CURRENT SELECTION

Human and dog CTA evidence

602 human CTA proteins

646 loci / 18,255 Unsafe / 1,006 Other

Every human protein classified. Dog selection remains unassessed.

Sources, expression profiles, single-cutoff filtering and measured tissue evidence.

Safety, cancer and Reproductive evidence

Assessment	Question	Current rule
Safety	Concerning normal expression?	Absolute RNA / IHC and clinical rules
Cancer expression	Expressed in a measured cancer?	One sample >0 TPM/pTPM or reviewed tumor protein
Reproductive evidence	Expression in a qualifying branch?	Bulk/cell RNA ≥5, reliable reproductive protein, or embryonic RNA
Coverage	Are required panels measured?	Combined HPA bulk / brain and GTEx for every locus

Only safety establishes Unsafe. Coverage and normal review holds remain explicit in assessment_status.

Diffuse expression is an identity question

Term	Meaning
N	HPA v23 normal-tissue nTPM
R	Highest Reproductive nTPM in HPA v23 (adult/placental)
N / (N + R)	Average across 29 organs / 38 contexts; thymus excluded
Active boundary	Average >10%, with a supported positive R and complete comparison

MTCP1: Unsafe (Rare vital-tissue cell expression); diffuseness 24.53%. Embryonic concentration remains unassessed.

Required and optional RNA coverage

Requirement	Interpretation
Combined normal bulk	HPA consensus / brain and GTEx; measured zero counts
All member loci measured	Missing identities are not replaced by paralogs
Missing required coverage	Blocks CTA; assessment status records the gap
Cell / immune / cornea panels	Missing coverage allowed; measured evidence counts

The fixed GTEx core is an inspectable profile, not a separate admission gate. Rare-cell coverage remains incomplete.

ZNF730: panel composition matters

Calculation	HPA v23	GTEx v11
Reproductive reference (testis)	4.9 nTPM	4.91414 TPM
Full source panel	9.53%: 29 anatomy units	1.65%: 23 anatomy units
Same 20 anatomy units	6.36%	1.71%

Thymus is excluded from both sides of the concentration score

Thymus exclusion changes HPA 10.34% → 9.53%. Independent safety evidence keeps ZNF730 Unsafe.

Published candidate lists

Paper	Proteins	CTA	Unsafe	Other	Unassigned
da Silva et al. (2017)	1,017	411	466	140	0
Wang et al. (2016)	988	269	683	36	0
Chang et al. (2019)	985	281	676	28	0
Gong et al. (2021)	729	346	343	40	0
Bruggeman et al. (2018)	692	12	679	1	0
Jamin et al. (2021)	469	33	429	7	0
Loriot et al. (2025)	202	83	113	6	0
Carter et al. (2023)	103	53	47	3	0
Hirbo et al. (2015)	78	8	67	3	0
Seager et al. (2024)	15	10	3	2	0
Sasaroli et al. (2011)	13	0	13	0	0
Garg et al. (2007)	5	3	2	0	0

Union: 2,514 proteins / 2,570 loci. 480 selected proteins have an annotation; 122 do not.

Broader published nominations

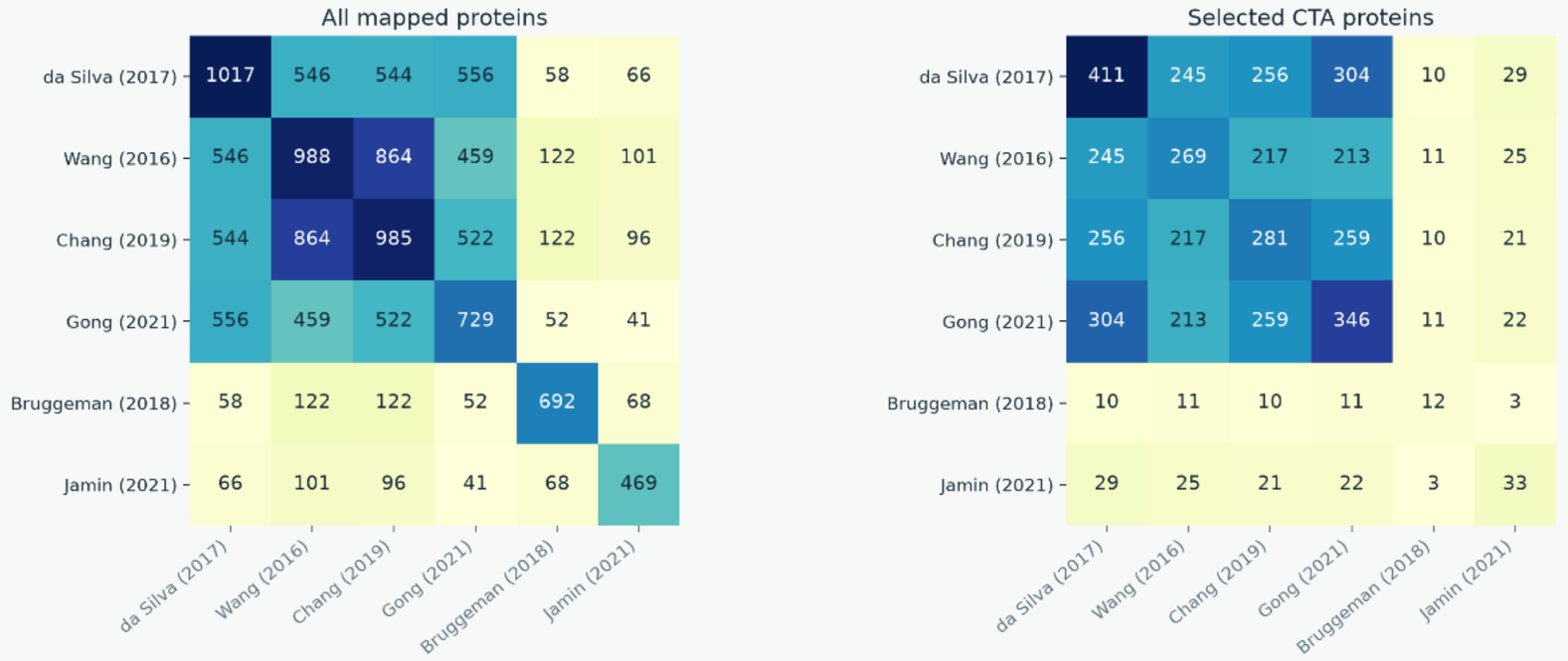
Paper	Published scope	Identity accounting
Sasaroli et al. (2011)	13 detailed tumor-vascular candidates	13 matched genes
Hirbo et al. (2015)	86 placenta-enriched genes, including 30 fast-evolving	82 matched, 4 unresolved
Compact source view	Bai, Boël and Bradley covered by larger sets	All memberships retained
Garg et al. (2007)	Still uniquely contributes SPAG9	Retained in main view

