

Reproducibility in Methods Publications: An example

Mingfu Shao
June 3, 2019



PennState
College of Engineering

**ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE**



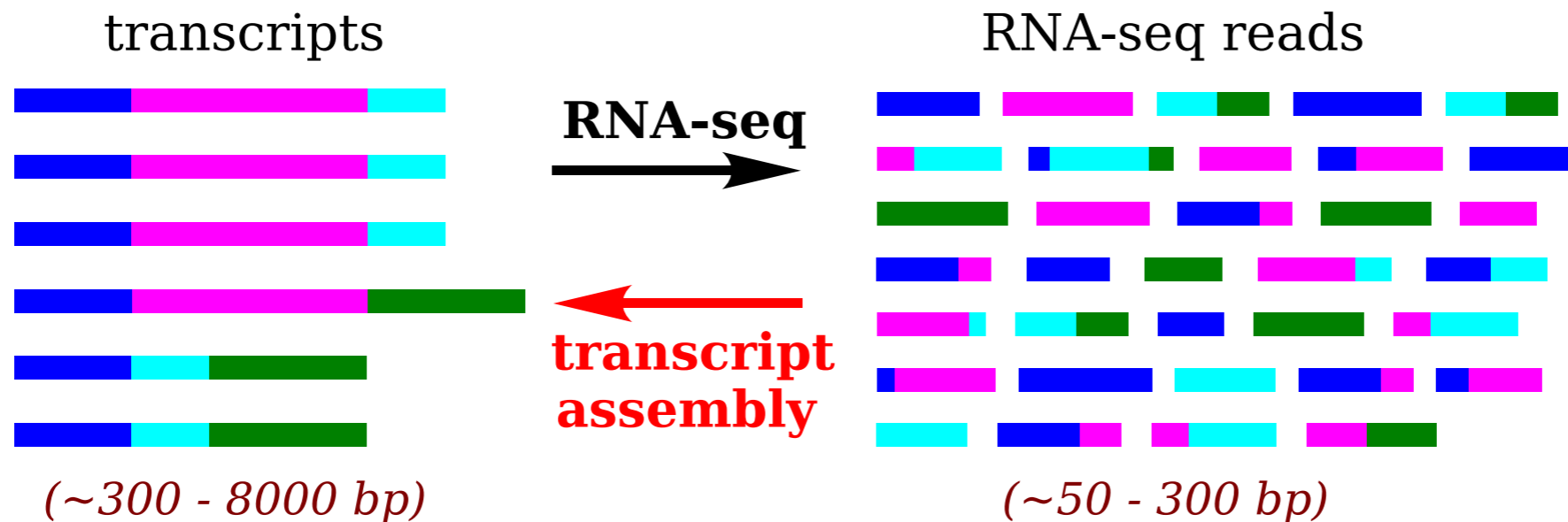
PennState
Huck Institutes of
the Life Sciences

Overview

- **Goal:** discuss good practices to facilitate users reproduce the results in methods papers
- **Example:** Mingfu Shao, and Carl Kingsford. Accurate assembly of transcripts through phase-preserving graph decomposition. *Nature Biotechnology* 35.12 (2017): 1167.

Components of Methods Paper

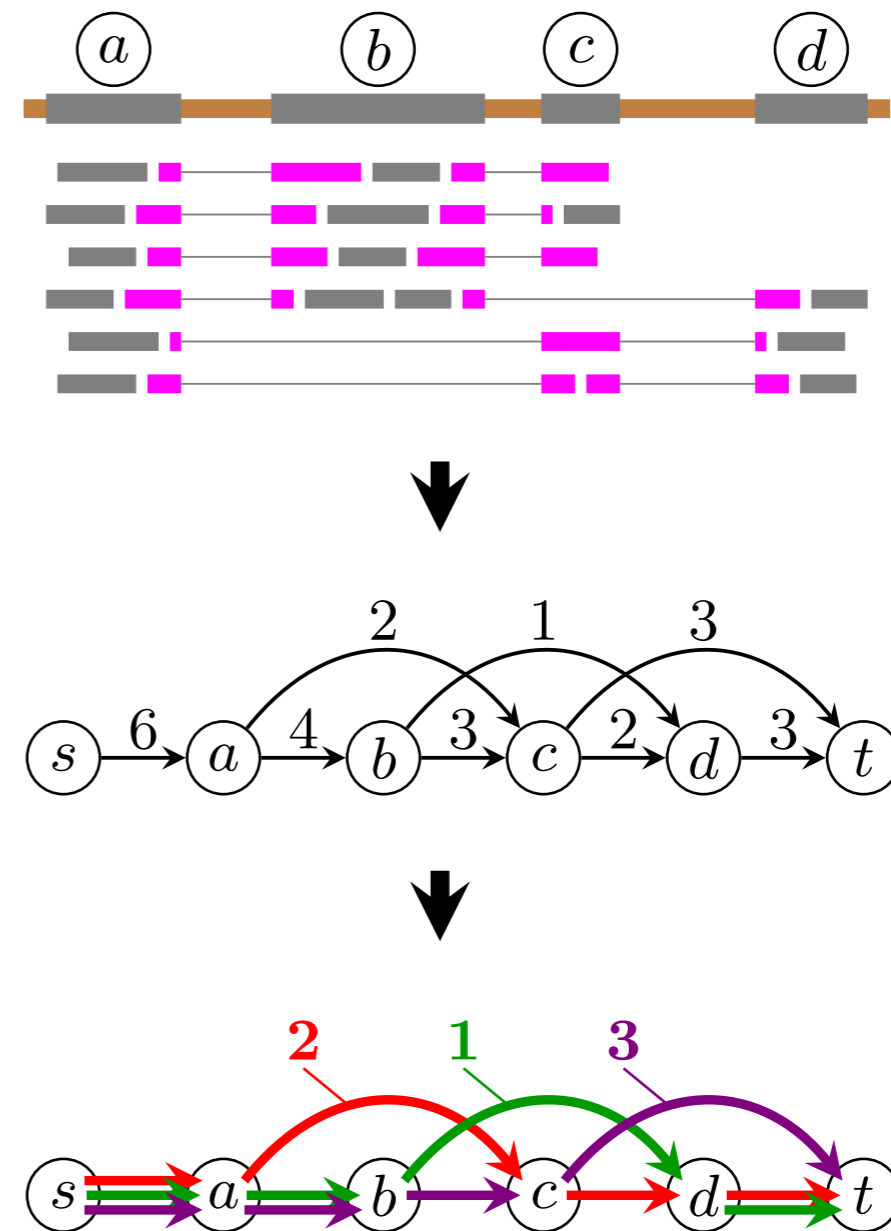
- **Problem statement**
 - Why this problem is important?
 - Challenges
 - State-of-the-art



