PCSK9 — V4 Target Opportunity Report

Whitespace & Trajectory — public edition

Target: PCSK9 — proprotein convertase subtilisin/kexin type 9 (HGNC 20001; UniProt **Q8NBP7**), read together with its receptor-side partner **LDLR** (HGNC 6547; UniProt P01130) **Date:** 2026-07-29 · **Substrate:** 780 distinct discovery rows → 578 gate-admitted → **291 ingested patent families / 1,258 member publications**; **280 of 291 are on-pathway** after a per-family claim-scope read

> *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" means the **absence of strong structural claim coverage in this corpus**, which is **necessary but not sufficient** for opportunity. Compiled-genus breadth, genus-overlap fractions and sequence identity are **lower bounds**; an INDETERMINATE verdict counts as a candidate, never as clearance. **A lapsed patent is not public-domain and is not uniformly revivable** — 10 of the 17 on-pathway lapses here sit *inside* the ~24-month revival window. No expiry proxy carries any finding anywhere in this document. **An empty cell cannot distinguish "nobody tried" from "everybody failed."** Verify any specific family with counsel before relying on it. Method and limits: Part III.

PART I — ACT ON THIS

1. Actionable summary — where the room is

Ranked by how much the evidence supports acting **on the landscape**. **Every row is a candidate for counsel-led investigation. None is a clearance, an FTO position, or a recommendation to practise.**

#	Opportunity	Status	Why it's actionable	Watch / design around
1	**PCSK9 CHRD surface** (the C-terminal cysteine–histidine-rich domain,	**PARTIAL — 5 families, 1 live grant, no broad live genus**	The validated blocking lane (the Asp-374 / EGF-A groove) carries **136 families and 54 live grants**. The CHRD	The **only** live grant on the node is a **functional** genus — a pocket plus a binding constant — not a structural one. That is broad in one direction

#	Opportunity	Status	Why it's actionable	Watch / design around
	subdomains M1/M2/M3) — the surface the oral wave is actually moving to		carries **5 and 1**. AstraZeneca's oral small molecule `US-11603523-B2` claims a binding pocket defined by **PCSK9 residues Val589 and Ser636 — both in the CHRD**, not in the groove every approved agent targets. *(Residue placement verified independently against UniProtKB Q8NBP7: length 692, 589 = V, 636 = S, C-terminal domain 450–692.)* Literature support is **moderate, with no clinical validation**: the M2 subdomain is required for extracellular PCSK9's activity on surface LDLR, and a nanobody binding M1/M3 inhibits function (PMID 36566984)	and fragile in another. AstraZeneca's *structural* position is **pre-grant** (an 883-compound published application) or **unreadable** (a bibliographic-only 2025 WO filing). **And until 2026-05-12 the node carried TWO functional-genus grants:** IRCM/Montreal's `US-8673850-B2` claims *any* polypeptide ligand binding the CHRD or its M2 subdomain — it lapsed on 2026-05-12 and is **inside the revival window**
2	**PCSK9 autocatalytic maturation / furin site** — block the intramolecular ER event rather than the extracellular interface	**THIN — 1 family, 0 live grants, no genus**	The cleanest node-level thinness in the grid, on the **strongest literature of any thin node**: autocatalytic Gln-152 cleavage is *strictly required* for maturation; the Arg-218 furin site is inactivating; F216L and R218S are human gain-of-	The one family is a **live pending US application** claiming an antibody to the epitope at residues **209–218** — terminating exactly at Arg-218. Blocking an intramolecular ER event is materially harder than blocking a surface. And early mechanism chemistry is disproportionately

#	Opportunity	Status	Why it's actionable	Watch / design around
			function variants acting through cleavage state	concentrated in the **218 discovery rows this run could not ingest at all**
3	**Reversible / tunable epigenetic repression of the PCSK9 locus**	**OPEN cell inside a PARTIAL node**	The epigenetic node is the **fastest-rising mechanism in the landscape — 0 → 0 → 1 → 7** families by era. All 8 families currently bin to *permanent* silencing; **none of the readable ones claims reversible or tunable repression**	**Three reasons this may not be a real gap, stated in full.** (i) **The cell may already be occupied by the very company the thesis cites** — Omega Therapeutics' family is binned to permanent silencing at **low confidence on title-only evidence**, and its own binning note records that the direction *"could equally be expression knockdown if the controller is transient rather than heritable"*; the coherence-mask rationale that **created this cell** reasons that reversible repression *is* expression knockdown **and that Omega is one of this node's families**. (ii) **4 of the 8 families have no readable claims**, so "none claims reversible repression" is a statement about the readable half. (iii) Two sponsors just declined Europe — **Chroma Medicine 2026-04-29, Omega Therapeutics 2026-05-20**
4	**PCSK9 accessory-partner node —**	**PARTIAL — 3 families, ZERO live grants**	**No granted coverage at all**: both Aarhus University	**Weak-to-moderate literature with an explicit recorded**