Vascular and placental nominations retain their scope. Paper membership never bypasses the quantitative funnel.

Overlap between published lists

Overlap between published lists

The six largest mapped paper sets; full pairwise tables are downloadable



Diagonal cells give each paper's set size. Off-diagonal cells count shared proteins, not independent validation studies.
Citations and original source-table IDs remain linked in the source inventory.

Reproductive evidence: measured branches

Qualifying route	Entrants	Final CTA
Adult/placental evidence only	4,686	400
Embryonic-only evidence	459	92
Both measured routes	12,546	110

RNA and reliable reproductive IHC establish association

Entrants count positive expression before the other admission gates. Reproductive presence and profile concentration are separate questions.

Normal and cancer measurement sources

Evidence	Source	Quantity
Normal / reproductive bulk	HPA v23 consensus / brain + GTEx v11	nTPM / TPM
Normal tissue cells	HPA v23 clusters, at least 20 reads	nTPM
Immune cells	HPA v23 / Monaco / original DICE	Native nTPM / TPM
Adult cornea	Collin et al., GSE155683	CP10K
Normal protein	HPA v23 + curated external IHC	Staining / reagent confidence
Cancer	132 compatible OncoRef cohort views + HPA	TPM / protein-coding TPM
Developmental	Tyser gastrulation + Petropoulos preimplantation	Stage / lineage mean TPM

DICE and corneal coverage are optional. Measured counterevidence still contributes.

CURRENT POLICY

RNA cutoffs

Ceiling	Missing / unusable	Approved –	Supported –	Enhanced –
Vital	4	8	16	32
Protected	8	16	32	64
Manageable	16	32	64	128

One boundary: above-cutoff Vital / Protected is Unsafe

Differentiated bowel cells and fibroblasts use Protected. Crypt progenitors use Vital. Equality passes.

Normal-tissue hazard ontology

Tier	Examples	Key interpretation
Vital	AT1/AT2, club/basal, crypt progenitors	Function and regenerative reserve
Protected	Ciliated/goblet, fibroblasts, skin	Persistent injury and uncertain control
Manageable	Thyroid, mature B cells	Defined compensation precedent
Disposable	Cutaneous pigment, hair follicles	Narrow anatomical allowance

Bulk retains organ concern. Recognition and reproductive membership are separate. Protected rare-cell findings create a review hold.

Clinical evidence adjusts RNA ceilings

With clinical credit	Missing / unusable	Approved –	Supported –	Enhanced –
Vital	8	16	32	32
Protected	16	32	64	64
Manageable	32	64	128	128

IHC grades stay unchanged. Trials do not stack.

Protein and clinical-injury exclusions remain. This is protein-profile credit with tested epitope limitations, not clinical clearance.

Serious events and antigen specificity

Targeted CTA epitope	Important event	Specificity attribution
MAGE-A3 / a3a	Two fatal cardiac injuries	Titin cross-recognition implicated
MAGE-A3/A9 / NCI receptor	Three neurological cases; two deaths	MAGE-A12 / EPS8L2 attribution unresolved
NY-ESO-1 / LAGE-1A, c259	Late cardiac death; two GBS cases	Peptide-specific mechanism not established
NY-ESO-1 / TBI-1301	Grade 3 lung injury	Unresolved; no credit for this product
MAGE-A4 / afami-cel	Aplasia and stroke deaths in phase 1	No demonstrated antigen-specific mechanism
MAGE-A10 / ADP-A2M10	Grade 4 CRS; fatal aplasia after repeat treatment	Conditioning favored for marrow injury

Clinical trial and SAE review. Seriousness differs from severity; this is a selected summary, not a pooled event rate.

Other records an unmet admission gate

Reason	Meaning	Proteins
Nonreproductive	No qualifying reproductive evidence	570
Translation	Selected primary translation requires review	362
Nested	Nested short reference-CDS product	28
Held	Target gates pass; normal review hold remains	27
Contained	At least 95% of 9-mers occur in Unsafe proteins	15
Diffuse	Reproductive concentration fails	4

Normal RNA level, tumor support and Disposable expression remain separate attributes.

Evidence attributes stay independent

Question	Fields / interpretation
How high in cancer?	Native TPM / HPA pTPM maxima; two identified tumors >=10 TPM is a support flag
How high in adult normal?	Separate nTPM / TPM / CP10K summaries; raw unsupported observations remain visible
Adult Disposable pattern?	Permitted adult expression with low measured adult expression elsewhere
Developmental expression?	Separate lineage RNA and measured-stage coverage; never changes adult tissue permission

cta_selection.profiles() exposes these fields. Low / Weak / Normal are no longer competing classes.

Reproductive cell evidence

Route	Positive evidence	Additional requirements
Identified lineage	≥5 nTPM with ≥20 reads	Mapped reproductive lineage; all protein loci measured
Indirect site	Generic/immune cell in reproductive tissue	Complete outside cluster panel; every value <5 nTPM; no normal-induction conflict
Examples	CAGE1 208.9; PRDM9 61.8; FOXR2 5.2 nTPM	Spermatids / spermatocytes / spermatogonia
Counterexample	IFNA2 and IFNA14 in breast macrophages	Reviewed normal immune protein → Unsafe

Indirect support identifies the sample site, not the producing cell. Normal and bulk-concentration gates still apply.

