

Supplementary information for: Multiple network alignment via multiMAGNA++

Vipin Vijayan and Tijana Milenković *

Department of Computer Science and Engineering, ECK Institute for Global Health,
and Interdisciplinary Center for Network Science and Applications (iCeNSA)
University of Notre Dame, Notre Dame, IN 46556, USA

SUPPLEMENTARY FIGURES

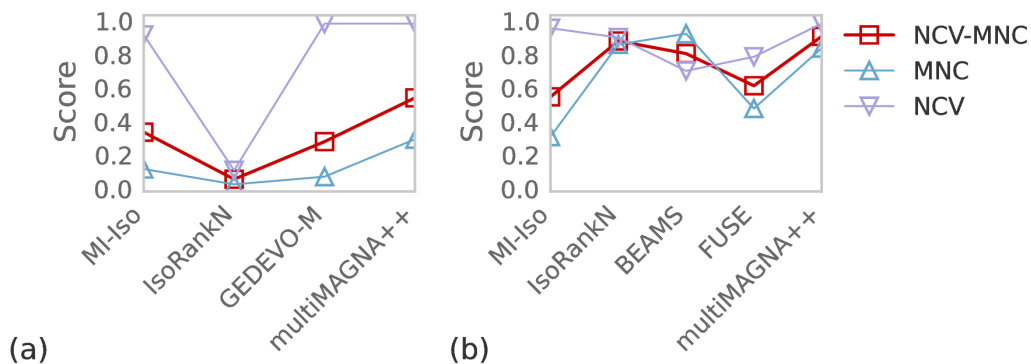


Figure S1: NCV-MNC, NCV, and MNC for the Yeast+%LC network set for (a) topology-only alignments and (b) topology+sequence alignments.

