

RECORDED MODEL BEHAVIOR

Resolve the dose dependence of AP-1 transcription

4/9 criteria passed

GPT-6-ASTRA · SAVED SOLVER AND DETERMINISTIC GRADING

Source review is ready. Exact saved scorer and submission are preserved; packet and replay checks do not prove absence of all scientific defects.

The selected publication contains the original saved solver attempt and deterministic grading. No written reasoning is graded.

FAILED CRITERIA IN THE SAVED GRADE

C-002 - Equal-dose mean distance between treated and vehicle AP-1 compositions, in percentage points: Shared deterministic runtime: incorrect_value (mean_composition_distance)

C-006 - Total AP-1 signal in the 10 nM condition divided by vehicle: ratio of the condition means of summed nuclear-coding-pool-relative RNA signal: Shared deterministic runtime: incorrect_value (pool_ratio_10nm)

C-007 - Total AP-1 signal in the 100 nM condition divided by vehicle: ratio of the condition means of summed nuclear-coding-pool-relative RNA signal: Shared deterministic runtime: incorrect_value (pool_ratio_100nm)

C-008 - Total AP-1 signal in the 1000 nM condition divided by vehicle: ratio of the condition means of summed nuclear-coding-pool-relative RNA signal: Shared deterministic runtime: incorrect_value (pool_ratio_1000nm)

C-009 - Descriptive fraction of library-level total AP-1 signal variation explained by the five condition means, including vehicle: Shared deterministic runtime: incorrect_value (pool_dose_eta2)

Read runs/gpt-6-astra/evaluation/outputs/submission/, deterministic-results.json, judgments.json and attempts/0001/ for the unedited answer and source checks.