Reviewed normal-cell knowledge

Evidence	Consequence
Established normal production	Vital/Protected protein or localized RNA plus orthogonal cell function can establish Unsafe
Unresolved normal findings	Sparse RNA, uncertain species/stage or isoforms remain review holds
Reproductive protein	Supplies presence below the RNA floor; safety and concentration still apply
Search and curation	All 611 baseline Other genes searched; accepted findings remain separate from search leads

Rare normal cells and activation states count. Expression exclusions do not assert clinical injury.

Historical findings retain explicit dispositions

Case	Current result	Reason retained
MAGEC2	Other:Held	Purkinje-cell reagent specificity remains unresolved
LGALS13	Other:Held	Adult/fetal and galectin identity remain unresolved
PLCZ1	Other:Held	Pancreatic RNA is unresolved; CFTR staining is not PLCZ1 protein
MAGEA12	Unsafe	Independent rare-cell exclusion; clinical peptide attribution remains unresolved

A completed review can retain uncertainty. Unresolved evidence requires review without establishing Unsafe.

Every candidate receives a classification

Candidate records	Count	Current treatment
Previously classified	19,604	All previous classifications unchanged
Restored: Other:Translation	247	Translation admission fails; sequence review required
Restored: Unsafe	12	Normal exclusion takes precedence
Complete inventory	19,863	All records enter the same assessment

Translation is a CTA admission gate. Insufficient support cannot hide normal-tissue evidence.

Expression-screen matches and counterevidence

Numerical screen	Count	Interpretation
Low normal expression profile	66	Preliminary matches; every independent exclusion still applies
Normal-cell counterevidence	6	Unsafe despite a routine-panel profile match
Stored adult/reproductive RNA = 0	40	Cancer present and required normal bulk coverage complete
Also developmental RNA = 0	0	18 adult-zero records lack developmental measurements

Stored RNA zeros do not establish protein absence. CTA counts are not cancer-only protein counts.

Selected translations require qualification

Record	Current classification	Selected translation
SERPINA2	Other:Translation	Incomplete primary LoF record; alternate locus cannot rescue it
MROH5	Unsafe	Interrupted primary record; normal exclusion still takes precedence
SELENOV	Other:Translation	Selenocysteine requires review under the AA20 contract

Qualification reviews the selected reference sequence. Original interrupted and alternate records remain available.

Sequence roles preserve every occurrence

Role	Source records	Meaning
Exclude	230,648	Applicable normal exclusion takes precedence
Review	13,512	Unresolved decision, isoform or translation
Allow	1,375	Qualified assessed representative under this policy

245,535 original records and all source occurrences survive. AA20 segments retain coordinates and never bridge ambiguity.

High tumor expression does not establish CTA identity

Protein	Cancer TPM	Adult atlas nTPM	Current reasoning
DBX1	297.5	0.8; thymic T cells	No qualifying reproductive evidence
OR2B2	246.3	0.5; pons	No qualifying reproductive evidence
KRTAP20-2	147.6	1.4; skin	Skin expression; reproductive identity unestablished
OR10A4	138.0	1.9; basal ganglia	Normal nasal RNA requires review

Disposable describes allowed adult tissues. Developmental expression is a separate measurement.

Other accumulates as identity gates fail

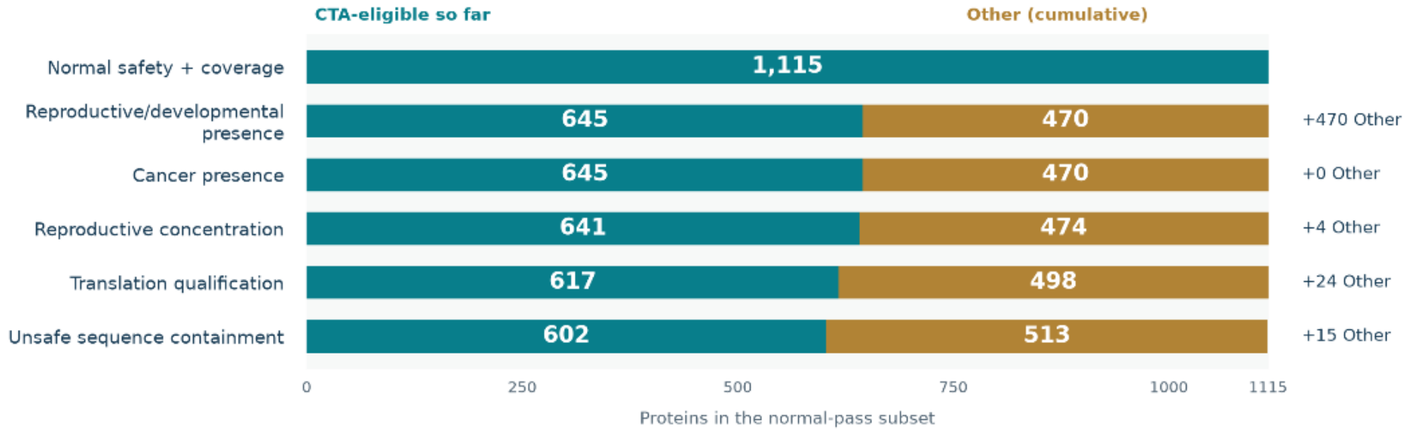
Where proteins become CTA or Other

OncoRef 1.31.0 | 19,863 representative proteins | Normal assessment first

602	18,255	513	493
CTA	Unsafe	Other	Unresolved
		passed the normal screen, failed a target gate	normal screen unresolved; also Other, not rejected

Follow the 1,115 proteins that pass normal safety and coverage

18,255 Unsafe and 493 Other proteins with an unresolved normal screen stay outside this track; none is unassessed.