*To whom correspondence should be addressed

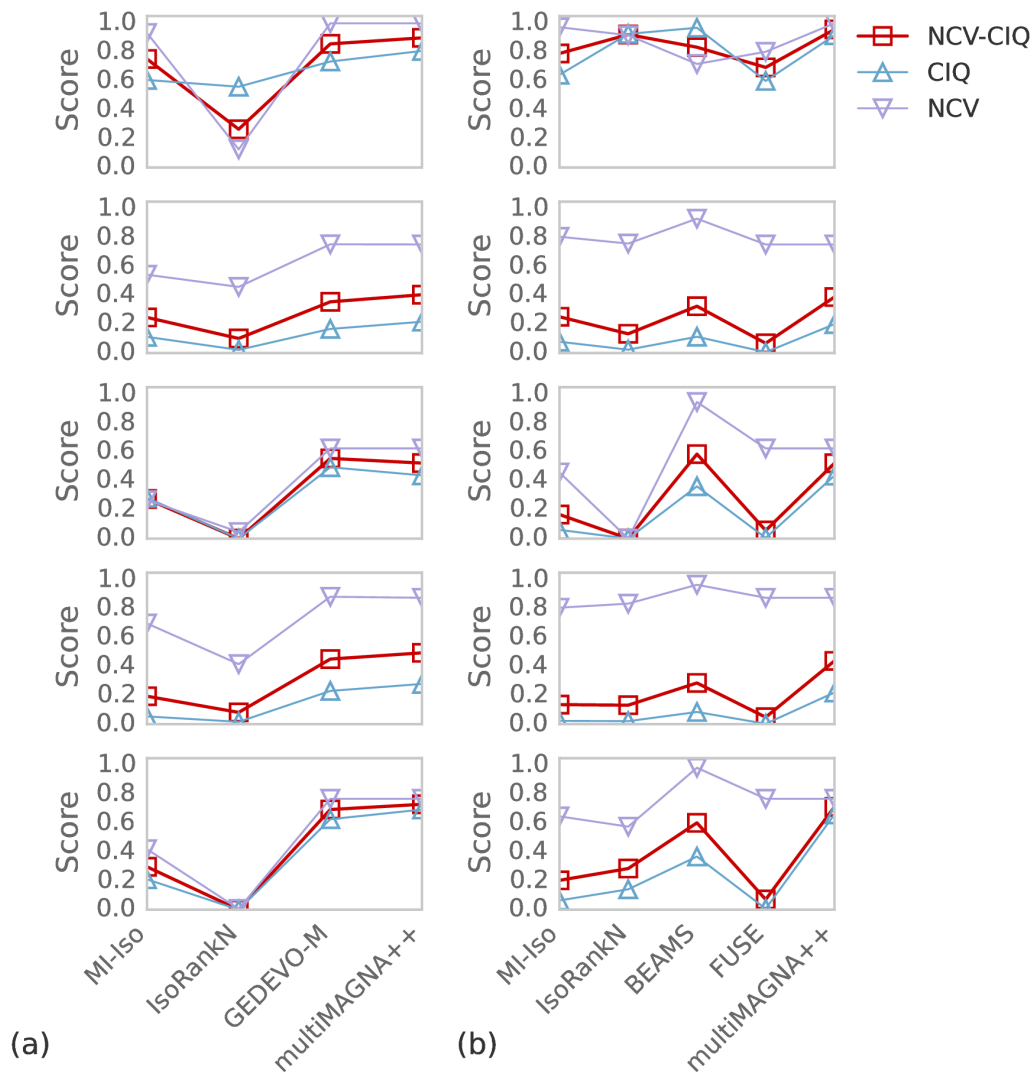


Figure S2: NCV-CIQ, NCV, and CIQ for the five network sets (Yeast+%LC, PHY1, PHY2, Y2H1, and Y2H2 from top to bottom) for (a) topology-only alignments and (b) topology+sequence alignments.

