From the candidate universe to CTAs

OncoRef 1.35.0 | 8 November 2026

Research candidate selection, evidence status and verified sequence use.

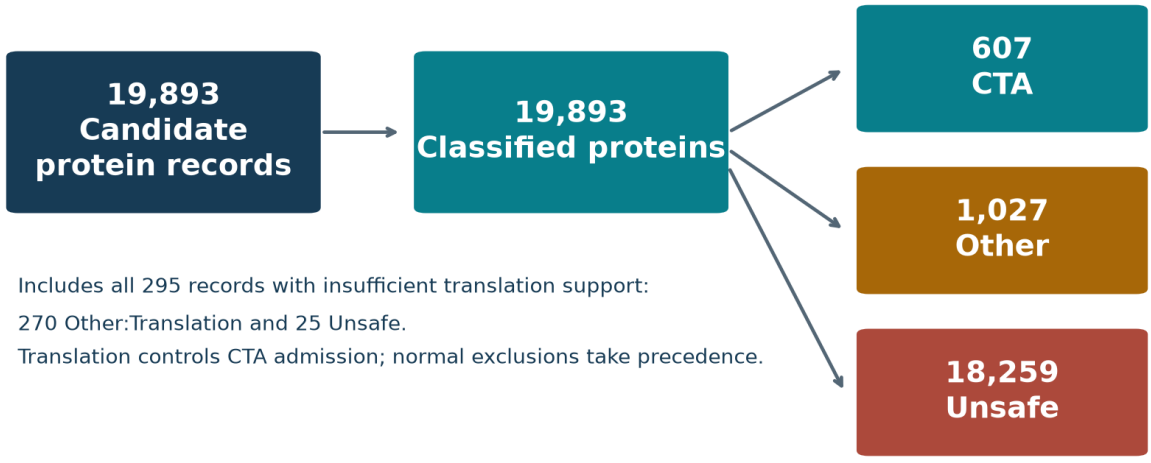
What OncoRef calls a CTA

An OncoRef CTA candidate is a reference-resolved protein with supported cancer expression and qualifying reproductive or developmental association that passes the named normal-tissue evidence, coverage and selection rules. CTA admission requires qualifying coding annotation or supported ncORF evidence and a complete, approved MANE Select or Ensembl canonical representative, with any conflicting coding evidence resolved. Every input record is assessed, including those with insufficient translation support.

The current human policy produces 607 CTA, 18,259 Unsafe and 1,027 Other proteins. It evaluates 19,893 candidate proteins in the complete inventory, counting identical selected reference sequences together and retaining every member locus. Splice isoforms and post-translational modifications are outside this inventory.

CTA admission combines cancer expression, reproductive or developmental association, translation qualification and normal-tissue assessment. The biological categories identify the tissue, cell population or developmental program associated with each protein.

Every candidate record receives a classification



Identical selected reference sequences count together. Member loci remain separately traceable.

From candidates to CTAs

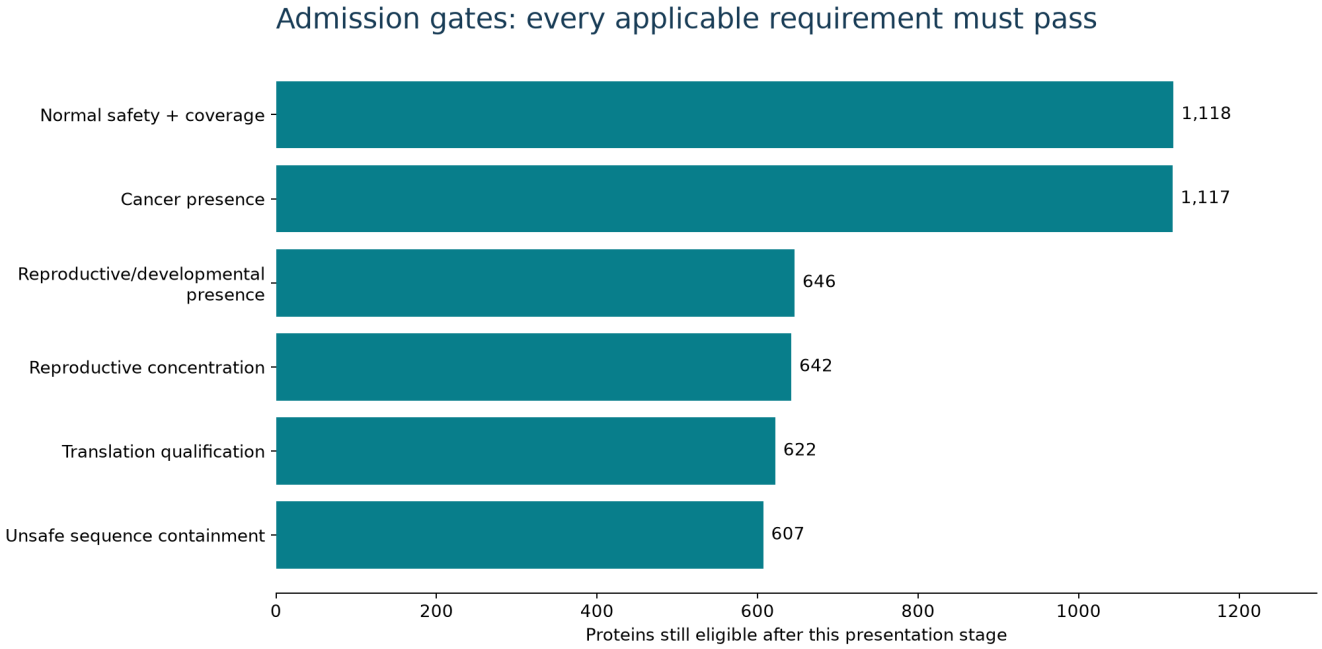
All 19,893 candidate records enter the same normal-evidence and admission path. The 295 records formerly removed at entry stay visible: 285 have only NMD/CDS-not-defined reference transcripts and 10 have only single-study Ribo-seq support. Translation support is assessed at CTA admission.

CTA admission requires qualifying translation support and the separately qualified exact reference record. Applicable normal exclusion takes precedence and establishes Unsafe; otherwise insufficient translation qualification yields Other:Translation and sequence review. The translation rule accepts qualifying coding annotation and supported noncanonical ORFs.

Apply all normal-tissue evidence independently. Accepted exclusion evidence establishes Unsafe. Unresolved findings or missing required coverage prevent admission and require review; they do not establish Unsafe. Every target must also pass reproductive/developmental association, cancer presence, concentration, translation qualification and sequence-containment rules.

The chart uses cumulative counts. Of 1,118 proteins passing normal assessment, cancer presence retains 1,117; reproductive/developmental presence then narrows these to 646. The previous order showed 645 at both gates because every reproductive/developmental survivor already met cancer presence. One compatible positive tumor RNA sample or reviewed tumor protein makes cancer presence a broad gate, not an abundance or prevalence requirement.

Every applicable evidence layer is still evaluated, including after an early exclusion. Unsafe takes precedence. Display order changes neither final classifications nor the priority used to assign Other subtypes.



Published nominations through the funnel

All 15 registered nomination papers appear in the companion funnel. Their union contains 2,514 proteins: 487 finish as CTA, 1,876 as Unsafe and 151 as Other. A protein can occur in several papers, so individual paper rows must not be summed.