#	Opportunity	Status	Why it's actionable	Watch / design around
	a third surface, neither EGF-A nor CHRD		grants lapsed (2024-08-20 and **2026-01-13**), the later one carrying a broad compiled genus. Two of the three families recite the **HSPG-binding surface at residues 78–167** *(the third claims a GRP94-derived polypeptide)*	**negative: PCSK9 has no obligate partner protein.** Sortilin, APLP2, CAP1, GRP94/HSP-family and annexin A2 were each searched, and each is modulatory, refuted, or unidentified — the "protein X" of the original genetics remains unidentified. **The thinness is partly the expected consequence of the biology**, and **both lapses are inside or near the revival window**
5	**Oral / route-chemistry convergence** — a second-source pre-positioning pattern	**Time-shaped signal, not a grid cell**	Oral small molecules are the **only strictly rising therapeutic format**: **0 → 7 → 8 → 12** by era, while full-length antibody filings collapse **28 → 13 → 3**. The synthetic-route / crystalline-form node rises **0 → 2 → 3 → 7**. On a target whose first-generation antibodies are approaching the end of their exclusivity runway, a simultaneous rise in route and polymorph filings is a recognisable generic/second-source pattern	The **readable** route/intermediate series is **flat at 0 → 2 → 3 → 3** — 4 of the 7 frontier families on that node are Chinese filings with zero claim text. Blocking-mechanism chemical space is **macrocycle-dominated**, not oral-small-molecule-dominated (see §3). And **61% of the compiled-genus overlap edges in this corpus are INDETERMINATE** — precisely the ones that would decide any design-around
6	**Read the**	**Largest**	**30 of 70 frontier-**	Ex-US claim translation

#	Opportunity	Status	Why it's actionable	Watch / design around
	frontier at all — the instrument gap	**actionable item, but it is a fix to the INSTRUMENT, not a position in the landscape — which is why it ranks last on a landscape criterion**	**era (2023–26) families — 43% — have no claim text.** 81 families corpus-wide could not be read (57 CN / 19 WO / 5 EP). This directly gates the epigenetic node (4 of 8), the AZD0780 combination stream (6 families, entirely unreadable), the CHRD node (1 of 5) and the secretion-trafficking node (**all 3 unreadable, 2 of them live Chinese grants**)	is the highest-value single investment available to this analysis. **The frontier is not thin; it is unreadable** — and unreadable in a jurisdiction-specific way
—	**The Asp-374 / EGF-A blocking interface** — generic entry	**AVOID**	**COVERED and crowded**: 136 on-pathway families, **54 live grants**, 6 broad live genera, 25 genus-bearing families. This is the mechanism litigated to the US Supreme Court in *Amgen v. Sanofi*	On this cell's own families the largest holders are Regeneron (12) and Merck Sharp & Dohme (8), with AstraZeneca and Salubris at 4 each — and **41 of the 136 have no resolved assignee**, so even that is a lower bound
—	**Formulation / dosing-regimen IP around the approved antibodies**	**AVOID (as discovery space)**	The **best-architected covered cell at any material scale** — 34 families and 4 broad live genera	A lifecycle observation, not a discovery opportunity

The one-sentence read: the validated lane is closed and well defended, and **the open room is on the surfaces the approved drugs do not touch** — the **CHRD** the oral wave is already moving to on five families, the **autocatalytic-maturation** event with no granted coverage, the **HSPG surface** whose only two grants have just lapsed, and the **reversible** corner of the fastest-rising epigenetic node — with the single largest caveat being that **43% of the frontier cannot be read at all**, so every count here is a lower bound.

2. Whitespace map — the coherent grid

The landscape was projected onto a scope grid built from four externally grounded axes — **mechanism node × direction × modality/chemotype × indication** (§II-A). A cell is a **whitespace candidate** when it is *scientifically coherent* but *structurally thin*: few on-pathway families, no broad **live-granted** compiled genus, and/or coverage held only by lapsing patents.

40 coherent cells: 6 COVERED · 6 PARTIAL · 7 THIN · 21 OPEN. A further **72 empty cross-product cells are incoherent** — category artifacts such as "an agonist of the transcript node" — and are excluded from the whitespace count rather than presented as opportunity.

Verdict rule, applied in code: - **COVERED** — ≥1 live grant AND (a broad live genus at compiled recall ≥ 0.5 OR ≥5 live grants) - **PARTIAL** — families present, but no broad live genus and <5 live grants - **THIN** — ≤2 families and no broad live genus - **OPEN** — zero families in this corpus

> **The coherence mask is load-bearing, and its first revision was wrong in a way that hid a live grant.** The first mask marked the diagnostic node's *inhibitor* direction incoherent — which suppressed a **granted, live** French public-sector family (`US-12253528-B2`, CNRS / CHU Nantes / INSERM: a PCSK9 theranostic for severe allergic disease). The revised mask fixes that, makes the residual direction coherent at all 14 nodes, and adds the epigenetic *expression-knockdown* cell. The grid went **32 → 40 coherent cells** and the defensible candidate count went **5 → 6**. *This is a change in the mask, not in the data — and it is disclosed rather than quietly re-baselined.*

COVERED cells — where not to build

node × direction	on-path families	live grants	raw register `granted_active`	broad live genus
`pcsk9_ldlr_egfa_interface` × antagonist_inhibitor	**136**	**54**	60	6
`pcsk9_delivery_formulation` × antagonist_inhibitor	34	11	13	4
`pcsk9_transcript` × expression_knockdown	32	12	14	1
`pcsk9_biomarker_dx` × diagnostic_none	21	5	5	0
`ldlr_axis_partner` × receptor_side_potentiation	4	3	3	1
`pcsk9_targeted_degradation` ×	1	1	1	1

node × direction	on-path families	live grants	raw register `granted_active`	broad live genus
expression_knockdown ⚠				

⚠ `pcsk9_targeted_degradation` is an axis value **added after the corpus was seen**, so a "COVERED" verdict on a single family is an artifact of the node having been created to house that family. **It is not a finding.**

Note the gap between the **raw register** column and the **live grants** column. Propagating lapse and Chinese adverse-outcome events out of the register rollups moves the whole landscape from **117 raw `granted_active` to 101 live-granted** and from 154 raw pending to **130 live-pending**. Reading the rollups alone would overstate live coverage by roughly 14% on grants — and would have shown candidate #4 as *having* granted coverage when it has **none**.