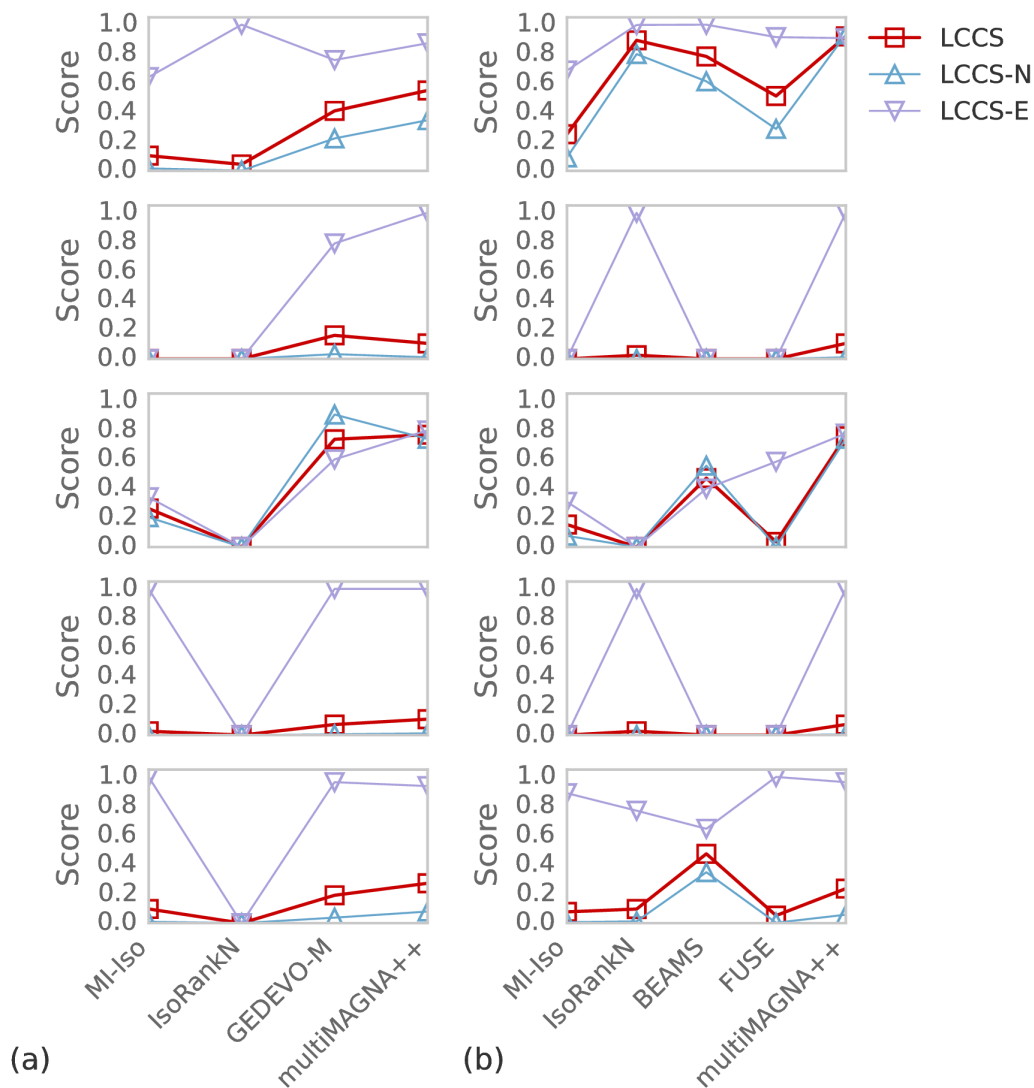


Figure S3: LCCS for the five network sets (Yeast+%LC, PHY1, PHY2, Y2H1, and Y2H2 from top to bottom) for (a) topology-only alignments and (b) topology+sequence alignments.