The 17,379 candidates without a registered-paper annotation undergo the same assessment and contribute 120 CTAs. Prior nomination is recorded separately from the admission gates. The evidence guide identifies additional publications for proteins outside this registry.

A publication's measured evidence can support reproductive/developmental association, tumor expression, translation or normal-tissue counterevidence when it meets the corresponding evidence rule. A list label alone does not establish those facts. Germ-cell, testis-biased, placental and tumor-vascular lists retain their original scope alongside CTA nominations.

Additional cancer-embryo catalogs

Supplementary-table matching found 11 distinct proteins from the 122 across these two catalogs. These program memberships are separate from the release's 15-paper nomination collection.

Chew et al., 2019: DUX4 Suppresses MHC Class I to Promote Cancer Immune Evasion and Resistance to Checkpoint Blockade

Table S3: high-confidence DUX4 targets: 4 matches: DUXA, DUXB, LEUTX, PRAMEF25/6. DUX4-induced gene program evaluated in human cancers; target definition also used induced myoblast experiments.

PRAMEF25 matches the historical WI2-2994D6.2 entry by Ensembl ID and is one member of the shared PRAMEF25/6 protein. DUXB matches by symbol; its source Gene ID is also DUXB. The paper reports DUXB expression in healthy tissues.

Smith et al., 2023: DUX4 expression in cancer induces a metastable early embryonic totipotent program

Table S3A: zygotic gene activation and eight-cell-like gene sets: 9 matches: C19orf85, CPHXL, DUXA, LEUTX, MBD3L2B, OR10G3, TAF11L12, ZNF280A, ZNF878. Single-cell RNA sequencing of KLE and SUSA cancer cell lines using published zygotic gene activation and eight-cell-like gene sets.

Exact gene-symbol matches to the 122. Membership records the program evaluated in cancer cells; gene-level measurements and effect sizes need separate extraction. MBD3L2 is not substituted for MBD3L2B, and H3Y is not mapped to H3Y1 in this count.

Gene-set source: Taubenschmid-Stowers et al., 2022: 8C-like cells capture the human zygotic genome activation program in vitro.

Matched source rows, download links and file hashes

Every registered nomination paper follows the same admission gates

	Input	Normal assessment	Cancer presence	Repro. / dev.	Concentration	Translation	Final CTA
da Silva et al. (2017)	1,017	512	512	432	432	427	418
Wang et al. (2016)	988	292	292	279	279	276	274
Chang et al. (2019)	985	292	292	289	289	287	287
Gong et al. (2021)	729	370	370	357	356	354	351
Bruggeman et al. (2018)	692	12	12	12	12	12	12
Jamin et al. (2021)	469	39	39	38	38	37	34
Loriot et al. (2025)	202	86	86	86	86	84	83
Carter et al. (2023)	103	54	54	54	54	54	54
Hirbo et al. (2015)	78	10	10	10	10	10	8
Seager et al. (2024)	15	11	11	11	11	11	11
Sasaroli et al. (2011)	13	0	0	0	0	0	0
Bradley et al. (2020)	9	0	0	0	0	0	0
Garg et al. (2007)	5	3	3	3	3	3	3
Boël et al. (1995)	2	1	1	1	1	1	1
Bai et al. (2016)	1	0	0	0	0	0	0
Registered-paper union	2,514	592	592	505	504	497	487
No registered-paper annotation	17,379	526	525	141	138	125	120
Complete inventory	19,893	1,118	1,117	646	642	622	607

Counts are cumulative. Color shows the fraction of each row's input remaining.
Paper rows overlap and must not be added. The union and unannotated rows partition the full inventory.
Nomination scopes include CTA, germ-cell, placental and vascular lists. Membership records provenance; measured evidence determines admission.

Download every paper-by-gate count and its source scope

Registered nomination papers and extraction scope

da Silva et al. (2017): At least 90% normal-tissue RNA attributable to testis; tumor positivity not required; Gene-labelled tumor proteomics; unique-locus peptide assignment is not established; RSEM >1 in at least 10% of samples in a tumor type; RNA nomination

Wang et al. (2016): Cancer-associated RNA nomination; no antigen validation asserted

Chang et al. (2019): Testis-specific coding gene expressed in at least 1% of TGCT; no broad somatic-cancer claim

Gong et al. (2021): Normal reproductive-tissue enrichment; no tumor validation; Placental RNA enrichment; no tumor-antigen validation

Bruggeman et al. (2018): GC RNA nomination across cancer types; no antigen validation asserted; GC RNA nomination; inclusion depends on TGCT under this study's thresholds

Jamin et al. (2021): CT-pattern microarray nomination; tumor upregulation with no corresponding normal-tissue detection; Core CT annotation; underlying microarray probes may recognize several loci

Loriot et al. (2025): Cancer-germline RNA nomination; strict and preferential classes remain distinct

Carter et al. (2023): Final CT-expression set without detectable thymic expression; functional vaccination evidence is not assigned to every gene

Hirbo et al. (2015): HPA v13 placenta-enriched nominations; 30/86 fast-evolving subset retained as metadata; not tumor validation

Seager et al. (2024): Member of the published 17-target CTA profiling panel; includes differentiation antigen MLANA

Sasaroli et al. (2011): Tumor-vascular candidate list; includes secreted and transmembrane proteins; not a CTA-restricted list

Bradley et al. (2020): CPA expression nominations; HLA peptide and antigen-specific T-cell validation for VGLL1

Garg et al. (2007): Ovarian-cancer RNA/protein expression and patient antibody responses; no HLA presentation or T-cell validation claimed

Boël et al. (1995): Original BAGE tumor-antigen nomination and autologous CTL recognition; contemporary paralog resolution is not inferred

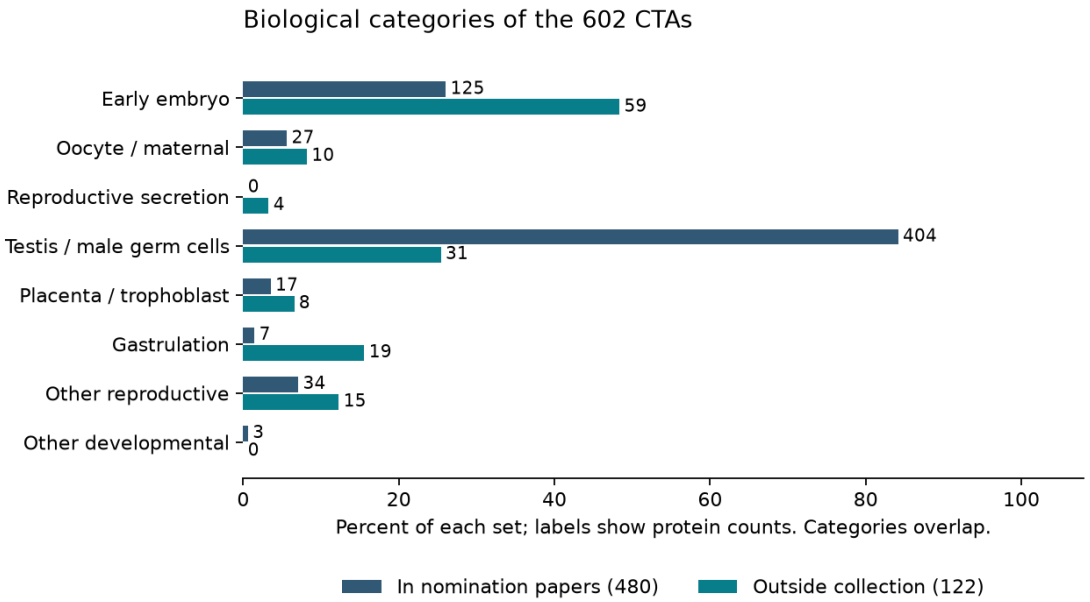
Bai et al. (2016): Ovarian-tumor and vascular expression/function; no normal-tissue exclusivity claim

Biological categories in the historical 602-CTA snapshot

The historical 1.31.0 literature tables categorize 602 CTAs by early embryo, oocyte and maternal function, reproductive secretion, testis and male germ cells, placenta and trophoblast, and gastrulation. A protein can belong to several categories. Current reference choices and coding holds are listed in the transcript-selection report.