THIN cells — ≤2 families, no broad live genus

node × direction	families	live grants	lapsed	reading
`pcsk9_catalytic_maturation` × antagonist_inhibitor	**1**	**0**	0	**the sharpest node-level whitespace — candidate #2**
`pcsk9_chrd_surface` × modulator_unspecified	1	0	0	part of candidate #1
`pcsk9_secretion_trafficking` × expression_knockdown	2	**2**	0	**an information absence, not whitespace** — both are Chinese families with **zero claim text ingested**
`pcsk9_secretion_trafficking` × modulator_unspecified	1	0	0	information absence, same cause
`pcsk9_delivery_formulation` × receptor_side_potentiation	2	0	0	genuinely thin
`pcsk9_delivery_formulation` × modulator_unspecified	1	0	0	residual bin
`pcsk9_biomarker_dx` × antagonist_inhibitor	**1**	**1**	0	**recovered by the revised mask** — the French public-sector theranostic grant the first mask hid

*(The accessory-partner cell behind candidate #4 is formally PARTIAL, not THIN: 3 families, **0 live grants**, 2 lapsed.)*

OPEN coherent cells — and why most of them are not opportunities

21 coherent cells hold zero families. **Most are not opportunities, and saying so is the point.** The split sums exactly to 21:

n	OPEN cells	verdict on the verdict
11	the **residual** direction bin, at 11 different nodes	**NOT findings.** A cell empty because nothing fell into a catch-all is evidence the binning was decisive — good news about the method, and nothing about the landscape
3	*agonist / activator* at the transcript, gene-locus and epigenetic-locus nodes	**NOT opportunities.** There is **no therapeutic rationale for raising PCSK9** in hypercholesterolemia. Empty for the right reason
1	`pcsk9_targeted_degradation` × antagonist_inhibitor	**excluded** — ⚠ post-corpus axis value; its cells are empty by construction
3	`pcsk9_delivery_formulation` × three directions	**candidates with a binning caveat** — the delivery node absorbs families whose claims are about a formulation *of* something, and that "something" was often binned to delivery rather than to the mechanism delivered. Their emptiness is partly a classification consequence
1	**`pcsk9_secretion_trafficking` × antagonist_inhibitor**	**a real candidate** — the adjacent play to #2: block PCSK9 *secretion* rather than its maturation. It sits beside a *populated* cell (2 Chinese families, **both live grants, both unreadable**), which is what makes it a real gap rather than an axis artifact — and the same unreadability caps how strongly it can be claimed
1	**`pcsk9_epigenetic_locus` × expression_knockdown**	**a real candidate, and one the first mask had hidden** — candidate #3
1	`ldlr_axis_partner` × antagonist_inhibitor	a real but narrow candidate — the receptor-side node holds only decoy/potentiation filings; nothing claims *inhibiting* a partner receptor

Net: 6 of the 21 OPEN cells are candidates; 15 are artifacts of therapeutic logic, residual binning, or the post-corpus axis revision. Publishing "21 open cells" would be an overclaim. **Naming and dismissing the other 15 is the single most common step a whitespace map skips in order to overstate itself.**

3. The actionable chemical core — where the compiled genus actually is

Small-molecule and other chemical families are the most actionable subset for an inventor, because their scope is defined by a compiled Markush **genus** whose **breadth** — a recall figure that is always a **lower bound** — indicates how much analog space is fenced.

Convention: the family-level best recall is the ONLY recall figure used. Per-claim recall is null on most families and was never substituted (that substitution is the exact bug that falsely upgraded thin cells to covered on an earlier target in this series).

43 on-pathway families carry a compiled genus. 42 have a recall value; median 0.513; 23 of 42 are broad at ≥ 0.5. Three families' recall is **absent, not zero** — one is null with a failed compile, and two report exactly 0.000 at precision 1.000 **with a failed compile status**. A failed compile is not a measurement of nil coverage.

Where the actionable chemistry sits

- **The largest and best-defined compiled genera on the *blocking* mechanism are peptide macrocycles**, not oral small molecules — a Merck macrocycle genus spans 350 compounds at recall 0.629, and two Ra Pharmaceuticals macrocycle genera span 214 and 97 compounds at recall 0.509 / 0.804 with precision 0.96.
- **The oral small-molecule genera on that same node are materially narrower** — 49 and 34 compounds.
- **Contrast the CHRD filing behind candidate #1: 883 compounds but recall only 0.320, and not granted.**
- **Compound count is a poor proxy for position.** The ten largest compound collections in this corpus sit behind low recall, pre-grant status, a lapse, or **no compiled genus at all**. The clearest case is the *granted* CHRD family `US-11603523-B2` — **412 compounds and no compiled genus**, so it cannot be dismissed on breadth, and candidate #1 turns on exactly that claim's scope.