470 lack reproductive evidence / are nested; 4 fail concentration; 24 need translation review; 15 have Unsafe sequence containment.

Cancer presence moves 0 additional proteins into Other; 645 remain eligible.

Rare vital-tissue cell expression (at least 1% in a named Vital population in the existing brain/spinal-cord/heart/lung panel) is the only Unsafe evidence for 121 proteins.

Reproductive includes placental and embryonic/fetal contexts. Unresolved proteins pass no CTA gate but are not Unsafe; Other:Held is the subset passing every target gate. The figure changes display order only; all final categories exactly match the released classification.

Every filtering layer can contribute



Final: 602 CTA / 18,255 Unsafe / 1,006 Other.

Removals at each filter: 1 of 2

Layer	First removals	Eligible	Unsafe
Association	2,175	17,688	0
Cancer expression	8	17,680	0
HPA normal protein	12,044	5,636	10,161
Required RNA coverage	102	5,534	10,161
Reproductive concentration	1,007	4,527	10,161
Translation qualification	118	4,409	10,161
Normal bulk	3,434	975	17,639
HPA tissue clusters	232	743	18,004
HPA immune	12	731	18,031
Monaco immune	11	720	18,052

Sequential removals overlap. Final categories use every applicable evidence layer.

Removals at each filter: 2 of 2

Layer	First removals	Eligible	Unsafe
DICE immune	15	705	18,096
Adult cornea	2	703	18,101
Accepted normal-protein evidence	3	700	18,103
Reviewed normal-cell knowledge	3	697	18,134
Clinical safety exclusion	0	697	18,134
Rare vital-tissue cell expression	74	623	18,255
Rare Protected-cell expression review	6	617	18,255
External validated normal protein	0	617	18,255
Unsafe sequence containment	15	602	18,255
CTA and Other assignments	0	602	18,255

Sequential removals overlap. Final categories use every applicable evidence layer.

IHC matching to measured RNA compartments

Context	Evidence and current ceiling
PRAME colon undifferentiated cells	20.3 nTPM / 96 reads; Enhanced-negative glandular IHC; Vital 32
PRAME adrenal bulk	3.6 nTPM; incomplete IHC coverage; Vital 4 → 8 with IMA203 credit
CTAG2 bronchial basal cells	24.9 nTPM / 58 reads; Enhanced-negative respiratory IHC; Vital 32
Unrelated cells in the same tissue	No negative-IHC credit without exact or declared compartment coverage

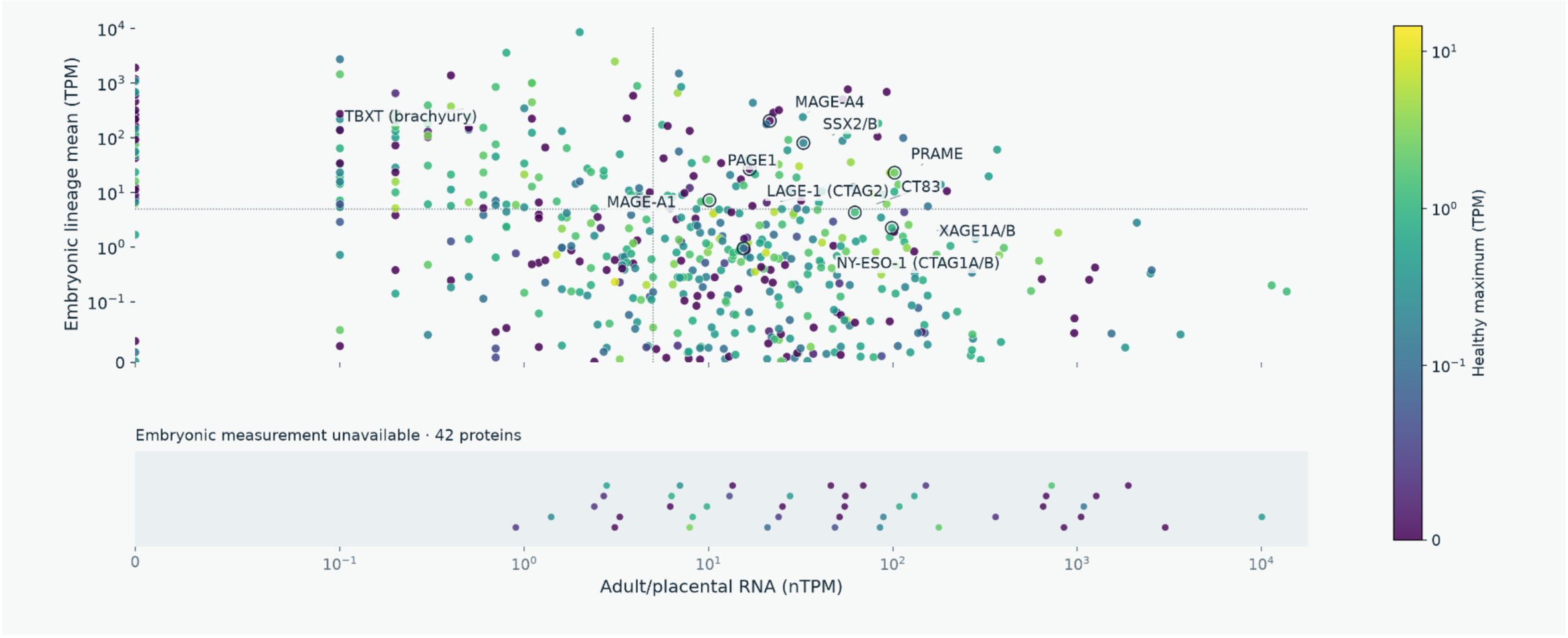
PRAME and CTAG2 pass this policy. Parent-compartment sensitivity and PRAME's Low adrenal protein remain explicit limitations.