Each tag links to a publication or a supported RNA observation. The detail view records the species, tissue, cell population or developmental stage. Biological categories are separate from the admission route and the CTA, Unsafe or Other classification.

Reproductive secretion currently has four manually annotated proteins: PATE1, WFDC8, RNASE9 and SEMG2. This count reflects the focused literature review; the other CTAs have not been systematically checked against reproductive-secretion catalogs.



Early embryo: 184 of 602 CTAs; 59 of the 122 outside the paper collection.

Oocyte / maternal: 37 of 602 CTAs; 10 of the 122 outside the paper collection.

Reproductive secretion: 4 of 602 CTAs; 4 of the 122 outside the paper collection.

Testis / male germ cells: 435 of 602 CTAs; 31 of the 122 outside the paper collection.

Placenta / trophoblast: 25 of 602 CTAs; 8 of the 122 outside the paper collection.

Gastrulation: 26 of 602 CTAs; 19 of the 122 outside the paper collection.

Other reproductive: 49 of 602 CTAs; 15 of the 122 outside the paper collection.

Other developmental: 3 of 602 CTAs; 0 of the 122 outside the paper collection.

Filter the CTA table by biological category and open the supporting evidence

Normal-tissue consequences and expression evidence

Vital, Protected, Manageable and Disposable describe the consequence and tolerability of injury to a specified tissue/cell compartment. Assignments use immediate function, repair capacity, persistent injury, anatomy and inflammatory effects.

The central registry distinguishes 39 bulk tissues and 68 cell populations, with source-linked rationale. A cell-level exception requires resolved tissue and cell identity. An unresolved or mixed label cannot earn a narrower allowance; reproductive/developmental membership is a separate anatomical axis.

The evidence records distinguish RNA, surface protein, peptide-HLA presentation, receptor recognition and tissue injury. Missing measurements and measured zeros have separate values.

Vital: Highest protection priority: plausible life-threatening failure, major irreversible neurological/sensory loss, or loss of indispensable regenerative support.

Engineered T-cell treatment caused fatal cardiac injury in the cited human study. AT2-cell depletion caused lethal pulmonary edema in the cited mouse study.

Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma; Targeted Type 2 Alveolar Cell Depletion: A Dynamic Functional Model for Lung Injury Repair

Protected: Persistent functional or structural injury is a major concern; compensation is incomplete or direct immune-depletion evidence is insufficient.

The cited studies document persistent salivary injury, ciliary dysfunction and impaired airway clearance.

Sicca Syndrome Associated with Immune Checkpoint Inhibitor Therapy; Lung function from school age to adulthood in primary ciliary dyskinesia; Parent stem cells can serve as niches for their daughter cells

Manageable: A defined compartment has credible replacement, compensation or supportive-treatment precedent; acute inflammation and persistent recognition still require review.

Clinical B-cell depletion is managed with infection monitoring and immunoglobulin replacement when needed. Thyroid hormone replaces thyroid endocrine function.

Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia; Immune-Related Thyroiditis with Immune Checkpoint Inhibitors

Disposable: Narrow conditional allowance for loss of a specified compartment; does not imply painless, reversible or inflammation-free ablation.

Gallbladder removal is a clinical example of tolerated organ loss. Melanocyte-targeted T cells caused eye and ear injury, so the cutaneous category is restricted to skin.

Treatment for Gallstones; Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen

Central tissue registry, compartment rationales and primary references

Evidence has multiple independent meanings

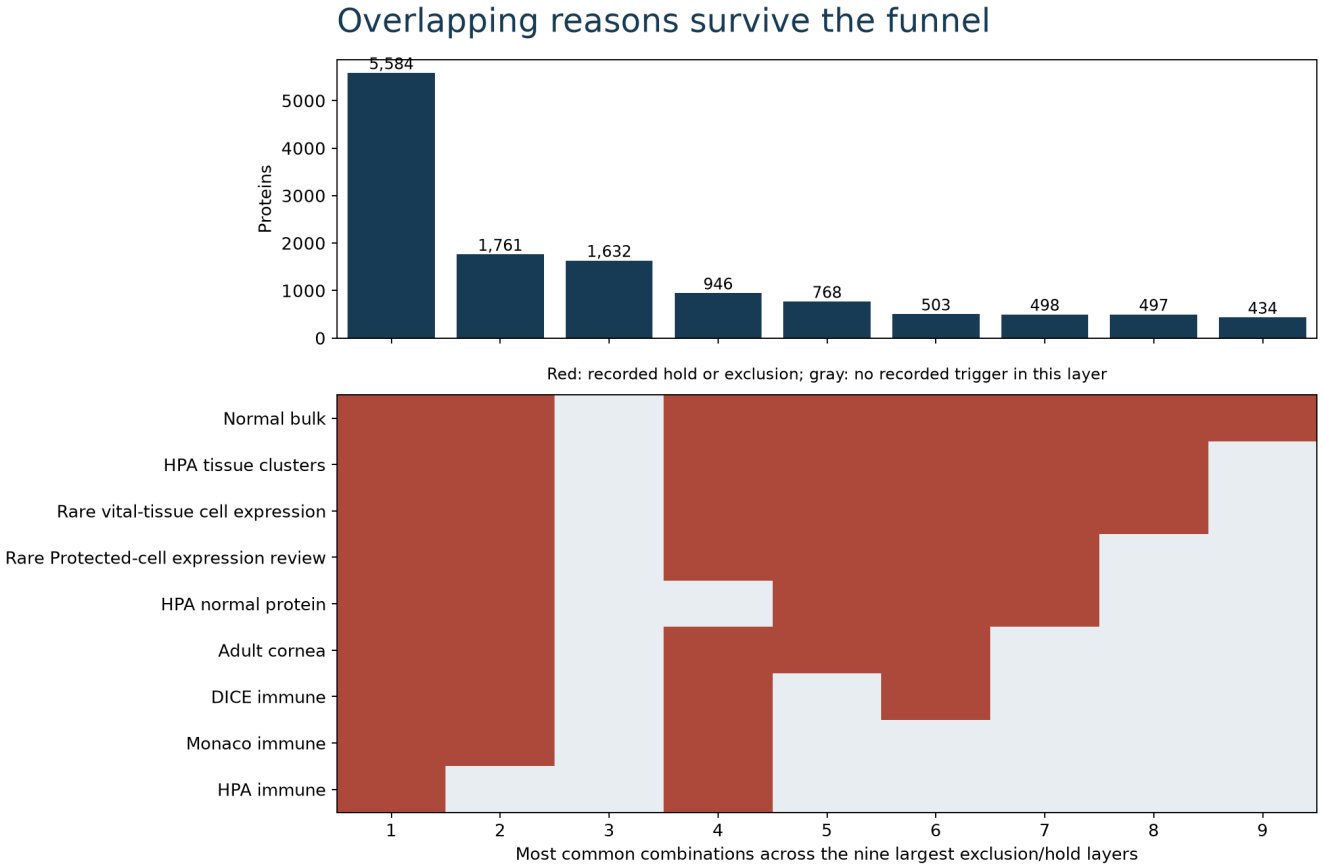
Checks complete (screened) means the release's required assessment is complete. It describes completion, so CTA, Unsafe and Other proteins can all have this status. No sequence exclusion (allow) means the selected protein has a resolved identity, completed assessment and no Unsafe assignment. This can apply to CTA or Other; CTA admission has additional requirements. These labels describe the released checks, not every tissue, isoform or later publication.