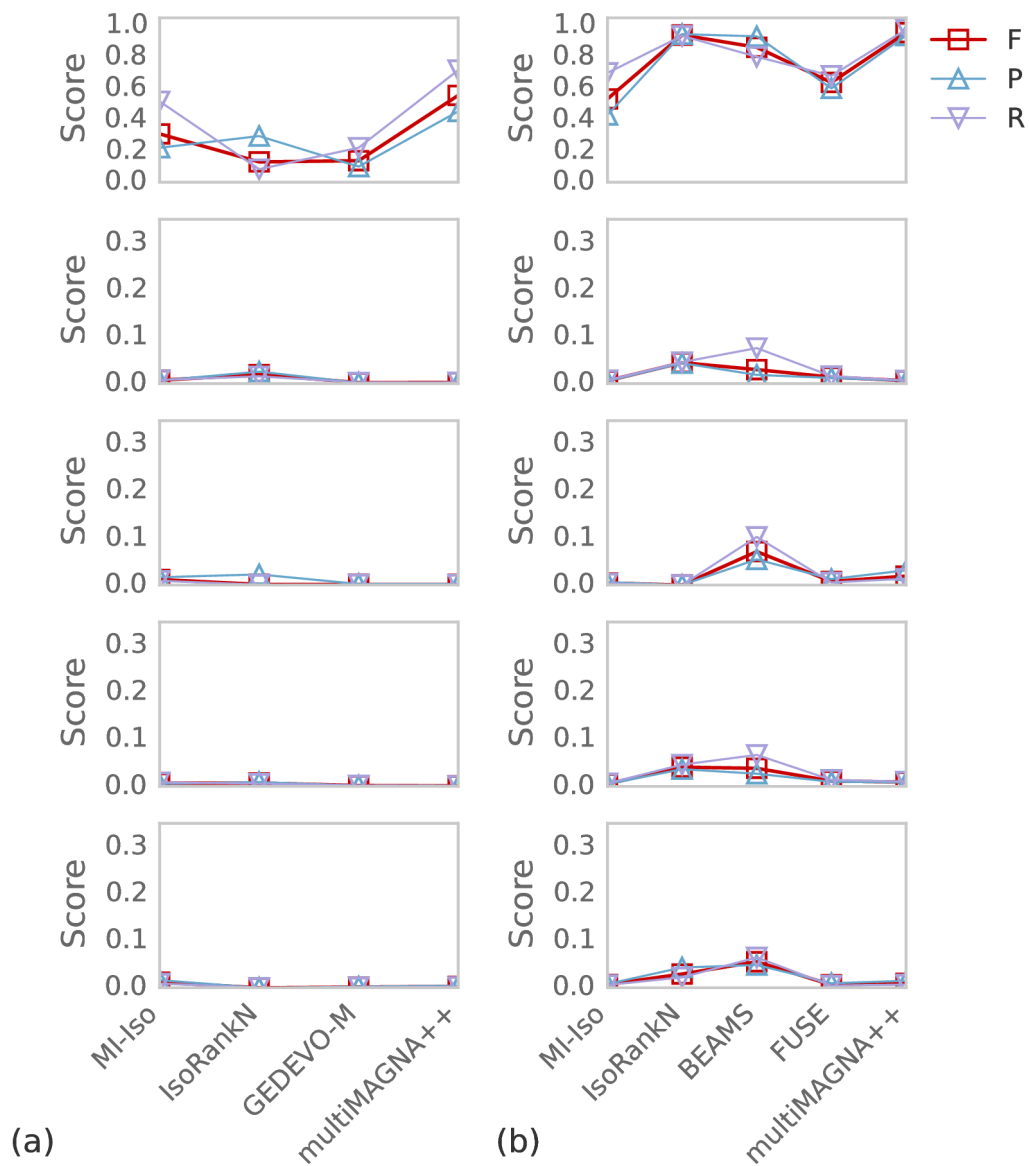


Figure S4: F-score for the five network sets (Yeast+%LC, PHY1, PHY2, Y2H1, and Y2H2 from top to bottom) for (a) topology-only alignments and (b) topology+sequence alignments.

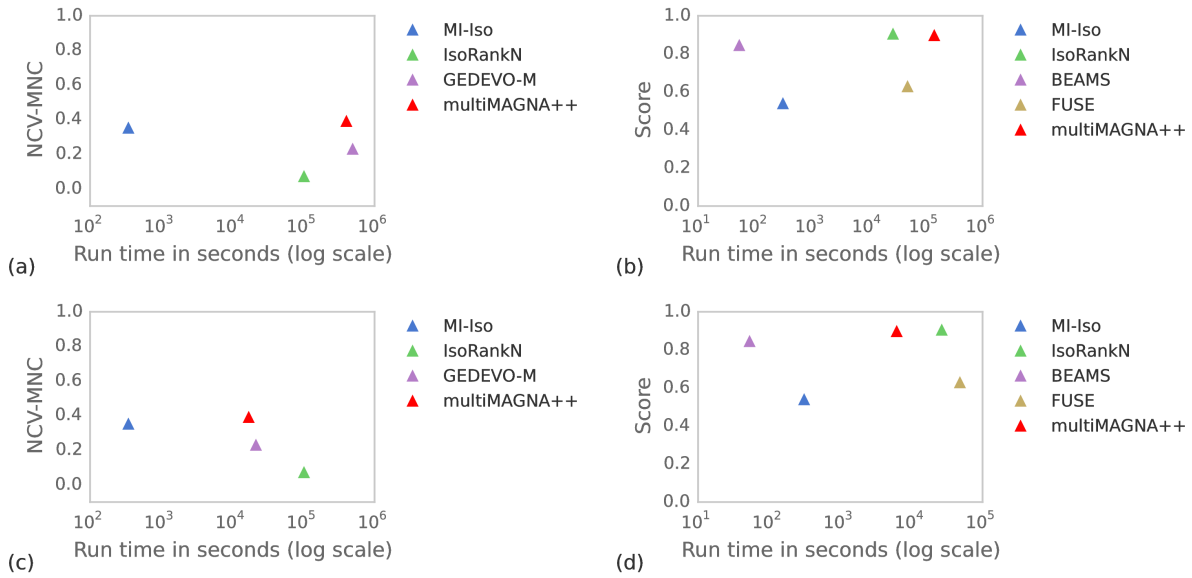


Figure S5: NCV-MNC as a function of time when using **(a-b)** a single thread and **(c-d)** 64 threads, for **(a,c)** topology-only alignments and **(b,d)** topology+sequence alignments, for the three network sets with more than two networks (we leave out these results for PHY2 and Y2H2 that have two networks each).