Three degeneracy cases, named rather than buried

A high recall is not automatically strong coverage:

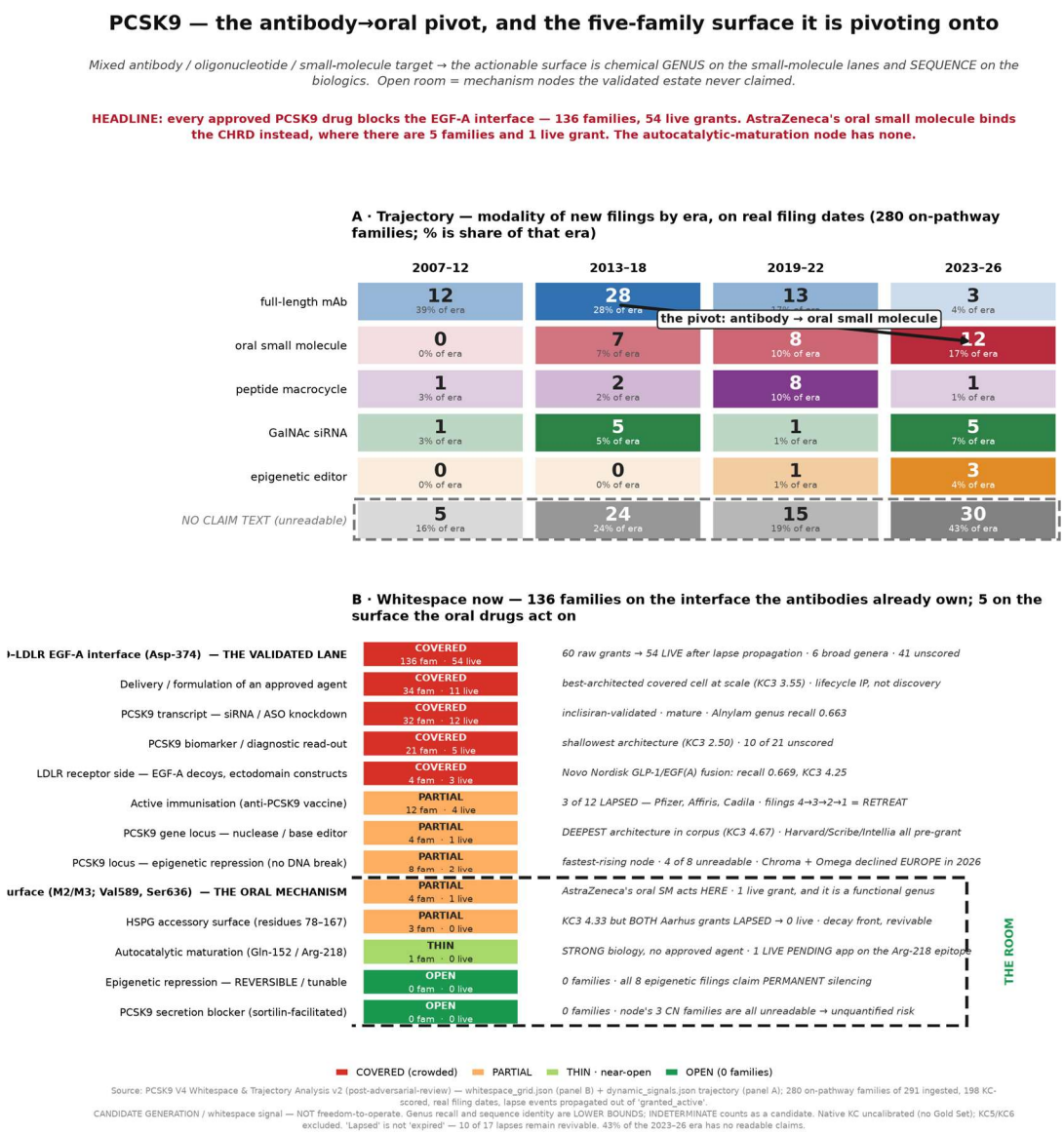
- One family reports **recall 1.000 at precision 0.005** — a **degenerate, non-discriminating genus** that captures everything because it excludes almost nothing. That is why precision is read alongside recall throughout, and never recall alone.
- A second reports **recall 1.000 at precision 0.45** (pending, so it never enters a broad-live count).
- **Precision does not detect an all-INDETERMINATE compile.** One family's genus carries a gate of `n_in: 0, n_indeterminate: 85, n_members: 85` and eleven provisos each ending *"membership is INDETERMINATE"* — image-only cores and unresolvable R-groups. **A zero-real-membership check was NOT applied to the genus layer in this run** and is recorded as a named method gap, not silently patched.

None of this is a design-around instruction, and a narrow genus is not a safe harbour. 61.2% of the 10,194 genus-overlap edges in this corpus are unresolvable INDETERMINATE — and those are precisely the ones counsel would need to resolve before any design-around could be relied on. The usable graded remainder has a **median overlap fraction of 0.125**.

---

4. Trajectory — where this target is going

Date basis: the earliest member **filing date** from each family digest. Earliest priority date is **null on all 291 family digests**, so filing date is the anchor — a conservative, late-leaning substitute for priority. **No expiry proxy is used anywhere in this analysis.** All 280 on-pathway families are dated; **zero undated**.



PCSK9 whitespace & trajectory heatmap

*Figure — **PCSK9: the antibody→oral pivot, and the five-family surface it is pivoting onto. Panel A** plots the format of new filings by era on real filing dates (share of era) for 6 of the 21 format values, with the arrow running from the antibody peak in 2013–18 to the oral-small-molecule peak in 2023–26 — it marks the **pivot**, not the crossover point, which falls between 2019–22 (13 antibody v 8 oral) and 2023–26 (3 v 12). The dashed **"NO CLAIM TEXT"** row is the honesty row — the unreadable share of each era, peaking at **30 of 70** families in 2023–26. **Panel B** plots **13 node × direction cells spanning 12 of the corpus's 14 mechanism nodes** (the epigenetic locus appears twice, under permanent silencing and expression knockdown), with the dashed **"THE ROOM"** box over five cells: CHRD × antagonist (PARTIAL), accessory-partner × antagonist (PARTIAL), catalytic-maturation × antagonist (THIN), epigenetic × expression-knockdown (OPEN) and secretion-trafficking × antagonist*