Review completion, interpretation and policy effect are separate. A completed review can retain an unresolved interpretation. A rejected protein claim remains visible with its rationale; rejecting that claim does not prove normal-tissue absence.

All 196 historical findings have linked current dispositions. MAGEC2, LGALS13 and PLCZ1 retain material holds. Each record lists the source, reagent, tissue, developmental stage and unresolved question.

Fatal neurological injury in the MAGE-A3-reactive TCR cohort remains product-specific evidence. The MAGEA12 and EPS8L2 explanations are both retained as unresolved. MAGEA12 remains Unsafe under an independent rare-cell finding. Neither receptor causality nor favorable clinical credit transfers across a gene family.

Morgan et al. 2013: neurological injury report · Martin et al. 2021: MAGEA12/EPS8L2 re-examination



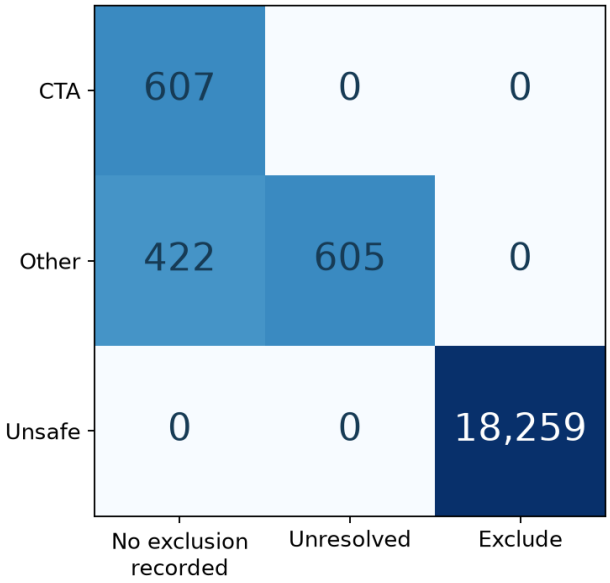
Translation, identity and downstream sequence use

Translation qualification asks whether the exact selected primary-locus record supports target sequence use. Complete coding annotation, start/stop codons, biotype and sequence integrity are assessed separately from normal-tissue hazard. Qualification applies to the selected primary-locus record.

The verified bundle retains original sequences, every source occurrence, qualification reasons and coverage gaps. It supplies contiguous AA20 segments with original coordinates and never joins across ambiguity, internal stops or different records. Selenocysteine-containing records require review under the AA20 sequence contract.

Across occurrences, exclude takes precedence over review, then allow. Unsafe sequence remains exclusion evidence even in an incomplete record. Allow means no exclusion from the assessed record under this policy. Exact matching does not validate an approximate similarity cutoff or patient-HLA decision. The historical Unsafe plus additional-forbidden assets retain their older policy identity.

Classification and sequence role answer different questions



Research parameters and coverage limits

Without usable matching protein staining, the upper RNA limits start at 4 for Vital, 8 for Protected and 16 for Manageable tissues, in the source's own TPM or nTPM units. Approved, Supported and Enhanced negative IHC respectively multiply the starting limit by 2, 4 and 8. Reviewed product-specific clinical evidence can add one doubling within the same cap. A value must be above the limit to trigger this RNA rule. Corneal CP10K uses separate limits. These screening thresholds are set by OncoRef policy.

Cancer presence requires one compatible tumor sample above zero or reviewed tumor protein. Reproductive RNA uses a 5-unit criterion with assay-specific support requirements; reliable reproductive IHC and reviewed protein provide separate routes. Adult/placental diffuseness above 10% can block identity, never establish Unsafe.

The rare-cell panel uses a 1% detection screen with cell-support rules. Detection frequency is not abundance. Coverage is limited by sampled organs, donor age, activation states, paralog mapping and assay resolution. Optional panels remain disclosed when missing. A 95% Unsafe 9-mer overlap blocks practical candidate admission, but is not a clinical cross-reactivity validation.

The complete gate table below retains the actual executable rules and source IDs. Human thresholds are not transferred to dog or mouse; their native nominations remain unassessed until supported by a species-specific policy.

Does CTA mean expressed only in cancer?

No. CTA admission requires cancer expression plus qualifying reproductive or developmental association. The CTA total is not a count of proteins absent from all normal tissues, including testis, placenta, embryo and fetus. Passing normal-tissue thresholds is not a claim of zero normal expression.

The current data do support an observed-expression count: 40 cancer-expressed records have complete required bulk coverage and zero stored adult/reproductive RNA summaries. Of these, 22 have developmental measurements and 0 also have zero developmental RNA; 18 lack developmental measurements. Stored zeros are not proof of absent protein, and tiny positive RNA values are not validated protein production.

The separate low-expression screen finds 66 preliminary matches. Independent normal-cell evidence excludes 6 of them. A numerical profile match is not a final cancer-only designation; remaining matches retain explicit review or coverage requirements. The inventory audit lists every record, counterevidence source and measurement gap.

Cancer-specific altered sequences and peptides require a separate analysis from these reference-protein counts. A mutation, fusion or alternative transcript can create a target sequence even when the source gene is normally expressed; evidence for one such sequence cannot establish exclusivity of every product of the gene.

Reproduction and version changes

Selection policy: oncoref.cta-selection.2026-11-08.1. Role policy: oncoref.sequence-roles.2026-11-04.1. Translation policy: oncoref.translation-qualification.2026-11-04.1.

The 1.29.0 comparison remains pinned. The evidence and translation PRs also retain adjacent baselines, separating changes in category, subtype, role, assessment status and reasons. Counts are recalculated from the evidence and rules.

Use the public decision, finding and sequence APIs. The final release receipt binds the decision exports, ontology, curation, qualification, bundle and generated figures. Complete and cached scientific replays must be byte-identical. Historical reports keep their original inputs and policies.

Release 1.35.0 applies MANE-first reference selection, coding-evidence review and recomputed protein identities. Its classifications and sequence roles use the current policy. The 602-protein literature analysis remains a labelled 1.31.0 snapshot.

