

GIAB TR

Adam English

Benchmark v6.26 quick check

- Received benchmarking results from
 - Hipstr (Helia)
 - Nistv4 (Justin / Nate0)
 - Verkko (Justin / Nate0)
 - Medakka (Sean)
- Answering one question today: How often do they all disagree with Tier1?

Regions where all programs have same state

| hipstr | nistv4 | verkko | medaka | region count | % of Tier1 |
|--------|--------|--------|--------|--------------|------------|
| TN | TN | TN | TN | 944,904 | 57.67% |
| - | TN | TN | TN | 542,241 | 33.09% |
| TP | TP | TP | TP | 34,283 | 2.09% |
| - | TP | TP | TP | 15,920 | 0.97% |
| FP | FP | FP | FP | 202 | 0.01% |
| - | FP | FP | FP | 71 | 0.00% |
| - | FN | FN | FN | 63 | 0.00% |
| FN | FN | FN | FN | 38 | 0.00% |
| - | FN,FP | FN,FP | FN,FP | 9 | 0.00% |
| FN,FP | FN,FP | FN,FP | FN,FP | 8 | 0.00% |

Hipstr doesn't evaluate every region

All programs have same state in **93.85%** of regions

Every program disagrees with the benchmark in **391 (0.02%)** Tier1 regions

Replicate State of 391 'all disagree'

| Eichler | Li | Align | Count | % of 'all disagree' |
|-----------|-------|-------|-------|---------------------|
| TN | TN | TN | 77 | 19.69% |
| TP | TP | TP | 74 | 18.93% |
| FP | TN | TN | 74 | 18.93% |
| TN | FP | TN | 72 | 18.41% |
| TP | FN | TP | 36 | 9.21% |
| TP | FN,FP | TP | 21 | 5.37% |
| FN | TP | TP | 19 | 4.86% |
| FN,FP | TP | TP | 16 | 4.09% |
| uncovered | TN | TN | 2 | 0.51% |

38.62% of 'all disagree' have 3x agreement across replicates
The rest have 2x agreement across replicates

| repl_state | FN,FP_TP_TP | FN_TP_TP | FP_TN_TN | TN_FP_TN | TN_TN_TN | TP_FN,FP_TP | TP_FN_TP | TP_TP_TP | _TN_TN |
|------------------|-------------|----------|----------|----------|----------|-------------|----------|----------|--------|
| evaluation_state | | | | | | | | | |
| FN | 3 | 17 | 0 | 0 | 1 | 5 | 32 | 43 | 0 |
| FN,FP | 5 | 0 | 0 | 0 | 0 | 7 | 3 | 2 | 0 |
| FP | 8 | 2 | 74 | 72 | 76 | 9 | 1 | 29 | 2 |

Eval FP -> 146 have 1 replicate agreeing

Eval FN -> 49 have 1 replicate agreeing

Eval FN,FP -> 12 have 1 replicate agreeing

Boundary issues -> 50 have positive/negative flipping

~257 / 391 (65.7%) I suspect as possibly having assembly or comparison problems.

And 119 are where Replicates Agree.