*(OPEN). Panels read their cell counts and the trajectory matrix **live from the underlying data at build time**; the annotation strings are hard-coded and can drift.*

Mechanism node by era

node	pre-2007	2007–12	2013–18	2019–22	**2023–26**
pcsk9_ldlr_egfa_interface	0	12 (39%)	54 (53%)	39 (51%)	**31 (44%)**
pcsk9_transcript	0	5 (16%)	11 (11%)	7 (9%)	9 (13%)
pcsk9_delivery_formulation	0	6 (19%)	14 (14%)	10 (13%)	7 (10%)
pcsk9_epigenetic_locus	0	0	**0**	1 (1%)	**7 (10%)**
pcsk9_process_intermediate	0	0	2 (2%)	3 (4%)	**7 (10%)**
pcsk9_biomarker_dx	1	3	7	7	4 (6%)
pcsk9_host_immunisation	0	4 (13%)	5 (5%)	2 (3%)	**1 (1%)**
pcsk9_chrd_surface	0	1	0	2	2 (3%)
pcsk9_gene_locus	0	0	2	1	1 (1%)
ldlr_axis_partner	0	0	3	1	**0**
pcsk9_accessory_partner	0	0	**3**	0	**0**
pcsk9_secretion_trafficking	0	0	0	2	1
pcsk9_catalytic_maturation	0	0	0	**1**	0
pcsk9_targeted_degradation ⚠	0	0	0	**1**	0
era total (on-pathway)	**1**	**31**	**101**	**77**	**70**

All five columns sum exactly to their era totals.

Format by era — the sharpest signal in the analysis

format	2007–12	2013–18	2019–22	**2023–26**
full-length antibody	12	28	13	**3**
oral small molecule	0	7	8	**12**
peptide macrocycle	1	2	**8**	1
GalNAc-siRNA	1	5	1	5
antisense oligonucleotide	2	3	1	**0**
epigenetic editor	0	0	1	**3**
nuclease editor	0	1	1	1
vaccine / immunogen	4	3	2	**1**
formulation / device	1	9	7	**1**

format	2007–12	2013–18	2019–22	**2023–26**
synthetic route / intermediate	0	2	3	**3**
no claim text ingested	5	24	15	**30**

Unlike the node table, ***this table's columns do not sum to the era totals and are not meant to*** *— it shows 11 of the 21 format values and drops the single-family pre-2007 column. Ten smaller format values are omitted.*

Trajectory statement: *PCSK9 is a* ***validated, closed, and now-decaying antibody target*** *transitioning to an* ***oral small-molecule and genetic-medicine*** *landscape. Full-length antibody filings fell ~90% from their 2013–18 peak (28 → 13 →* ***3****) and are now outnumbered four-to-one in the frontier era by oral small molecules (****12****), the only strictly rising therapeutic format. The* ***epigenetic editors are the fastest-rising mechanism node*** *(0 → 0 → 1 →* ***7****) — while two of their sponsors declined Europe in 2026 and none of the readable families claims reversible repression.* ***Synthetic-route and crystalline-form filings rise in parallel****, a recognisable second-source pre-positioning pattern on a target whose first-generation exclusivity runway is ending. The* ***vaccine sub-lane is the clearest sponsor retreat*** *— filings 4 → 3 → 2 →* ***1****, with the Pfizer/Rinat, Affiris and Cadila/Zydus grants all lapsed. Entrants shift from a Western-originator land-grab on the EGF-A interface to a* ***China-weighted frontier — which is also where claims are unreadable****, so the entrant shift is real but amplified by what this pipeline can and cannot read.*

The single most important caveat on this section: *no claim text* accounts for **30 of 70 families in the current era — 43% of the frontier cannot be read at all**, and it is unreadable in a jurisdiction-specific way. **The frontier is not thin; it is unreadable.** Any trajectory claim about 2023–26 is a claim about the readable minority.

The decay front — and why "lapsed" is not "free"

17 non-payment lapse events sit on on-pathway families (19 across the whole ingested corpus). All are reported as public-register facts, with no inference about intent.

date	patent	node	inside the ~24-month revival window?
2020-01-21	US-7300754-B2 (INSERM)	biomarker / dx	**no** — register reads `expired`
2020-04-21	US-9266961-B2	EGF-A interface	no
2022-02-01	US-8598139-B2 (Alnylam)	transcript	no
2023-01-17	US-8889144-B2 (Pfizer/Rinat)	host immunisation	no
2023-11-07	US-9127280-B2 (Univ Osaka)	transcript	no
2024-	US-9220762-B2 (Affiris)	host	no

date	patent	node	inside the ~24-month revival window?
02-27		immunisation	
2024-04-09	US-9255154-B2 (AlderBio)	EGF-A interface	no
2024-08-20	US-10688119-B2 (**Aarhus Univ**)	**accessory partner**	**yes**
2025-02-04	US-10858446-B2 (Pharmaexplorer)	EGF-A interface	**yes**
2025-09-16	US-9707155-B2	delivery / formulation	**yes**
2025-09-30	US-9717649-B2	delivery / formulation	**yes**
2025-10-14	US-20200207743-A1 (Cardio Therapeutics)	EGF-A interface	**yes**
2026-01-13	US-11173177-B2 (**Aarhus Univ**)	**accessory partner**	**yes**
2026-01-27	US-8598320-B2 (**Merck Sharp & Dohme**)	EGF-A interface	**yes**
2026-05-12	**US-8673850-B2 (IRCM / Montreal)**	**CHRD surface**	**yes**
2026-05-26	US-11286309-B2 (Zhejiang Blue Shield)	EGF-A interface	**yes**
2026-07-07	US-11325945-B2 (Cadila / Zydus)	host immunisation	**yes**