Complete executable gate rules

Reproductive or developmental expression

Full Reproductive bulk ≥5 nTPM OR supported reproductive-cell RNA ≥5 nTPM (≥20 reads; indirect sites additionally need complete outside-cluster coverage and every outside value <5, with no curated normal-induction conflict) OR reliable mapped positive reproductive IHC for every member without an antibody conflict OR reviewed human reproductive protein covering every selected member OR supported developmental stage/lineage mean ≥5 TPM; ≥20 cells, ≥5 positive and ≥10% positive at ≥1 TPM. Presence evidence never overrides normal exclusions.

Sources: HPA v23 / GTEx consensus; HPA v23 normal-tissue IHC; HPA v23 tissue/cell clusters; Published human normal immune induction contexts; Tyser human gastrulation atlas; Petropoulos preimplantation atlas

Cancer expression

At least one measured compatible tumor sample >0 in native TPM or HPA pTPM OR reviewed reliable tumor protein. This establishes presence, not therapeutic abundance; do not add maxima from different samples.

Sources: OncoRef source-mapped tumor union; HPA v23 cancer sample measurements

HPA normal protein

Reliable non-Low positives in Vital/Protected supply Unsafe evidence; Manageable positives block CTA only; uncertain evidence remains unresolved. Low is flagged.

Sources: HPA v23 normal-tissue IHC

Required RNA coverage

Combined normal bulk (HPA consensus, HPA brain, GTEx v11) is the required RNA coverage panel: every nonreproductive tissue concept measured for every protein member. Missing coverage in the separate GTEx 52-context core, HPA tissue-cell, immune, DICE and cornea panels does not itself block admission; observed normal evidence from these panels still applies.

Sources: HPA v23 / GTEx consensus; HPA v23 human brain, 193 subregions; Direct GTEx v11 tissue medians; HPA v23 tissue/cell clusters; HPA v23 immune summaries; Monaco immune summaries via HPA v23

Reproductive concentration

HPA v23 average normal/(normal+Reproductive maximum) above 10% blocks the adult/placental route when its reference is positive and the comparison is complete. Use 38 contexts/29 anatomy units, including Disposable and excluding thymus; thymus remains Protected for hazard assessment. Supported embryonic evidence remains independent; its contrast is unassessed. This is not safety evidence.

Sources: HPA v23 / GTEx consensus

Translation qualification

CTA admission requires qualifying coding annotation or supported ncORF evidence and an approved reference protein: MANE Select first, Ensembl canonical second, with a complete primary-locus CDS and AA20 sequence. An incomplete designated transcript or unresolved coding conflict requires review; a shorter or alternate-locus protein cannot silently replace it. Every member must share the approved sequence. All candidates still undergo normal-evidence assessment; translation review does not establish Unsafe.

Sources: Ensembl 115 coding translations; Reviewed coding and representative-protein findings

Normal bulk

One inclusive tissue/IHC ceiling: above it in Vital/Protected supplies Unsafe evidence; above it in Manageable blocks CTA, and above 128 (the top Manageable ceiling) supplies Unsafe evidence. No second RNA boundary.

Sources: HPA v23 / GTEx consensus; HPA v23 human brain, 193 subregions; Direct GTEx v11 tissue medians

HPA tissue clusters

RNA requires ≥20 reads and uses the single ceiling. IHC credit requires an exact or declared compartment match; unrelated cells in the same tissue cannot inherit it.

Sources: HPA v23 tissue/cell clusters

HPA immune

Central cell-specific tiers with native nTPM limits and no matched negative-IHC credit; reproductive/placental location remains neutral.

Sources: HPA v23 immune summaries

Monaco immune

Central cell-specific tiers with native TPM limits and no matched negative-IHC credit.

Sources: Monaco immune summaries via HPA v23

DICE immune

Third-highest reported TPM value; central cell-specific tiers. Unpaired group distributions retain bounds. Missing coverage alone does not exclude.

Sources: DICE immune distributions

Adult cornea

Recurrent native CP10K: ≥5 positive cells and ≥20 counts in ≥2 sample groups. Nominal RNA limits /100; not a TPM conversion. Missing coverage alone does not exclude.

Sources: Collin adult corneal atlas

Accepted normal-protein evidence

Accepted positive protein findings exclude unless extracted Low intensity retains a flag; unquantified does not mean negative.

Sources: Reviewed normal protein and clinical exclusions

Reviewed normal-cell knowledge

Reviewed adult human normal protein, or localized RNA with orthogonal normal-cell function, in Vital/Protected supplies Unsafe evidence. Sparse RNA, uncertain species/stage or unresolved identity produces a review hold. Activation and rare-cell frequency do not exempt normal cells. Every finding retains its source and limitations.

Sources: Reviewed human gene knowledge: normal cells and reproductive protein

Clinical safety exclusion

Retain reviewed product-specific clinical injury counterevidence and causal-attribution limits; no admission shortcut from clinical use.

Sources: Reviewed normal protein and clinical exclusions

Rare vital-tissue cell expression

At least 1% detection in a centrally assigned Vital population in the existing brain/spinal-cord/heart/lung CELLxGENE panel supplies Unsafe evidence. At least 3,000 cells and 10 expressing cells are required. Excluded or unmeasured populations are coverage limits, not allowances. Detection frequency is distinct from abundance.

Sources: CELLxGENE rare-cell screen

Rare Protected-cell expression review

At least 1% detection in a centrally assigned Protected population in the existing CELLxGENE panel blocks CTA admission without itself establishing Unsafe. The same cell-support criteria apply. Source age restrictions and detection artifacts remain limitations.

Sources: CELLxGENE rare-cell screen

External validated normal protein

Use categorical/quantitative staining only within verified reagent, tissue and cell scope; no inherited HPA reliability label.

Sources: Curated external staining observations

Unsafe sequence containment

A would-be CTA with at least 95% of its 9-mers also present in Unsafe proteins becomes Other:Contained: Unsafe sequence is blacklisted, so almost no peptide survives. Computed over all protein-coding transcripts of Ensembl 115. This is a sequence rule, not expression evidence.

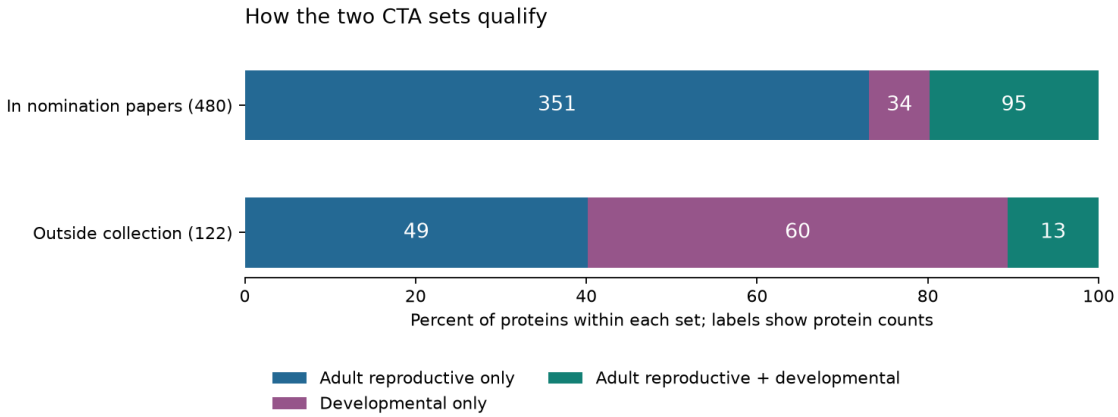
Sources: Ensembl 115 coding translations; Sequence overlap with released Unsafe proteins

CTA and Other assignments

CTA requires all applicable admission gates, required coverage and resolved normal evidence. Unsafe exclusion takes precedence. Other includes assessed noncandidates, held/incomplete evidence, nested microproteins and sequence-contained candidates. Subtype and assessment_status answer different questions; an identity failure cannot erase a normal-evidence hold. Sequence roles require that independent status.