Selected CTA: adult/placental versus embryonic RNA



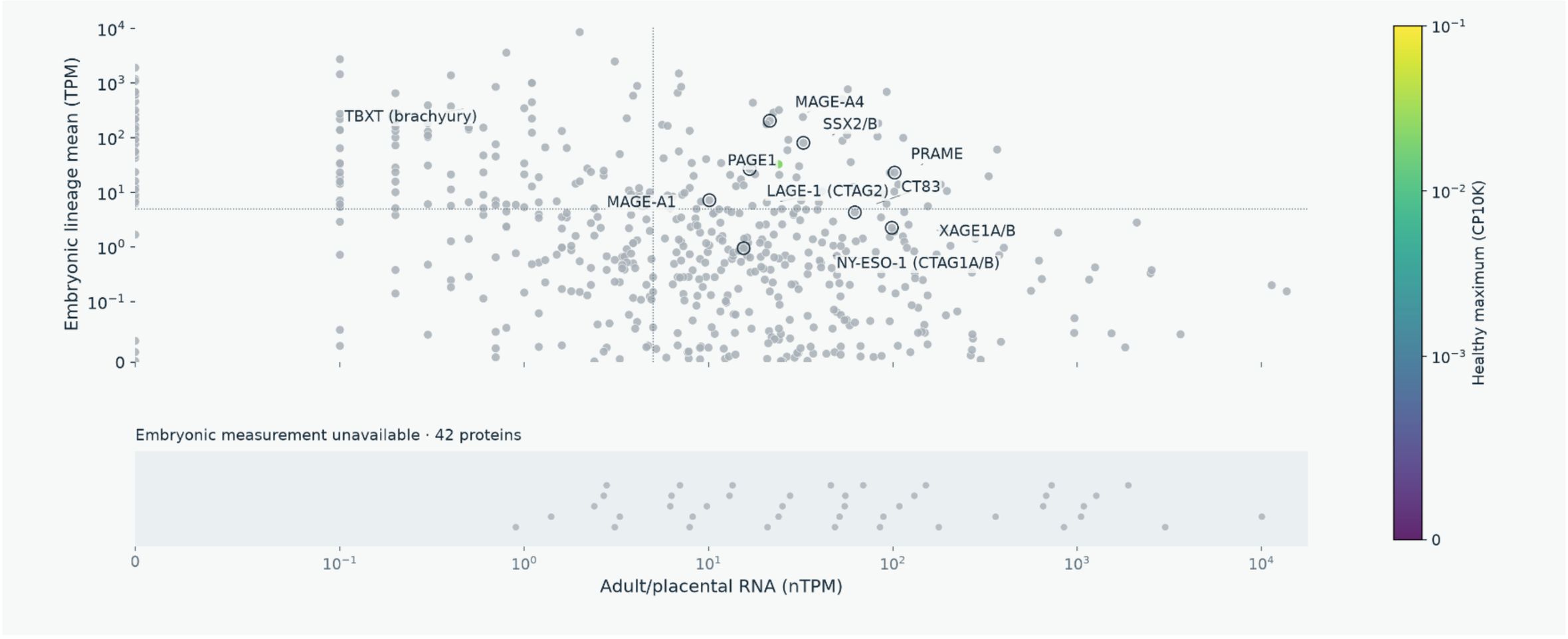
Gray: healthy measurement unavailable. The lower band has no embryonic-expression value.

Selected CTA: healthy RNA in TPM



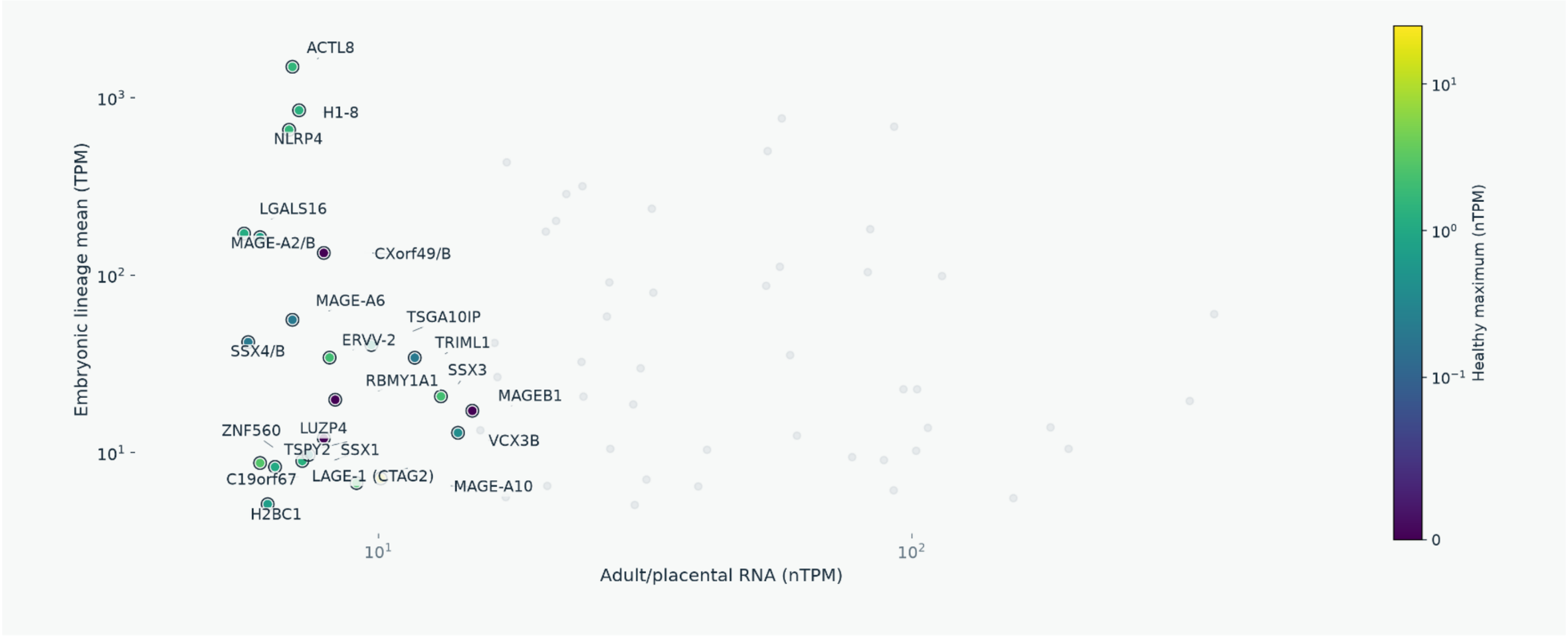
Gray: healthy measurement unavailable. The lower band has no embryonic-expression value.