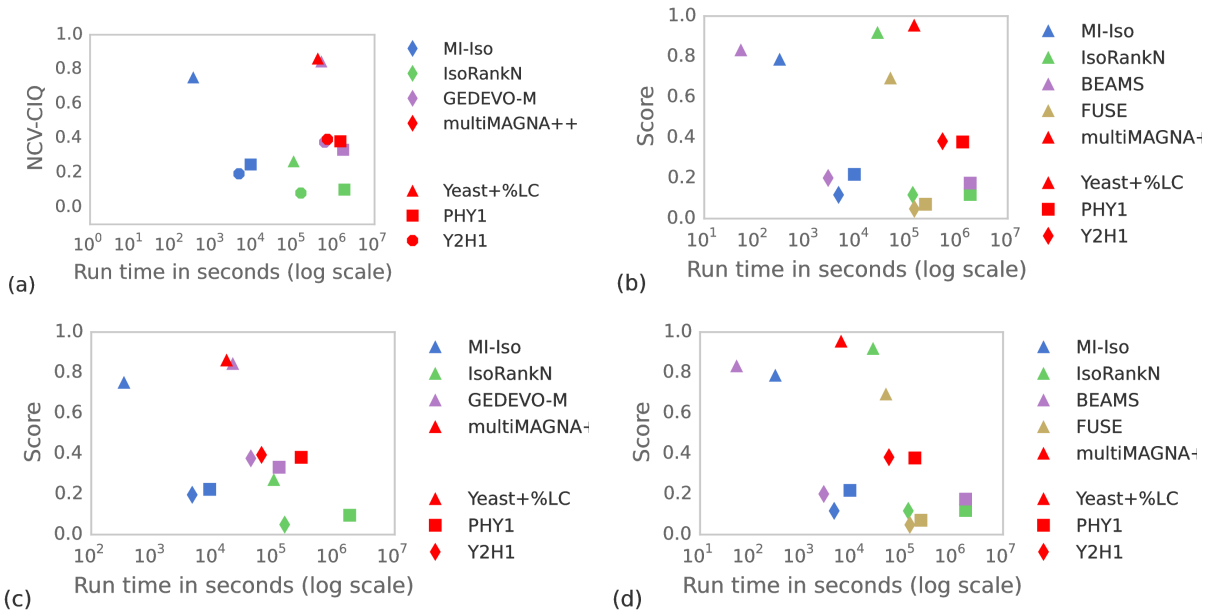


Figure S6: NCV-CIQ as a function of time when using **(a-b)** a single thread and **(c-d)** 64 threads, for **(a,c)** topology-only alignments and **(b,d)** topology+sequence alignments, for the three network sets with more than two networks (we leave out these results for PHY2 and Y2H2 that have two networks each; Section 2.2.1). For equivalent results for the remaining measures, see Supplementary Figures S8 and S9.