*Three rows are listed by publication number without an assignee: two have **no assignee on record** in the source data, and one is omitted for publication-gate reasons only — all three are public-register facts and the full attribution is in the internal edition.*

The correction that matters most for candidates #1 and #4. A blanket "revival is petitionable for roughly 24 months" is **false for 7 of these 17**. **10 are inside the window and are therefore emphatically not free** — including *both* Aarhus grants under candidate #4 and the IRCM CHRD functional genus under candidate #1. Exactly one on-pathway family reads `expired` on the register, and because this run bars any estimated-expiry figure from carrying a finding, the honest statement is that *the register reads expired and this analysis does not rely on the estimate*. **A lapsed patent is not public-domain.**

Census scope: this counts **only** the US event string for non-payment lapse. One further on-pathway family carries a `lapsed` register verdict with no such US event and is therefore **not** in the 17. "17" is a **US-scoped lower bound**, not a complete decay census.

Four other decay signals, distinct from a lapse and reported separately: - **EP non-entry — 16 families carrying 17 events.** PCT retreat, not abandonment of the invention. The two that matter most are **Chroma Medicine (2026-04-29)** and **Omega Therapeutics (2026-05-20)**, both on PCSK9 epigenetic-regulation families, in the fastest-rising node — candidate #3. - **24 Chinese families** were rejected, withdrawn or deemed abandoned after substantive examination **while still carrying a `pending` rollup verdict**. - **Two terminal disclaimers**, one of them on the inclisiran Orange Book family (2026-06-02). - **The vaccine sub-lane retreat** above.

Combination adjacency — and why no claim about it is supportable

23 on-pathway families recite a *named* combination partner; a further 10 hit only the generic word "combination" with no partner named — **33 records, 23 named.** The distribution is thin and concentrated: **statin (14 filings)** dominates, spanning the antibody, oral small-molecule and siRNA nodes; then **ezetimibe (3)**, **GLP-1 (3)** — including a **GLP-1 / EGF(A) fusion**, the most structurally unusual claimed object in the landscape — **PD-1 (2)**, **seladelpar (2)**, **PD-L1 (1)** and **bempedoic acid (1)**. The substrate tokenises only the generic `statin`, so no specific-statin claim is asserted here.

The combination axis is materially under-read in this run, because the two most current combination estates — AZD0780 + statin and AZD0780 + ezetimibe — are both **unreadable WO families with no claim text**. **Any statement that combination space is open would be unsupportable.**

PART II — THE SCOPE-COORDINATE SYSTEM

II-A. How "whitespace" is defined here

Whitespace only exists relative to a coordinate system, and if that system is derived from the patent set the analysis is circular. Four axes, **externally grounded** (UniProt / NCBI Gene / peer-reviewed literature, plus Wikidata SPARQL) and operator-gated, frozen **before** any family was binned:

- **mechanism node (15):** the EGF-A/Asp-374 blocking interface · the CHRD surface · autocatalytic maturation · secretion/trafficking · accessory partner · transcript · gene locus · epigenetic locus · targeted degradation · host immunisation · delivery/formulation · process & intermediate · biomarker/diagnostic · the LDLR-axis partner · plus an off-pathway sink
- **direction (8):** antagonist/inhibitor · agonist/activator · expression knockdown · permanent silencing · host immunisation · receptor-side potentiation · diagnostic · plus a residual bin
- **modality / chemotype (21):** full-length antibody · binding-protein scaffold · peptide macrocycle · oral small molecule · natural product · GalNAc-siRNA · antisense · nuclease editor · epigenetic editor · viral-vector gene therapy ·

vaccine/immunogen · EGF(A) decoy · lipid nanoparticle · formulation/device · assay reagent · synthetic route/intermediate · unresolved categories · …

- **indication (12):** primary hypercholesterolemia / dyslipidemia · monogenic FH · ASCVD / CV risk · elevated Lp(a) · metabolic-hepatic · inflammation/immune · oncology · other non-CV · platform/none · …

Ontology revision is disclosed, not hidden — including the limits of the disclosure. The axes were frozen before binning, which is what licenses reading an empty cell as a real absence. One revision was then made *after* the corpus was seen: two axis values were **added** because families had no legal home (independence **compromised for those cells only** — their population counts are not independent evidence, and both are flagged ⚠ wherever they appear), and one catch-all was **split** into "no claim text" and "functional genus unspecified" (independence **intact**, since a split re-labels families already binned). **Three coherent grid cells derive from the added values and none is treated as a finding.** The reader is also entitled to know that the "no claim text" label — which carries this document's most-repeated headline, **43% of the frontier** — post-dates the corpus, even though the split itself is independence-intact.

Wikidata grounding (SPARQL, both directions of the gene→disease property, agreeing at 6 rows each): disease anchors are **familial hypercholesterolemia** and **ADH3/FHCL3** from PCSK9, plus three LDLR-linked items — **lipid metabolism disorder**, **abdominal aortic aneurysm** and **cardiovascular disease**. **Two of those three are disease *classes*, not conditions, and their LDLR links are data-quality artifacts** — a gene is not "genetically associated with" the class *cardiovascular disease*, and they are flagged as such. The grounding also carries **two first-class recorded gaps**: **ASCVD / CV risk has no genetic-association path from PCSK9** in Wikidata (it is literature-grounded only, so the Wikidata anchor *under-states* the indication axis), and **no Wikidata item exists for elevated lipoprotein(a)** despite its commercial prominence. One marketed PCSK9 antibody has **no Wikidata item at all** — confirmed by SPARQL non-hit *and* a direct entity-search cross-check returning zero. Those are coverage gaps in the reference, not absent biology or absent drugs.