Selected CTA: normal corneal RNA

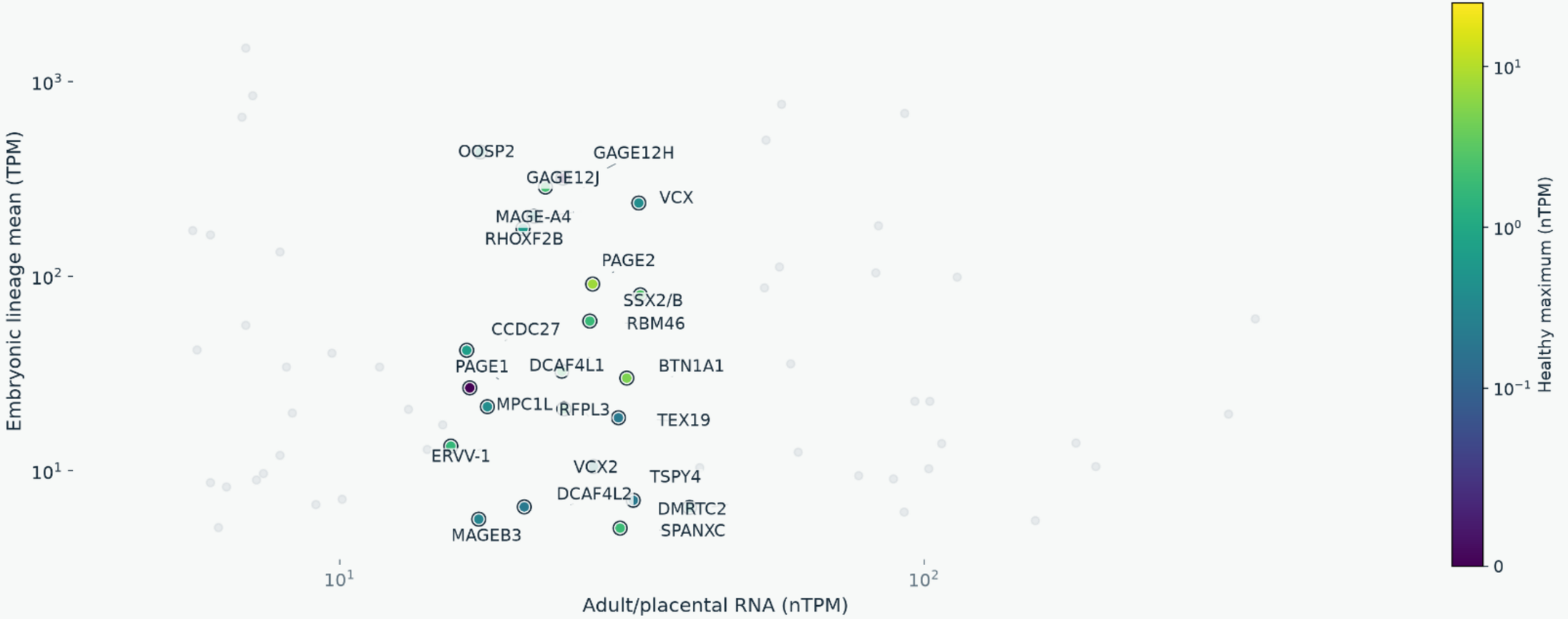


Gray: healthy measurement unavailable. The lower band has no embryonic-expression value.

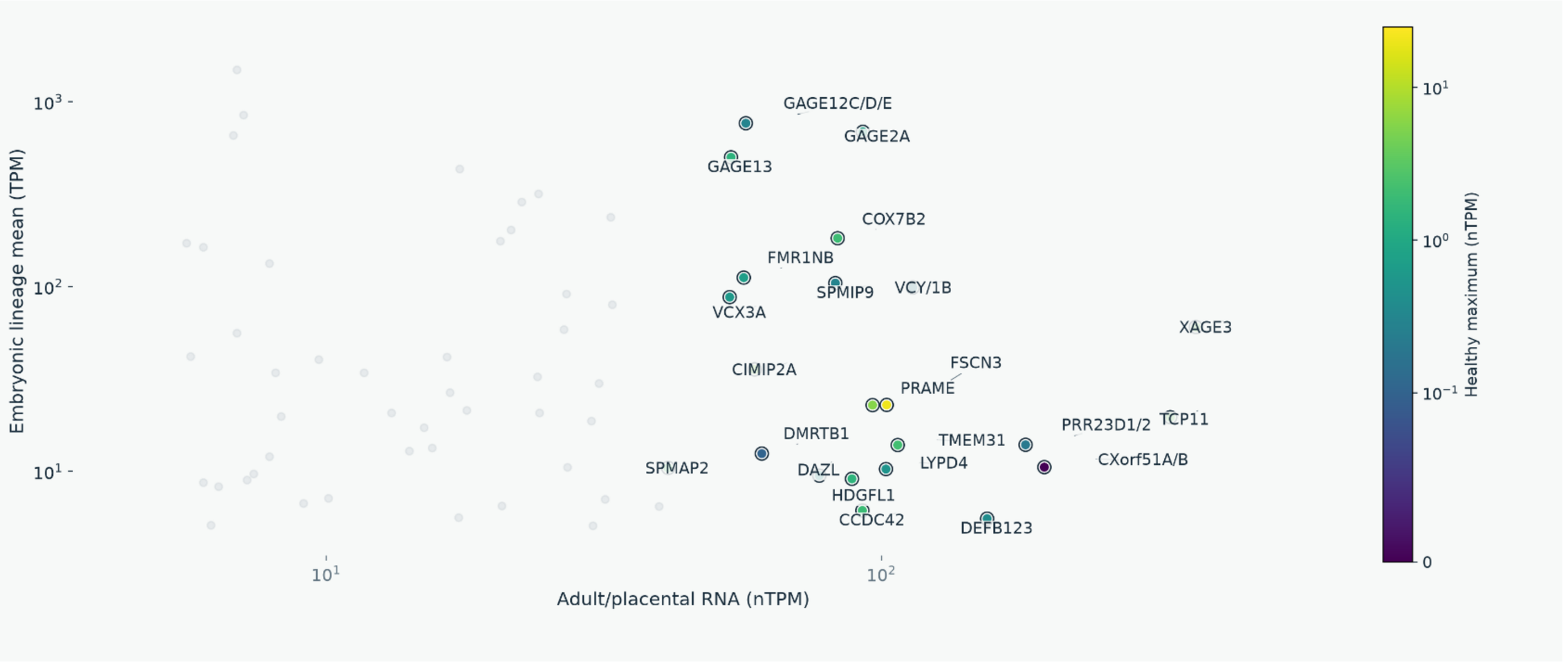
High adult/placental and embryonic RNA: labels 1 of 3



