → 2,552 raw rows → **1,736 distinct**
relevance gate	**1,513 admitted**; 36 rows held out across 15 assignee-collision classes
entrypoint resolution	**934 ingestable** (410 publication numbers / 524 INPADOC family ids) · **579 with no entrypoint of any kind**
ingest	933 submitted, 0 submission failures → **903 done / 30 hard-failed → 783 families**
claim-set read	**267 families** carry a machine-read claim set · **516 carry none**
claim-scope screen	**223 on-pathway · 44 reclassified off-pathway**

Assignee-collision discipline. Seven collision classes were expected; **fifteen** were found — including Tre Tau Engineering, TAU CARBON, TaskUnite, Telepathy Labs, Maplebear, Tau Motors, Alpha Tau, and the telecom sense of "Tracking Area Update". One false drop was caught and repaired (WO `67218458`).

The 579 no-entrypoint rows were the first pass's headline evidence for open room, and 61% of them were already in the corpus. A title-join against the ingested set resolves the pool: **174 already binned on-pathway · 104 ingested but KC-absent (repairable) · 66 ingested but KC-absent (structural) · 12 already screened off-pathway · 223 genuinely unread.** The first pass overstated its own unread residual by a factor of **2.6**, and it did so in the direction that made the landscape look emptier.

The 223 genuinely unread rows are blocked, not deferred by choice. They are US pre-grant/pending rows whose `pubnum` is null and whose identifier is an application number: `ingest_patent` cannot resolve them and `ingest_family` rejects them. No available entrypoint reaches them.

The absence ledger. Of the 516 KC-absent families: **314 structural** (ex-US bibliographic-only, no full-text member, `claims_total = 0`, CN-heavy — **not repairable by re-ingest**, needs ex-US claim translation) and **202 build/timing** (all have a full-text member; **repairable**). Re-ingest was submitted and accepted for all 202; **landing was not confirmed within this run**. The reviews independently established that repairing them would land roughly **281 reads on already-COVERED nodes against 21 on thin ones** — so the repair is worth doing for completeness, not because it is likely to change a verdict.

30 ingest requests hard-failed (18 Google-Patents 404 family resolution, 12 BigQuery snapshot lag), landing on five distinct 2026 lanes: MAPT siRNA formulation, MAPT-modulating oligonucleotides, tau degraders, acetylated-tau antibodies, NFT imaging. **The 2026 frontier in this report is incomplete by a named amount.**

III-B. The load-bearing reclassifications

1. **44 of 267 scored families reclassified off-pathway** on a per-family claim-scope read — including 7 WisTa-lineage methylthioninium families whose claims are chemical-synthesis process or non-tauopathy indication claims (COVID-19, hypoxemia) with no tau limitation. **The same estate therefore appears on both sides of the screen**, which is why the screen is per-family and claim-based rather than assignee-based.
2. **The coherence mask, 17 → 24 coherent cells** (2 struck, 9 added) — the widening that produced no whitespace (§2.1).
3. **The legal layer, 73 → 109 live grants**, rebuilt off `legal_status.verdict` rather than a USPTO application-status string (§II-B).
4. **The modality axis** — the KC analysis's §5.1 was rebuilt on the authoritative Stage-D claim-scope bin; the regex tag it originally led with agrees only **50.2%** of the time and produced four wrong lane findings.
5. **Six dual-direction overrides**, each with verbatim claim evidence; four candidate overrides explicitly refused (§2.3).

III-C. Adversarial audit — three reviews, two documents, 336 checks

Three impartial adversarial reviews, none by the author of the material under review. Review #1 (KC portfolio): verdict **REJECT — do not ship**; 5 BLOCKER / 16 MAJOR / 11 MINOR. Review #2 (whitespace): 3 BLOCKER / 15 MAJOR / 16 MINOR. Review #3 (axes and mask): 9 wrongly-excluded cells, 5 missing nodes, 6 missing modalities, 8 unreported measured results.

Both #1 and #2 confirmed that the arithmetic reproduced exactly. The failures were in what was *concluded* from correct arithmetic — which is the more dangerous class of error, and the reason a numeric spot-check is not a review.

The blockers, and where each is resolved: the live/lapsed field (§II-B) · the unreachable-estate evidence (§III-A) · a direction filter that could not distinguish the three therapeutic directions · the unreachable `CONTEXT_DEPENDENT` flag (§II-B) · two headlines resting on a tag the document itself declared unreliable (§III-B) · a pre-run corpus baseline built on five **saturated** probes · the Voyager superlatives. **All are resolved**; full record in `V4_Tau_whitespace/whitespace_audit_record_2026-07-31.md`.

Two defects were found inside the Stage-F fixes themselves, both by independent recomputation rather than by re-reading output: a rank-agreement statistic was computed with **ordinal ranks**, breaking ties by sort order on a corpus where values tie heavily (corrected to midranks), and a fence-exclusion list was **named after one of the two reasons it actually applied**, with its broadest row described in prose as "not live-granted" when it is live-granted and was excluded for a failed compile. Both are now asserted by the verifiers.

Verification. Two verification scripts gate the two internal analyses — **87 checks** and **258 checks** respectively. Neither reads its numbers from the prose: both recompute from the artefacts and then assert what the prose says, so a transcription error and a computation error fail separately. Several retractions are asserted **structurally** — an unreachable statistical threshold is re-enumerated from first principles, the baseline probe's saturation is re-checked, and withdrawn sentences are asserted **absent** — so they cannot be silently reinstated.