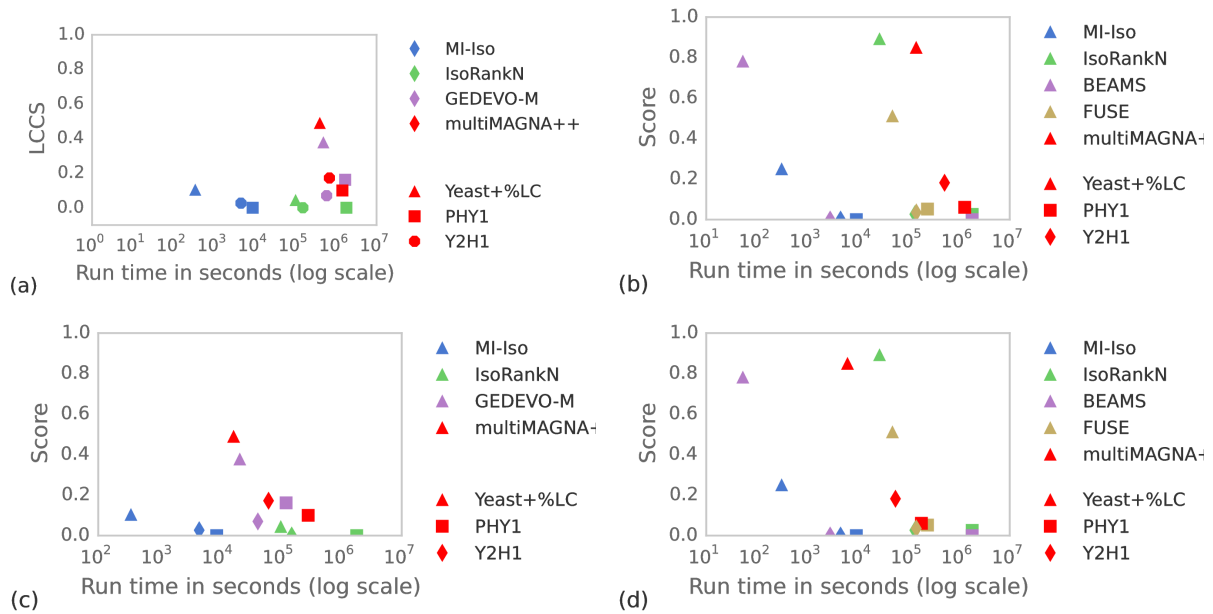


Figure S7: LCCS as a function of time when using (a-b) a single thread and (c-d) 64 threads, for (a,c) topology-only alignments and (b,d) topology+sequence alignments, for the three network sets with more than two networks (we leave out these results for PHY2 and Y2H2 that have two networks each).

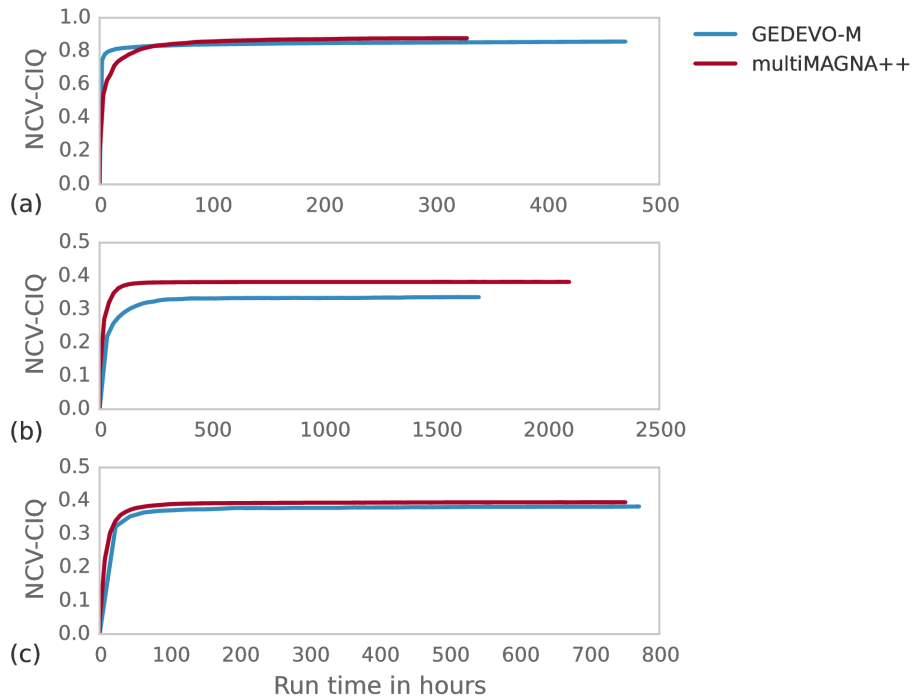


Figure S8: NCV-CIQ as a function of the number of hours spent by multiMAGNA++ and GEDEVO-M when using a single thread, for topology-only alignments of (a) Yeast+%LC, (b) PHY1, and (c) Y2H1 network sets. We only show these results for the three network sets with more than two networks; we leave out these results for PHY2 and Y2H2 that have two networks each. Both multiMAGNA++ and GEDEVO-M are run for 100,000 generations.