High adult/placental and embryonic RNA: labels 2 of 3



High adult/placental and embryonic RNA: labels 3 of 3



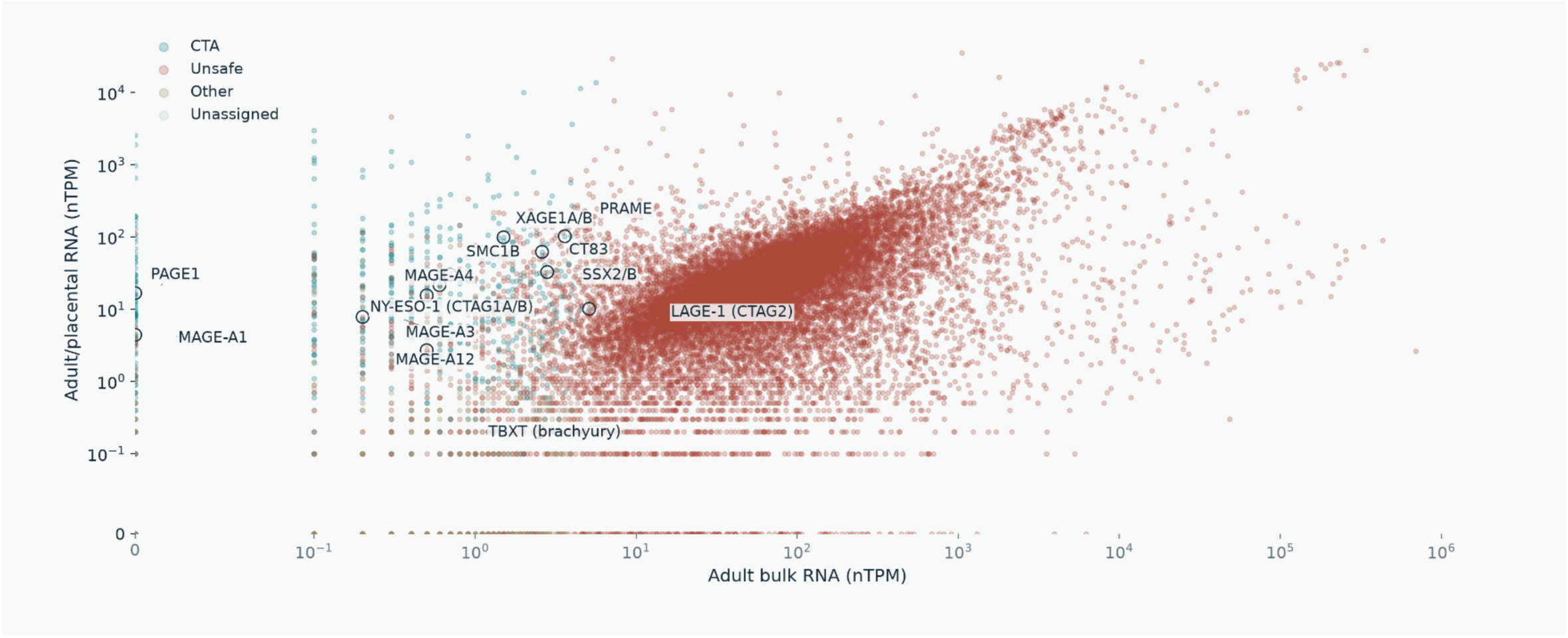
Complementary embryonic measurements

Source	Measured cells / embryos	Developmental scope
Tyser et al. (2021)	1,195 cells / 1 embryo	Gastrulation, Carnegie stage 7
Petropoulos et al. (2016)	1,519 cells / 88 embryos	Preimplantation, days 3–7
Recovered coverage	683 of 1,006 earlier missing proteins	Measured TPM, including familiar CTAs