A recorded negative literature finding shapes candidate #4, against this document's own interest: PCSK9 has no obligate cofactor. Sortilin, APLP2, CAP1, GRP94/HSP-family and annexin A2 were each searched; each is modulatory, refuted (one is directly contradicted between two published reports), or unidentified. The accessory-partner node is therefore thin partly *because of the biology* — and that is stated rather than sold.

PART III — METHOD & LIMITS

III-A. How the substrate was built

1. **Discovery — term-sliced, not jurisdiction-partitioned.** Synonym and code-name expansion on PCSK9 (verified gene/protein-database tier plus an operator-reviewed drug-name tier, with **enlicitide / MK-0616** added at the operator gate),

then **102 discovery calls across 7 term-slice groups**, 50 US and 52 ex-US. **1,284 raw rows → 780 distinct**; by jurisdiction US 311 / EP 195 / CN 173 / WO 101.

- **A verified negative worth recording:** the jurisdiction selector **does not partition** — three different single-jurisdiction requests returned *positionally identical* 100-row sets. They are aliases, not partitions. **Term slicing, not jurisdiction slicing, is what breaks the result cap.**

2. **The relevance gate.** 780 rows → **578 admitted / 202 dropped** (201 returned only by an indication-term *control* slice with no PCSK9 token anywhere; 1 a false-friend string collision).
   - **The load-bearing rule, learned by getting it wrong first:** *a title token may **admit** a row but must never be permitted to **reject** one that a target-specific query returned.* An earlier title gate systematically removed **method-of-treatment and dosing filings, polymorph and crystalline-form filings, and synthetic-intermediate filings** — three of the highest-value whitespace-adjacent categories on a validated target. A landscape built on a title filter under-counts them structurally, on every target.
3. **Ingest.** 360 targets, of which **218 had no ingest entrypoint at all** (US pre-grant/pending rows carrying only an application number). **362 requests submitted, 352 completed, 10 hard-failed** on upstream snapshot lag for 2025–26 publications. Result: **291 families / 1,258 members / 393 full-text members.**
4. **Off-pathway screen.** A per-family claim-scope read returned **279 on-target / 12 off-target**; 2 further families were excluded by recorded analyst override. The off-target modes were instructive: two omnibus "laundry-list" filings whose claims are a shotgun sequence genus over ~40 unrelated clones, a misresolved anchor claiming an unrelated synthase, and a pure string collision on a **thin-film-transistor** patent via an assignee code.
5. **Scope binning.** All **291** families binned onto the four axes by a **claim-scope read of each family's claims, summary and compiled genus** — never the title alone — run as 6 parallel shards, each family carrying an evidence string and a confidence grade (82 high / 116 medium / **82 low**). Zero illegal axis values. **280 on-pathway / 11 off-pathway**, agreeing with the independent off-target screen on 11 of 12 with one reasoned promotion.
6. **Structural layer.** Compiled-genus retrieval over the genus-bearing families (family-level best recall as the sole convention) plus scope-neighbour retrieval over 61 families; legal timelines re-joined so that lapse and adverse-outcome events propagate out of the live counts; real filing dates for all 280 on-pathway families.
7. **Reproducibility.** The whole chain runs from a single stage-ordered script whose order encodes three real dependencies discovered the hard way. Every script is idempotent, and two verifiers assert the numbers in these documents against the underlying data — one of which **parses the tables out of the prose** and diffs them cell-by-cell, because a verifier that recomputes a value the way the document computed it confirms the document rather than testing it.

III-B. The two load-bearing reclassifications

Both cut *against* an inflated landscape:

1. **11 of 291 families are off-pathway** on the claim-scope read (14 excluded from the analysis basis on the parallel screen). They co-mention PCSK9 but **claim something else.** Every whitespace, trajectory and grid number here is computed on the **280 on-pathway** families; the off-pathway set is retained only as combination-adjacency signal.
2. **Lapse and adverse-outcome propagation moved the coverage numbers materially** — raw `granted_active` **117 → 101 live-granted**; raw pending **154 → 130 live-pending**. Reading the register rollups alone would have shown candidate #4 as having granted coverage when it has none.

A third, smaller reclassification is disclosed for honesty rather than impact: the **coherence-mask revision** described in §2, which recovered a live grant the first mask had hidden.

III-C. Adversarial review

Every layer of this analysis was reviewed by an impartial adversarial reviewer whose brief was to find errors, and **every finding was applied.** The reviews were not cosmetic:

- The **whitespace/trajectory layer** returned *accept with fixes* on **14 must-fix and 12 should-fix** findings. Two changed conclusions: the coherence mask was found to be **hiding a live grant**, and the sequence-neighbour layer was found to be a **51% sample** because of an undocumented 50-rows-per-publication cap in the retrieval API. **Two of the reviewer's own findings were refuted on evidence** and are recorded as refuted — a review that cannot be wrong is not a review.
- The **combined report** returned **reject**, with five errors inside the decision surface. All 34 findings were applied. The most consequential was a **structural-edge targeting bug in the underlying knowledge graph that both prior reviews had missed** — a substantial share of its edges pointed at the source family itself, inside a headline edge count that both reviews had verified as *a count*. **A count is not an integrity check**; the build now asserts the properties the artifact must have. A second was an **assignee mis-attribution** that had supported a competitive-intelligence reading, now downgraded to a hypothesis pending a corporate-registry check.
- A third finding is disclosed in candidate #3 itself: the reasoning that **created** that cell implies the cell may already be occupied by the company whose positioning the candidate cites. It is stated in full rather than deleted or sold.