Sources: All released decision gates

Historical 1.31.0 review: 122 CTAs outside the paper collection



The historical 602-CTA snapshot includes 480 proteins in the 15-paper nomination registry and 122 outside it. These are the counts for the original literature review, not the current selection. The 122 represent 123 loci. The review found additional cancer-target publications for TBXT, SEMG2 and PATE1.

Of the 122, 60 qualify through developmental evidence alone, 49 through adult reproductive evidence alone and 13 through both. Only 34 of the 480 registered proteins use the developmental-only route. The biological categories record their tissue and developmental associations separately.

Median highest tumor RNA is 35.6 versus 33.5 TPM. All 122 pass the reference-coding translation rule. Their released gene-knowledge ledger contains no findings. Tumor prevalence and experimental protein detection are separate measurements.

Complete 122-protein table, route comparison, linked primary-literature findings and follow-up priorities

All 122 proteins have expression profiles in this report. The literature review covers 20 proteins and 19 publications or source records. Findings are linked to the released classifications; structured gene-level curation is pending.

TBXT - Missing cancer-target literature

Brachyury was tested as a cancer-vaccine target in a human phase I trial. This publication is missing from the nomination collection.

Phase I Trial of a Yeast-Based Therapeutic Cancer Vaccine (GI-6301) Targeting the Transcription Factor Brachyury

Trial record

Human clinical vaccine study. The product was a yeast-based vaccine; the trial did not test engineered CAR or TCR cells.

SEMG2 - Missing explicit CTA literature

The authors identify SEMG2 as a cancer-testis antigen and report tumor RNA and cancer-cell-line protein measurements.

SEMG1/2 augment energy metabolism of tumor cells

Human tumor data and cancer cell lines. SEMG1 and SEMG2 results require separate attribution.

SEMG2 - Normal-tissue conflict requiring curation

SgII transcripts were reported in trachea and kidney. Airway basal cells stained with semenogelin-reactive antibodies.

Semenogelin I and II, the predominant human seminal plasma proteins, are also expressed in non-genital tissues

Human tissue RT-PCR and immunohistochemistry. The antibodies recognize both SgI and SgII; the basal-cell staining is not specific to SEMG2.

PATE1 - Missing cancer/reproductive literature

The original PATE study reported expression in prostate cancer, normal prostate and testis.

PATE, a gene expressed in prostate cancer, normal prostate, and testis, identified by a functional genomic approach

Human tissue expression and functional genomics. The historical PATE name must be mapped to PATE1 at the exact locus.

LEUTX, DPRX - Early embryo regulation

LEUTX and DPRX regulate an early-embryo transcriptional program. The paper also examines cancer-cell-line expression.

The human PRD-like homeobox gene LEUTX has a central role in embryo genome activation

Human embryos and human embryonic stem-cell overexpression. Experiments include human embryo transcripts and embryonic stem-cell overexpression.

DPRX, DUXA, DUXB, TPRX2 - Early embryo regulation

DPRX, DUXA, DUXB and TPRX2 are included in a study of human preimplantation homeobox factors.

Characterization and target genes of nine human PRD-like homeobox domain genes expressed exclusively in early embryos

Human embryo transcripts and human stem-cell functional assays. Measurements include embryo transcripts and stem-cell functional assays.

LEUTX, ZNF280A, TPRX2, DUXB - Early embryo regulation

LEUTX regulates ZNF280A and TPRX2. The study reports TPRX2 transcription and function despite its historical pseudogene annotation.

Comprehensive characterization of the embryonic factor LEUTX

Human cell models, proteomics and transcriptional assays. The downstream response was measured after experimentally induced LEUTX.

LEUTX - Fusion-specific cancer evidence

A pediatric thoracic tumor contained a YAP1::LEUTX fusion.

Whole genome profiling of rare pediatric thoracic tumors elucidates a YAP1::LEUTX fusion in an unclassified biphasic embryonal neoplasm

Human tumor case study. The reported molecule is the fusion; this release uses the full-length reference LEUTX sequence.

WEE2 - Oocyte and maternal-effect biology

WEE2 mutations caused fertilization failure in human oocytes. The study measured WEE2 protein and tested experimental rescue.

Homozygous Mutations in WEE2 Cause Fertilization Failure and Female Infertility

Human genetics, affected oocytes and rescue experiments. The evidence concerns human oocytes and fertilization. WEE2 enters this release through developmental RNA measurements.

KHDC3L - Oocyte and maternal-effect biology

KHDC3L/ECAT1 was identified by proteomics in human oocytes.

Identification of Maturation-Specific Proteins by Single-Cell Proteomics of Human Oocytes

Human oocyte proteomics. ECAT1 and C6orf221 are historical aliases that require exact locus mapping.

NLRP14 - Oocyte and maternal-effect biology

NLRP14 deficiency was associated with female infertility, oocyte maturation defects and early embryonic arrest.

NLRP14 deficiency causes female infertility with oocyte maturation defects and early embryonic arrest by impairing cytoplasmic UHRF1 abundance

Human infertility investigation and mechanistic experiments. The study combines human infertility observations with mechanistic model experiments.

PABPN1L - Oocyte and maternal-effect biology

Pabpn1l knockout disrupted the maternal-to-zygotic transition and early embryonic development in mice.

PABPN1L mediates cytoplasmic mRNA decay as a placeholder during the maternal-to-zygotic transition

Mouse oocytes and embryos. The functional experiments used mice. PABPN1L and PABPC1L are different genes.

PATE1, WFDC8 - Reproductive tract secretion

Human reproductive-tissue proteomics identified epididymal enrichment of PATE1 and WFDC8.

In-depth mapping of human testicular and epididymal proteins and their functional association with spermatozoa

Human reproductive tissues and proteins. The study distinguishes epididymal protein production from association with sperm.

RNASE9 - Reproductive tract secretion

RNASE9 was cloned and localized in human epididymis; the study also reported sperm-associated protein.

Cloning, expression and location of RNase9 in human epididymis

Human tissue RNA and immunolocalization. Measurements include tissue RNA and immunolocalization.

MRGPRX1 - Normal-tissue conflict requiring curation

MRGPRX1 RNA was detected in human dorsal-root ganglia and in 79% of 85 sampled human sensory neurons; the study also tested agonist responses.

Activation of pruritogenic TGR5, MrgprA3, and MrgprC11 on colon-innervating afferents induces visceral hypersensitivity

Mouse mechanisms plus human dorsal-root-ganglion expression and functional assays. The released rare-cell panel covers brain, spinal cord, heart and lung. It does not include dorsal-root ganglia.

SP7 - Normal-tissue conflict requiring curation

SP7 was detected in human primary osteoblast cultures, including adult-derived cultures.

Expression of alternatively spliced isoforms of human Sp7 in osteoblast-like cells

Human primary osteoblast cultures, including adult cells. The study measured cultured osteoblasts and splice variants; in-vivo adult tissue levels need extraction.

