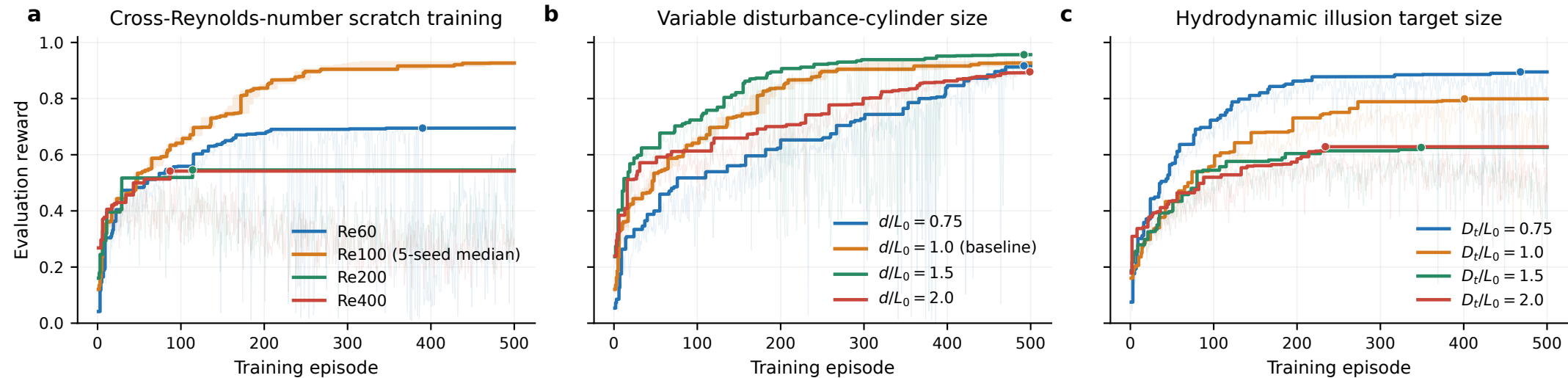


Best-so-far policy discovery across the expanded control matrix



Stair-step curves are incumbent best checkpoints; faint lines show exploratory evaluations. Re100 reports five-seed median and IQR.