

CURE Student Assessment Materials and Evaluation Rubrics

Course: Animal Development (upper-level undergraduate) | CURE: Testing the Ortholog Conjecture: Evolution of Gene Expression During Mammalian Embryonic Development

Purpose. Instructor-ready sample assignment instructions, scoring rubrics, and pre/post assessment items for a computational course-based undergraduate research experience (CURE). These are proposed materials and may be adapted to available datasets, class size, and student experience.

1. Assessment framework

Student learning is evaluated through conceptual understanding, reproducible computational practice, research design, interpretation, collaboration, and communication. Research progress is evaluated through usable datasets, validated homology annotations, completed comparisons, reproducible code, and well-supported conclusions—not whether a favored hypothesis is confirmed.

Assessment / evidence	Timing	Weight
Pre-CURE diagnostic (ungraded baseline)	Weeks 1–2	0%
Research question and hypothesis proposal	Weeks 3–4	15%
Data and annotation audit + computational notebook	Weeks 5–8	20%
Comparative analysis and reproducibility package	Weeks 9–12	25%
Final research report	Weeks 13–15	25%
Group presentation and individual reflection	Weeks 14–16	15%
Post-CURE assessment (compare with baseline)	Final week	0%

Weights are suggested and sum to 100%. The instructor should score individual contributions separately when group submissions are used.

2. Student research assignment: instructions

Research question and scope

Investigate whether orthologous genes show greater conservation of embryonic expression than paralogous genes, and whether gene duplication history is associated with expression divergence. Select one developmentally relevant gene family and formulate at least two contrasting, testable predictions. Expression similarity is evidence about transcriptional conservation, not direct proof of conserved function, natural selection, or neofunctionalization.

Required steps and student deliverables

1. Proposal: Define gene family, biological relevance, ortholog/paralog comparisons, developmental context, hypotheses, expected patterns, and a feasible analysis plan (1–2 pages).
2. Dataset selection: Retrieve public embryonic expression data from suitable sources (e.g., NCBI GEO, EMBL-EBI Single Cell Expression Atlas, EmExplorer, DevOmics, dbEmbryo, TEDD, as applicable). Record accession, species, tissue/cell type, stage, assay, sample size, processing status, and license/attribution.

3. Homology annotations: Obtain Ensembl orthology and paralogy assignments, record annotation release, identifiers, gene family membership, and any one-to-many or many-to-many relationships. Explain exclusions and ambiguities.
4. Computational analysis: Follow the instructor-provided Behura Lab workflow as appropriate, documenting software versions, commands, input files, parameters, normalization and matching decisions, and error checks. Extend or adapt steps for cross-species ortholog comparisons when needed.
5. Results: Generate a gene-family summary, an expression heatmap or stage profile, a similarity/divergence comparison, and at least one suitable statistical analysis. Report effect sizes and uncertainty when feasible; discuss developmental-stage and cell-type confounding.
6. Communication: Submit a reproducibility package, 4–6-page report, 8–10-minute group presentation, and individual contribution/reflection statement.

Minimum submission checklist

Research proposal; data provenance table; Ensembl annotation table; README; runnable scripts or notebook; environment/software versions; input/output directory manifest; at least two publication-quality figures; statistical summary; final report; presentation; individual reflection.

3. Scoring rubric: research proposal (15 points)

Criterion	Strong (full credit)	Partial / needs improvement	Points
Biological question	Specific, developmentally relevant and researchable	Question broad or weakly linked to development	0-4
Hypotheses	At least two falsifiable predictions with alternatives	Predictions unclear or not testable	0-4
Comparison design	Valid ortholog/paralog groups, species and stage plan	Insufficient control of comparison structure	0-4
Feasibility and clarity	Data sources and analysis steps realistic and clear	Missing data or workflow detail	0-3

4. Scoring rubric: computational notebook and data audit (20 points)

Criterion	Evidence expected	Points
Data provenance	Accessions, species, stages, tissues/cell types, sample sizes and source citations	0-5
Homology annotation quality	Ensembl release, stable identifiers, ortholog/paralog classification, ambiguity handling	0-5
Computational execution	Correct commands, documented preprocessing and meaningful quality checks	0-5
Reproducibility	README, versioned scripts, parameters, file manifest and rerun instructions	0-5

5. Scoring rubric: comparative analysis and reproducibility (25 points)

Criterion	Evidence expected	Points
Appropriate comparisons	Biologically justified ortholog/paralog and stage/cell-context contrasts	0-6
Quantitative methods	Appropriate normalization, similarity measures, statistical comparisons and controls	0-7
Figures and tables	Readable, labeled heatmaps/profiles, homology summary and statistical outputs	0-5
Reproducible outputs	Analysis reruns from documented inputs or a justified reproducible subset	0-5
Caveats	Identifies confounding, missing data and interpretation limits	0-2