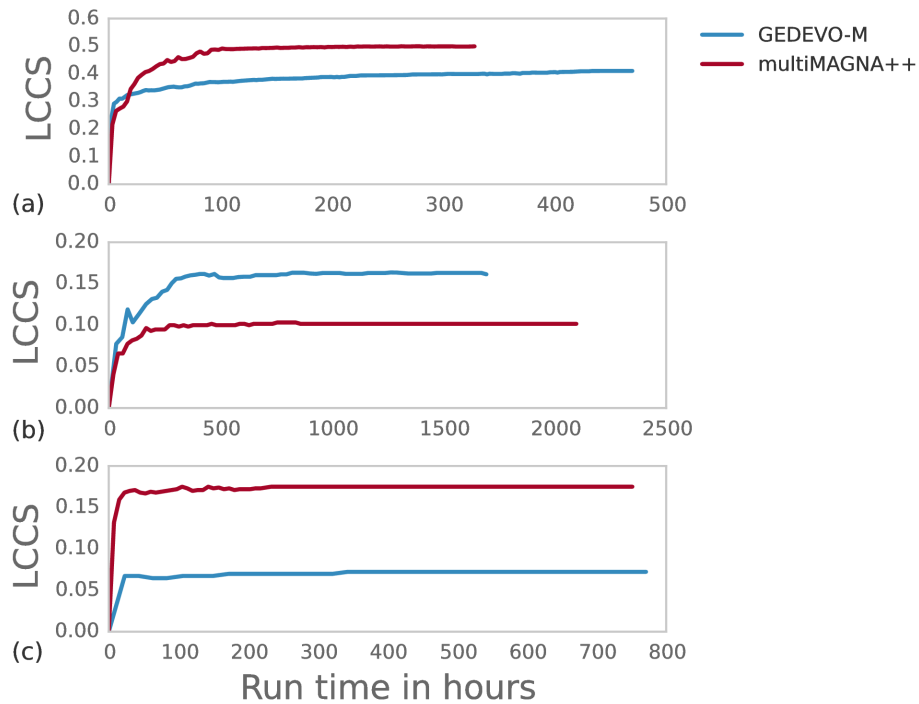


Figure S9: LCCS as a function of the number of hours spent by multiMAGNA++ and GEDEVO-M when using a single thread, for topology-only alignments of **(a)** Yeast+%LC, **(b)** PHY1, and **(c)** Y2H1 network sets. We only show these results for the three network sets with more than two networks; we leave out these results for PHY2 and Y2H2 that have two networks each. Both multiMAGNA++ and GEDEVO-M are run for 100,000 generations.

SUPPLEMENTARY TABLES

| Set | Species | Proteins | Interactions |
|-----------|-------------|----------|--------------|
| Yeast+%LC | Yeast+0%LC | 1,004 | 8,323 |
| | Yeast+5%LC | 1,004 | 8,739 |
| | Yeast+10%LC | 1,004 | 9,155 |
| | Yeast+15%LC | 1,004 | 9,571 |
| | Yeast+20%LC | 1,004 | 9,987 |
| | Yeast+25%LC | 1,004 | 10,403 |
| PHY1 | Fly | 7,887 | 36,285 |
| | Worm | 3,006 | 5,506 |
| | Yeast | 6,168 | 82,368 |
| | Human | 16,061 | 157,650 |
| PHY2 | Yeast | 768 | 13,654 |
| | Human | 8,283 | 19,697 |
| Y2H1 | Fly | 7,097 | 23,370 |
| | Worm | 2,874 | 5,199 |
| | Yeast | 3,427 | 11,348 |
| | Human | 9,996 | 39,984 |
| Y2H2 | Yeast | 744 | 966 |
| | Human | 1,191 | 1,567 |

Table S1: The five PPI networks sets that we use in our study. We know the true mapping for the Yeast+%LC network set unlike for the other network sets. For PHY2 and Y2H2, we only use the yeast and human PPI networks since the fly and worm networks are too small for analysis.

| Algorithms | Parameters |
|------------------------------|---|
| Topology-only alignments | |
| IsoRankN | K=30 thresh=1e-4 maxveclen=5000000 alpha=1.0 |
| MI-Iso | K=30 thresh=1e-4 maxveclen=5000000 alpha=1.0 |
| GEDEVO-M | beta=0.4 |
| multiMAGNA++ | m=CIQ p=15000 n=100000 e=0.5 a=1.0 |
| Topology+sequence alignments | |
| IsoRankN | K=30 thresh=1e-4 maxveclen=5000000 alpha=0.5 |
| MI-Iso | K=30 thresh=1e-4 maxveclen=5000000 alpha=1.0 |
| BEAMS | beta=0.4 alpha=0.5 |
| FUSE | k=[100,100,100,100] iter_num=1000 gamma=0.7 a=0.5 |
| multiMAGNA++ | m=CIQ p=15000 n=100000 e=0.5 a=0.25 |

Table S2: Parameters of the MNA methods. For the existing methods, we use parameters that were recommended in the methods' original publications. The parameters of multiMAGNA++ are the edge conservation measure (m), population size (p), number of generations (n), and the α parameter (a).