KLF14 - Normal somatic function requiring review

Human adipose studies connect KLF14 expression and function with adipocyte size, insulin resistance and type 2 diabetes risk.

Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition

Human adipose cohorts and primary cells, with mouse experiments. Adipose is Manageable in the tissue registry. The expression amount and cell compartment determine the policy result.

CALHM3 - Sensory-function follow-up

Calhm3 was required for ATP-mediated taste signaling in mouse experiments.

CALHM3 Is Essential for Rapid Ion Channel-Mediated Purinergic Neurotransmission of GPCR-Mediated Tastes

Mouse taste-cell genetics and heterologous channel experiments. The functional evidence is from mice and heterologous constructs. Adult human taste-cell localization remains uncurated.

BTN1A1 - Physiological-state coverage gap

The live HPA page reports high BTN1A1 staining in lactating breast glandular cells.

Human Protein Atlas: BTN1A1 staining in lactating breast

Current human tissue immunohistochemistry. The page is newer than the pinned HPA v23 export. Version, physiological state and antibody data need reconciliation.

Plain-language sequence decisions

Checks complete (screened): the release's required assessment is complete, regardless of whether it found an exclusion. No sequence exclusion (allow): the selected protein has a resolved identity, completed assessment and no Unsafe assignment. CTA and Other proteins can receive this role. Unresolved (review): one or more evidence questions remain. Exclude: retain the sequence as exclusion evidence. Exclude takes precedence over review and allow.

Missing measurements: Required normal tissues or member loci have missing observations.

Conflicting or uncertain normal evidence: A normal signal, reagent result or causal interpretation remains unresolved.

Uncertain identity: The gene, locus, transcript or source mapping is not resolved.

Uncertain translation: Coding support or translation completeness does not qualify.

Ambiguous sequence: Stops, ambiguous residues or non-AA20 characters prevent automatic exact-record use.

Not assessed: The record or isoform lacks an applicable decision; the selected representative cannot clear every isoform.

RNA-limit table and full decision explanations

Data-source guide

Measurement datasets, literature nominations, reviewed findings and derived rules are distinct. Blank source cells in earlier reports did not mean evidence-free decisions: rare-cell rules use CELLxGENE, and sequence containment uses Ensembl translations plus current Unsafe assignments.

Reviewed human gene knowledge: normal cells and reproductive protein

Measurement: source-native qualitative and quantitative observations. Scope: explicit per-finding stage and activation state. Exact gene/protein attribution; sparse RNA and unresolved species/isoforms remain review holds. Search leads are not accepted findings.

Durante et al. 2020 adult human olfactory atlas supplementary tables

Measurement: positive neurons; cluster detection fraction. Scope: adult. Exact reported genes enter knowledge review holds; sparse detection is not TPM, protein, donor replication or proof of absence for unlisted genes.

Verbeurgt et al. 2014 whole human olfactory mucosa

Measurement: native qPCR Cq; assay IDs and donor identity. Scope: adult; 26 donor samples. OR10A4 normal RNA supplies a reviewed hold. Native values are not TPM, protein or a clinical toxicity estimate; unlisted genes are not measured zero.

Verbeurgt et al. 2014 whole human olfactory mucosa

Measurement: estimated RNA copies per 20 ng total RNA. Scope: adult; 26 donor samples. OR10A4 normal RNA supplies a reviewed hold. Native values are not TPM, protein or a clinical toxicity estimate; unlisted genes are not measured zero.

Published human normal immune induction contexts

Measurement: qualitative reported RNA induction. Scope: adult; activated. Qualitative counterevidence to indirect reproductive-site restriction; no TPM or clinical-toxicity inference

HPA v23 / GTEx consensus

Measurement: nTPM. Scope: adult; placenta handled separately. Bulk source summaries; not exhaustive pediatric, pregnancy or rare-cell coverage

Tyser human gastrulation atlas

Measurement: TPM. Scope: Carnegie stage 7. One embryo; cell support does not establish independent-donor replication

Petropoulos preimplantation atlas

Measurement: TPM. Scope: 3–7 days, stage × inferred lineage. 1,519 measured cells; stage and embryo counts retained; not coverage of all fetal stages

HPA v23 human brain, 193 subregions

Measurement: nTPM. Scope: adult. Maximum subregion mean after within-context group sums; no automatic negative-IHC credit

Direct GTEx v11 tissue medians

Measurement: TPM. Scope: adult. 52 native whole-organ contexts; 14 microdissected and two cultured/transformed contexts remain outside this adapter. Exact whole-organ HPA negative-IHC grades apply; donor medians do not exclude donor outliers. Native reader preserves all 68 contexts.

OncoRef source-mapped tumor union

Measurement: TPM. Scope: cohort-specific. 132 compatible cohorts; source namespaces deduplicate views, not patients across studies

HPA v23 cancer sample measurements

Measurement: pTPM. Scope: cohort-specific. Protein-coding RNA denominator; no numerical protein or peptide abundance

HPA v23 tissue/cell clusters

Measurement: nTPM. Scope: source-labelled adult; reproductive locations explicit. At least 20 reads; missing cell types and disease states remain uncovered

HPA v23 immune summaries

Measurement: nTPM. Scope: adult. Sorted populations retain native units; no matched negative-IHC relaxation

Monaco immune summaries via HPA v23

Measurement: TPM. Scope: adult. Sorted populations retain native units; no matched negative-IHC relaxation

DICE immune distributions

Measurement: TPM. Scope: adult. 15 distributions; third-highest value; unpaired locus distributions retain bounds; optional coverage

Collin adult corneal atlas

Measurement: CP10K. Scope: adult. Two supported sample groups; optional coverage; CP10K is not length-normalized TPM

HPA v23 normal-tissue IHC

Measurement: HPA categorical level and reliability. Scope: adult. Reliability describes evidence support, not a calibrated false-negative probability

Curated external staining observations

Measurement: native categorical findings. Scope: source-specific. Reagent, tissue, cell type and validation scope remain explicit; no inherited HPA grade

Reviewed normal protein and clinical exclusions

Measurement: categorical reviewed findings. Scope: source-specific. Qualitative reported Low is distinct from HPA Low and numerical intensity; clinical evidence is product-specific

Letetresgene Autoleucel in Advanced/Metastatic Myxoid/Round Cell Liposarcoma

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Autologous T cell therapy for PRAME+ advanced solid tumors in HLA-A*02+ patients: a phase 1 trial

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Autologous T cell therapy for MAGE-A4+ solid cancers in HLA-A*02+ patients: a phase 1 trial

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

SPEARHEAD-1: a single-arm phase 2 trial of afamitresgene autoleucel in advanced synovial sarcoma and myxoid/round cell liposarcoma

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Identification of a Titin-Derived HLA-A1-Presented Peptide as a Cross-Reactive Target for Engineered MAGE A3-Directed T Cells

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Guillain-Barre syndrome observed with adoptive transfer of lymphocytes genetically engineered with an NY-ESO-1 reactive T-cell receptor

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

NCT03709706 protocol: NY-ESO-1/LAGE-1A shared peptide and c259 receptor

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

NCT03132922 protocol: MAGE-A4 c1032 target peptide

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

First-in-human dose escalation trial to evaluate the clinical safety and efficacy of an anti-MAGEA1 autologous TCR-transgenic T cell therapy in relapsed and refractory solid tumors

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Phase I clinical trial evaluating the safety and efficacy of ADP-A2M10 SPEAR T cells in patients with MAGE-A10+ advanced non-small cell lung cancer

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

NY-ESO-1-specific redirected T cells with endogenous TCR knockdown mediate tumor response and cytokine release syndrome

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Re-examination of MAGE-A3 as a T-cell Therapeutic Target

Measurement: reviewed clinical attribution. Scope: adult clinical cohort. Peptide/HLA/receptor scope; targeted review, not all trials; no species transfer

Ensembl 115 coding translations

Measurement: annotation / amino-acid sequence. Scope: source-specific. MANE Select precedes Ensembl canonical; the exact primary-locus CDS must qualify and conflicting coding evidence must be resolved. Reference selection does not prove tumor protein production.