III-D. Limits — read before relying

1. **Not FTO.** Candidate generation only. Genus recall and overlap fractions are **lower bounds**; INDETERMINATE is a candidate, never clearance. **Sequence identity is not a lower bound and is not scope-bearing on this target** — the upstream alignment-length defect was fixed and the layer re-harvested on 2026-08-23, which showed that **98.7%** of its edges reach publications with no tau relationship at all (§II-C). It is not an input to any cell verdict and supports no conclusion in this report.
2. **The family denominator is UNMEASURED.** 14 of 48 slices saturated at the retrieval cap. Every count is a **lower bound**; **no comparative landscape-size claim is made**.
3. **The null whitespace result is bounded by the 11 UNMEASURED cells and the 138 unaudited modality cells.** "No reportable whitespace on this grid, at this measurement depth" is the claim. It is not "no whitespace exists".
4. **77% of the corpus has no compiled genus, and 64% of the recall values that exist come from failed compiles.** A failed compile is **UNMEASURED — not narrow**.
5. **The claim-strength layer is uncalibrated and is not published here.** Two of its six axes are gated by how completely a specification could be machine-read, making them **lower bounds** and excluding them from every judgement; two more do not discriminate at the top of this corpus at all.
6. **Two internal detectors return zeros of different kinds, and only one is a finding.** One threshold is reachable on this corpus, so its zero is genuine; two others cannot be reached by any possible input, so their zeros are **UNMEASURED** and carry nothing.
7. **Dates are earliest member filings, not priority dates** — `earliest_priority_date` is null for all 223 — so every era assignment is a **late bound**.
8. **"Lapsed" ≠ "expired", and no revival window was computed.** 45 of 91 adverse-event families remain live-granted via a sibling.
9. **The China-affiliation flag is assignee-string inference and PROVISIONAL.** Directional only.
10. **The assignee rollup is not an ownership record.** 148 keys, 117 singletons; two cohorts under one corporate roof keyed separately; licensee/originator distinctions were wrong until checked (§4.3).
11. **A quarter of the basis is medium-or-low confidence on its own claim-scope coordinates** (high 100 / medium 98 / low 25), and the grid inherits that.
12. **314 structural absences cannot be repaired by re-ingest** (ex-US claim translation needed); **202 repairable absences were submitted but not confirmed landed**; **30 requests hard-failed** on five named 2026 lanes; **223 discovery rows are unreachable by any available entrypoint**.
13. **The KC composite carries no resolvable `provenance_id`** — the trust chain stops at `kc.scope` / `kc.scored_publication` / `kc.scored_model`. For **128 of 223** families the producer scored a member other than the card publication.
14. **Two ontology defects shaped what retrieval could see:** the verified drug tier for this target is entirely antibodies, so the intracellular-lowering cohort is drug-name-invisible; and `search_terms("diranersen")` mis-resolves to a prekallikrein / hereditary-angioedema programme and was rejected wholesale at the operator gate.

15. **No prior-generation anti-Tau analysis exists**, so there is **no reproducibility check** on the claim-reading producer for this target, and no prior-versus-current delta section. The pre-run corpus state is **UNMEASURED** (all five baseline probes saturated at `k = 40`).
16. **One disclosed binning inconsistency** is left in place: family `43662090`, on-pathway despite `scope_evidence` reading "No tau limitation in the claim".

III-E. What the next run should do

1. **Add the five missing discovery slices** that make five cells UNMEASURED for the cheapest possible reason — nobody asked. `detect` slices for `MAPT_mRNA`, `MAPT_splicing`, `MAPT_transcription` and `tau_translation`; a `redirect_modulate` slice for `tau_aggregation`.
2. **Audit the 138 `UNMEASURED_UNAUDITED` modality cells.** This is where undiscovered room, if it exists, is.
3. **Re-run the baseline corpus probe cap-free** — a `search_facets` CPC enumeration rather than a `k`-capped semantic sweep — so a first-cohort claim rests on a census instead of a ranking.
4. **Raise the discovery cap or narrow the slices** on the 14 saturated slices, so the family denominator becomes measurable.
5. **Confirm the 202 repairable KC absences landed**, and re-derive; expect completeness rather than verdict changes.
6. **Escalate ex-US claim translation** — 314 structural absences, CN-heavy, on a target where China-affiliated filings went 0 → 1 → 2 → 14.
7. **Fix the unreachable detector threshold in the shared engine, not just here.** It cannot fire on any target, and at least one other target's report builds a cross-target comparison on it.
8. **Ask the platform for an entrypoint that resolves US application numbers**, which is the only thing standing between this run and its 223 unread rows.

Method, in brief — and what is not published here

How the substrate was built is in §III-A above and is unabridged: 48 term-sliced discovery calls across four jurisdictions, a two-channel relevance gate, a 933-submission ingest, a per-family claim-scope read that placed every family on four axes, and a per-family legal read. **What each family claims was read by a model from the claim text**, not from an abstract or a title, and a quarter of the basis is medium-or-low confidence on its own coordinates.

What is deliberately withheld, per the series' publication rules: the claim-strength scoring layer and every per-family and per-assignee score table derived from it (uncalibrated, relative-ordering only — publishing it as a ranking would assert a precision that does not exist and would put named-party quality judgements on the record); the knowledge graph and scope ontology; the scored inputs and the scripts; and the internal audit records. The reading is published; the engine is not.

What that means for a reader. Every structural number here — family counts, live-grant counts, adverse events, compiled-genus breadth, filing-date trajectory, cell verdicts — is reproducible from the public record with patience. None of it is a valuation, none of it is freedom-to-operate, and the guardrails in the read-me apply to every sentence above without exception.

Figure: `_downloads/target_heatmaps/Tau_whitespace_heatmap.png`.