III-D. Limits — read before relying

- **Not FTO, not a clearance, not a valuation.** Whitespace = absence of *strong structural claim coverage in this corpus* (**necessary, not sufficient**). Compiled-genus recall, overlap fractions and sequence identity are **lower bounds**; INDETERMINATE counts as a candidate. **An OPEN or THIN cell is a hypothesis prompt, not clearance.**

- **43% of the frontier era is unreadable. 30 of 70** 2023–26 families have no claim text; **81 families corpus-wide** could not be read (57 CN / 19 WO / 5 EP). **The frontier is not thin; it is unreadable** — and unreadable in a jurisdiction-specific way.
- **218 of 578 gate-admitted discovery rows (38%) had no ingest entrypoint**, concentrated in the 2024–26 US genetic-medicine frontier and in two large continuation streams. **Every count in this document is a lower bound.**
- **10 named ingest failures** on upstream snapshot lag for 2025–26 publications — including an AstraZeneca crystalline-forms WO filing that sits **directly on candidate #1**. Named absences, never silent.
- **The AZD0780 combination sub-lane is unscored** — 6 WO families, entirely unreadable. **No claim about combination whitespace is supportable from this run.**
- **`lapsed` ≠ `expired` ≠ public domain.** 10 of the 17 on-pathway lapses are inside the ~24-month revival window, and the census is US-scoped.
- **The denominator is operator-defined.** Reachable space is set by the ontology axes; different axes → different whitespace. Three coherent cells derive from post-corpus axis values and none is treated as a finding.
- **Classification is a model read.** The 280/11 split and every per-family bin are claim-scope reads carrying an evidence string and a confidence grade — **82 of 291 at low confidence**. This document emphasises cell-level patterns over single-family calls, and the one candidate that turns on a single low-confidence bin says so.
- **Named method gaps, not silently patched:** (i) **no zero-real-membership degeneracy check was applied to the genus layer** — an all-INDETERMINATE compile can pass the precision screen, and a demonstrated case is named in §3; (ii) the sequence-neighbour layer is permanently at **51% capture** unless re-pulled per publication, and **no residue-level inspection** was performed on its 16 framework-scale candidate hits, so their benign explanation remains an untested hypothesis; (iii) the "a failed compile is not a measurement" rule is applied to the two zero recalls but **not** to the 17 non-zero recalls that also come from failed compiles — **including the 0.320 figure behind candidate #1**. That asymmetry is recorded, not resolved.
- **Two claims here are analyst context, not products of this run:** that this is the target litigated in *Amgen v. Sanofi* (true, but no litigation record was retrieved by the pipeline), and that the simultaneous rise in route and polymorph filings is "a recognisable second-source pattern" (pattern recognition, and the accompanying statement that a first-generation exclusivity runway is ending has **no term data behind it** — this run bars any expiry estimate from carrying a finding).
- **Corpus-bounded and US-anchored.** Whitespace is relative to the 291 ingested families; ex-US members materialise at bibliographic tier; **genuinely un-filed art is invisible by construction**, and an empty cell cannot distinguish "nobody tried" from "everybody failed."

III-E. What the next run should do

1. **Ex-US claim translation on the named 81-family unreadable list** (57 CN / 19 WO / 5 EP) — the highest-value single investment available. It directly gates candidate #3, candidate #6, the AZD0780 combination stream and the secretion-trafficking node (all 3 unreadable, 2 of them live Chinese grants).
2. **Resolve the 218 no-entrypoint US pending rows** — an application-number ingest path would open the 2024–26 genetic-medicine frontier this run can only name, not read.
3. **Re-pull the sequence layer per publication** to defeat the 50-row cap (476 true hits versus 245 captured), and **perform the residue-level check** that would confirm or refute the benign reading of its 16 framework-scale candidate hits.
4. **Apply a zero-real-membership degeneracy screen to the genus layer** alongside the existing precision screen, and re-run the broad-genus counts.
5. **Retry the 10 snapshot-lagged ingests**, priority on the AstraZeneca crystalline-forms filing that bears on candidate #1.
6. **Run the two corpus-independent searches the candidates require and this run did not:** a primary search on maturation / propeptide / Gln-152 / Arg-218 claim language plus a tractability read before any IP work on candidate #2 — an empty cell in *this* corpus is not evidence of an empty field — and **monitoring of the revival windows** on the two Aarhus grants and the IRCM CHRD grant, all three of which are inside the window.
7. **Pull the AstraZeneca CHRD prosecution histories** and test whether the granted pocket-plus-binding-constant claim reads on non-AstraZeneca chemistry. **This is the single highest-value counsel-led follow-up in the report.**

Public edition, 2026-07-29. Whitespace and trajectory computed on the 280 on-pathway families of 291 ingested, on real filing dates, with lapse and adverse-outcome events propagated out of the live counts. Both source analyses were independently reviewed by adversarial reviewers with every finding applied; the reviews, the audit records, the claim-strength scoring layer, the knowledge graph and the machine-readable inputs are part of the internal analyst deliverable and are not reproduced here. ***Candidate generation, NOT freedom-to-operate.***