Reviewed coding and representative-protein findings

Measurement: exact-reference, source-linked review. Scope: finding-specific. Gene or sequence scope, experimental setting and unresolved conflicts are retained.

CELLxGENE rare-cell screen

Measurement: % cells with detected RNA. Scope: source-specific. Existing brain/spinal-cord, heart and lung panel only; >=3,000 cells and >=10 positive cells. >=1% detection is a policy trigger, not an abundance measurement or proof of receptor recognition.

Sequence overlap with released Unsafe proteins

Measurement: % shared 9-mers. Scope: source-specific. Computed from Ensembl 115 coding translations and current Unsafe assignments. >=95% containment blocks CTA admission; it does not itself establish tissue expression.

All released decision gates

Measurement: classification / reasons. Scope: source-specific. Derived result, not an independent experimental source. Unsafe takes precedence over CTA admission; Other retains its evidence status.

Worked decisions and primary sources

PRAME - CTA

Checks complete; sequence: No sequence exclusion. Reasons: All declared gates pass.

Translation: Qualified selected primary-locus record.

unresolved / Context only: The historical blanket primary-stain hold is superseded by the current HPA level/reliability and compartment rules, including prior reagent eligibility. Original raw staining and unresolved specificity remain visible; this policy revision is not new validation of the reagent or evidence of absence. No second historical IHC veto is applied.

HPA v23 primary IHC; repeated family profiles are not independent studies

All mapped positive rows, profile notes and reagent summaries reviewed; no independent histology re-read or new experiment. Reliable gene-profile labels do not validate each raw stain. RNA/protein donors are unpaired. Retain observed intensities and identity ambiguity; the decision is a conservative research hold.

accepted / Context only: Reported Low intensity remains a flag without an automatic protein veto under the current policy; retain the observation and source-specific limitations.

33890816; 35973038

Primary 2021 abstract; 2022 normal-tissue methods/results inspected via primary article search excerpts. Full 2021 methods unavailable; abstract-level conservative acceptance, not proven therapeutic injury. IMA203 evidence concerns a selected peptide/HLA/TCR product and is preserved separately.

MAGEC2 - Other:Held

Normal evidence unresolved; sequence: Unresolved. Reasons: Reviewed normal-cell knowledge

Translation: Qualified selected primary-locus record.

unresolved / Unresolved: Material normal-tissue finding remains unresolved; block CTA admission and require sequence review without establishing Unsafe. Negative bulk profiles do not resolve assay identity.

HCA587 antigen expression in normal tissues and cancers: correlation with tumor differentiation in hepatocellular carcinoma

Primary abstract rechecked; publisher full text unavailable in this review. Original tissue and antibody controls remain unverified.. Abstract does not settle reagent specificity; no demonstrated target-mediated clinical injury. A neural protein conflict warrants a hold, but the historical antibody result does not establish target-mediated clinical injury or resolve MAGE-family cross-reactivity.

LGALS13 - Other:Held

Normal evidence unresolved; sequence: Unresolved. Reasons: Reviewed normal-cell knowledge

Translation: Qualified selected primary-locus record.

unresolved / Unresolved: Material normal-tissue finding remains unresolved; block CTA admission and require sequence review without establishing Unsafe. Negative bulk profiles do not resolve assay identity.

Isolation and sequence analysis of a cDNA encoding human placental tissue protein 13 (PP13)

Primary abstracts rechecked (PMIDs 10527825 and 19497882); original adult/fetal specimen labels, bands and reagent controls not independently re-inspected.. Adult versus fetal samples and original protein bands require full-paper adjudication. Conflicting assay identity is unresolved; neither a proven unsafe classification nor unconditional inclusion is justified by these abstracts alone.

PLCZ1 - Other:Held

Normal evidence unresolved; sequence: Unresolved. Reasons: Reviewed normal-cell knowledge

Translation: Qualified selected primary-locus record.

rejected / Context only: Preserve rejection of the PLCZ1 protein interpretation: the immunostaining measured CFTR. The separately retained qualitative pancreatic RNA finding remains unresolved.

Rescuing defective pancreatic secretion in cystic fibrosis by suppression of a newly identified gene

Primary full text: Methods, normal-tissue RT-PCR/Figure 1, and CFTR immunolocalisation. Original isoform mapping and quantitative expression remain unresolved.. Human pancreatic RNA is qualitative, not nTPM. Disease/cell-model functional experiments and mouse measurements do not establish healthy human protein abundance.

unresolved / Unresolved: Material normal-tissue finding remains unresolved; block CTA admission and require sequence review without establishing Unsafe. Negative bulk profiles do not resolve assay identity.

Rescuing defective pancreatic secretion in cystic fibrosis by suppression of a newly identified gene

Primary full text: Methods, normal-tissue RT-PCR/Figure 1, and CFTR immunolocalisation. Original isoform mapping and quantitative expression remain unresolved.. Human pancreatic RNA is qualitative, not nTPM. Disease/cell-model functional experiments and mouse measurements do not establish healthy human protein abundance.

MAGEA12 - Unsafe

Checks complete; sequence: Exclude. Reasons: Rare vital-tissue cell expression

Translation: Qualified selected primary-locus record.

SERPINA2 - Other:Translation

Translation unresolved; sequence: Unresolved. Reasons: Translation qualification

Translation: designated translation incomplete | incomplete annotation tag | incomplete start | incomplete stop | noncanonical start residue | nonqualifying transcript biotype

MROH5 - Unsafe

Checks complete; sequence: Exclude. Reasons: Translation qualification; Rare vital-tissue cell expression

Translation: ambiguous or interrupted sequence | designated translation incomplete | nonqualifying transcript biotype

SELENOV - Other:Translation

Translation unresolved; sequence: Unresolved. Reasons: Translation qualification

Translation: ambiguous or interrupted sequence | designated translation incomplete

GDPD4 - Other:Translation

Translation unresolved; sequence: Unresolved. Reasons: Translation qualification

Translation: unresolved translation evidence

SLCO1B3-SLCO1B7 - Other:Translation

Translation unresolved; sequence: Unresolved. Reasons: Translation qualification

Translation: unresolved translation evidence

FRG2 - Other:Translation

Translation unresolved; sequence: Unresolved. Reasons: Translation qualification

Translation: unresolved translation evidence

FRG2B - Other:Translation

Translation unresolved; sequence: Unresolved. Reasons: Translation qualification

Translation: unresolved translation evidence