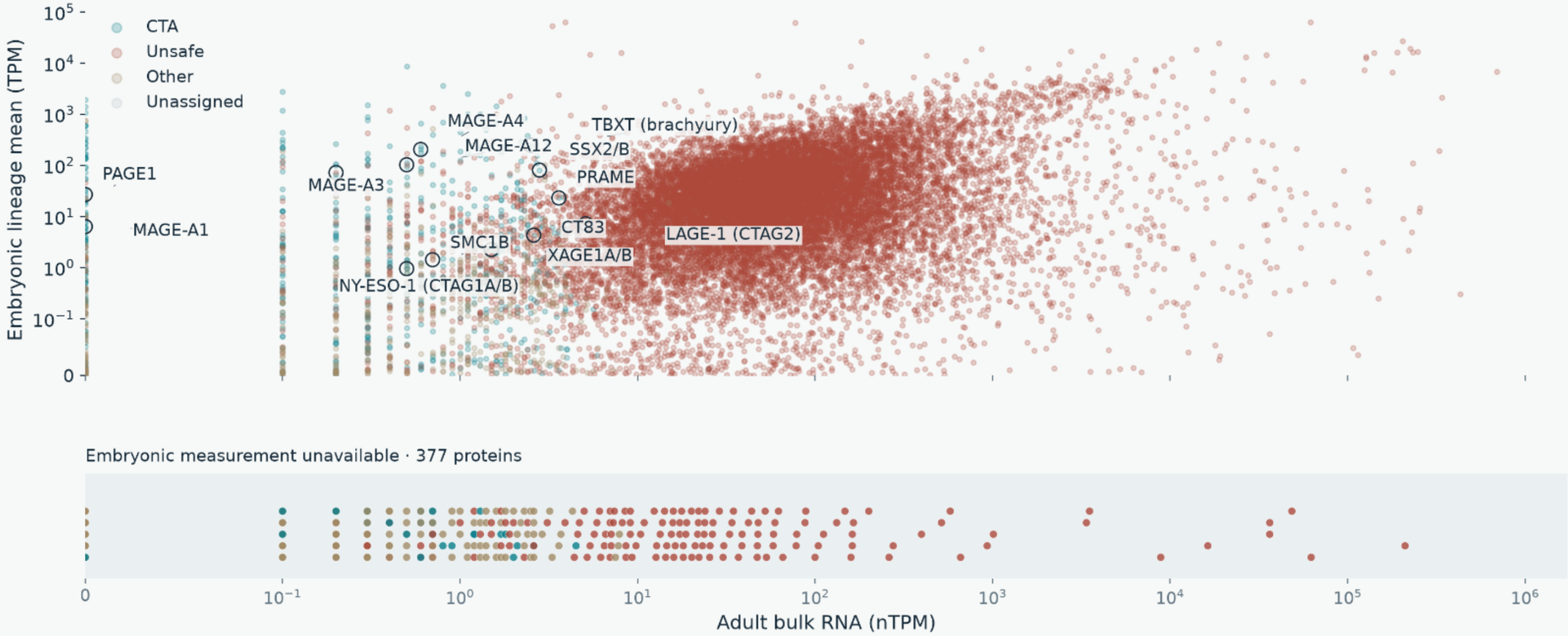
PRAME, MAGE-A4, CTAG1A/B, XAGE1A/B and CT83 now have TPM rows

Each source keeps its own stages. Neither atlas covers all fetal or pediatric development.

Widely studied antigens: adult/placental RNA



Widely studied antigens: embryonic RNA



Assessment status is separate from category

API status	Meaning	Human proteins
screened	Classification with a resolved applicable screen	19,277
incomplete_evidence	Classified with disclosed evidence gaps	105
normal_evidence_hold	Required normal coverage passes; a normal-screen hold remains	118
translation_review	Selected primary translation needs qualification	363
unassessed	No species-specific selection policy applied	0

Other includes review holds and sequence rules. No human protein has a null category.

Which sources have incomplete evidence

Source or unresolved evidence	Proteins
HPA v23 consensus bulk	88
Combined normal bulk: HPA consensus + HPA brain + direct GTEx	85
HPA v23 tissue-cell clusters	69
HPA v23 immune cells	45
Monaco immune cells via HPA v23	45
GTEx v11 core panel (52 contexts; informational, not required)	58
Combined Tyser + Petropoulos TPM	57
No tumor measurement: OncoRef + HPA	28
Measured tumor, incomplete member coverage	1
HPA v23 Uncertain normal-tissue IHC	20

Overlapping gaps within 105 proteins with incomplete evidence. DICE and corneal coverage are optional.

Dog evidence and integration gaps

Layer	Available evidence / remaining gap
Nominations	5 + 10 – 1 shared = 14 native nominations
Identity	13 unique mappings; one ambiguous MAGE-family assay
Normal / developmental RNA	Bgee TPM and BarkBase TMM-CPM need reference reconciliation
Tumor RNA	GSE304066 count/FPKM and reference conflicts remain unresolved
Normal protein	Limited study IHC; Pfizer pooled iBAQ is not cell-resolved IHC
Selection	All 14 unassigned / unassessed; no canine quantitative policy

Human thresholds, antibody grades and normal-tissue absence never transfer by orthology.

Curated oncofetal sources

Source	Scope and integration status
Tyser + Petropoulos	Current gastrulation and preimplantation TPM sources
Shaw et al. (2024)	157 cancer-specific targets; six oncofetal annotations. Exon-specific scope retained
TRIM71 liver study (2024)	Liver oncofetal signature; requires identity and measurement ingestion
Cardoso-Moreira et al. (2019)	Developmental organ trajectories with child/adult comparisons; not integrated

Published nominations need independent normal/tumor filtering and postnatal coverage.

Common API with native evidence underneath

```
results = cta_selection.results(species="human")
audit = cta_selection.audit(species="human")
audit.sources    audit.overlaps    audit.funnel
audit.explain("PRAME")
```

Field	Meaning
category	CTA / Unsafe / Other; classification gives the named category
assessment_status	Screened, evidence/translation review, or unassessed
classification / subtype	One categorical label; Other subtype always populated

Source-specific native units, tissue/cell context, confidence and coverage remain inspectable.

Every source resolves to a concrete version

```
catalog.ensure("gtex_tissue_rna")              # v11
catalog.ensure("gtex_tissue_rna", version="v10")
catalog.provenance("gtex_tissue_rna", "v11")
```

Shared choice	Source-specific meaning
Default / latest	Newest registered release; saved runs keep explicit pins
Concrete version	GTEx v11; HPA v25.1; a study snapshot or OncoRef build
Provenance	Version, content checksum, native unit, annotation and pipeline
Species	The same selector applies to native dog and mouse assets

This analysis pins GTEx v11 and HPA v23 normal panels. A newer public default does not rewrite it.