6. Scoring rubric: final research report (25 points)

Criterion	Evidence expected	Points
Introduction and hypotheses	Accurate conceptual framework and clearly stated predictions	0-5
Methods	Sufficient detail on data, Ensembl annotation and analytical choices	0-5
Results	Quantitative, well-labeled figures and statistical findings	0-6
Interpretation	Evaluates ortholog conjecture and duplicate-gene divergence without overclaiming	0-6

Criterion	Evidence expected	Points
Scientific writing	Clear structure, citations, limitations and next-step proposal	0-3

7. Scoring rubric: presentation and reflection (15 points)

Criterion	Evidence expected	Points
Scientific presentation	Clear question, approach, findings and legible visuals	0-6
Evidence-based discussion	Answers questions, acknowledges uncertainty and alternatives	0-4
Collaboration	Specific and equitable documented contributions	0-3
Individual reflection	Explains skills gained, challenges and next research question	0-2

8. Pre- and post-CURE assessment (same questions at both time points)

Instructions: Administer in weeks 1–2 and again during the final week. Allow 25–35 minutes. Use responses to measure change in learning, not to grade students on their starting knowledge. Score using the guide below; keep student identifiers consistent across administrations.

1. Definitions (4 points)

Define ortholog and paralog. Explain how speciation and gene duplication give rise to these relationships.

2. Conceptual prediction (4 points)

State the ortholog conjecture as a testable prediction about expression similarity, and describe one possible outcome that would challenge it.

3. Experimental design (4 points)

You have expression data for two mammalian species at several embryonic stages. Describe one valid ortholog-versus-paralog comparison and two confounders you would control.

4. Annotation reasoning (4 points)

Why should Ensembl annotation version and one-to-many orthology assignments be documented?

5. Computational reproducibility (4 points)

List four items that another researcher needs to rerun your analysis from public data.

6. Interpretation (4 points)

A paralog pair has divergent expression across embryonic cell types. What can be concluded, and what cannot be concluded about neofunctionalization or natural selection?

7. Quantitative reasoning (4 points)

Describe a suitable measure of expression-profile similarity and explain how you would compare its distribution between ortholog and paralog pairs.

8. Research communication (4 points)

Write a two-sentence conclusion for a hypothetical result in which ortholog expression similarity is higher on average, but varies substantially by embryonic stage.

Scoring guide (32 points total)

For each item: 4 = accurate, specific, complete and appropriately qualified; 3 = substantially correct with minor omission; 2 = partial understanding; 1 = minimal relevant content; 0 = absent or incorrect. Report mean pre/post score, paired score change, and item-level patterns. If class size permits, summarize uncertainty; avoid claiming a causal effect without a comparison group.

9. Research progress tracking template

Milestone	Verification / evidence	Status
Research question approved	Hypotheses and study design reviewed	Not started / In progress / Complete
Datasets located and audited	Accession/provenance table and quality notes	Not started / In progress / Complete
Homology assignments validated	Ensembl release and gene-pair table	Not started / In progress / Complete

Milestone	Verification / evidence	Status
Pipeline executed	Versioned code, logs and outputs	Not started / In progress / Complete
Analyses interpreted	Figures, statistics and alternative explanations	Not started / In progress / Complete
Research products shared	Report, presentation and reproducibility package	Not started / In progress / Complete

Instructor checkpoint questions: Can a second group reproduce the principal result? Are stage and cell-context comparisons defensible? Were missing data and ambiguous gene relationships reported? Do conclusions stay within the evidence? What new testable question emerged?

10. References and instructional resources

Islam M, Behura SK. Role of paralogs in the sex-bias transcriptional and metabolic regulation of the brain-placental axis in mice. *Placenta*. 2024;145:143–150. doi:10.1016/j.placenta.2023.12.019.

Behura Lab. Pipeline to model evolution of paralog expression. GitHub: https://github.com/sbehura392/FUNCTIONAL_GENOMICS_SCRIPTS_BEHURALAB/blob/main/Pipeline_to_model_evolution_of_paralog_expression.txt

Ensembl: <https://www.ensembl.org/> | NCBI GEO: <https://www.ncbi.nlm.nih.gov/geo/> | EMBL-EBI Single Cell Expression Atlas: <https://www.ebi.ac.uk/gxa/sc/>