Components of Methods Paper

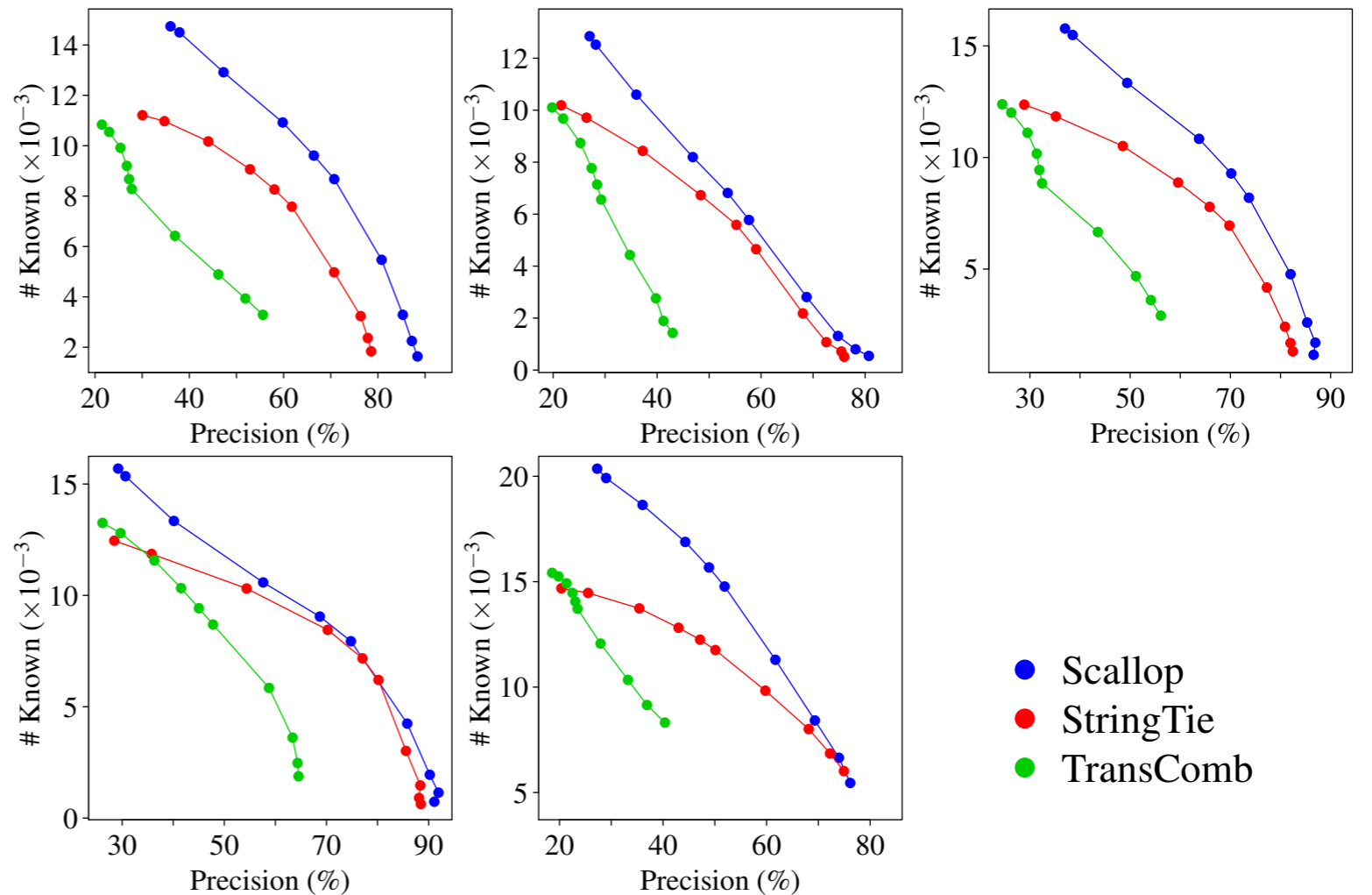
- **Methods description**

- Novelty
- Why does this method work?



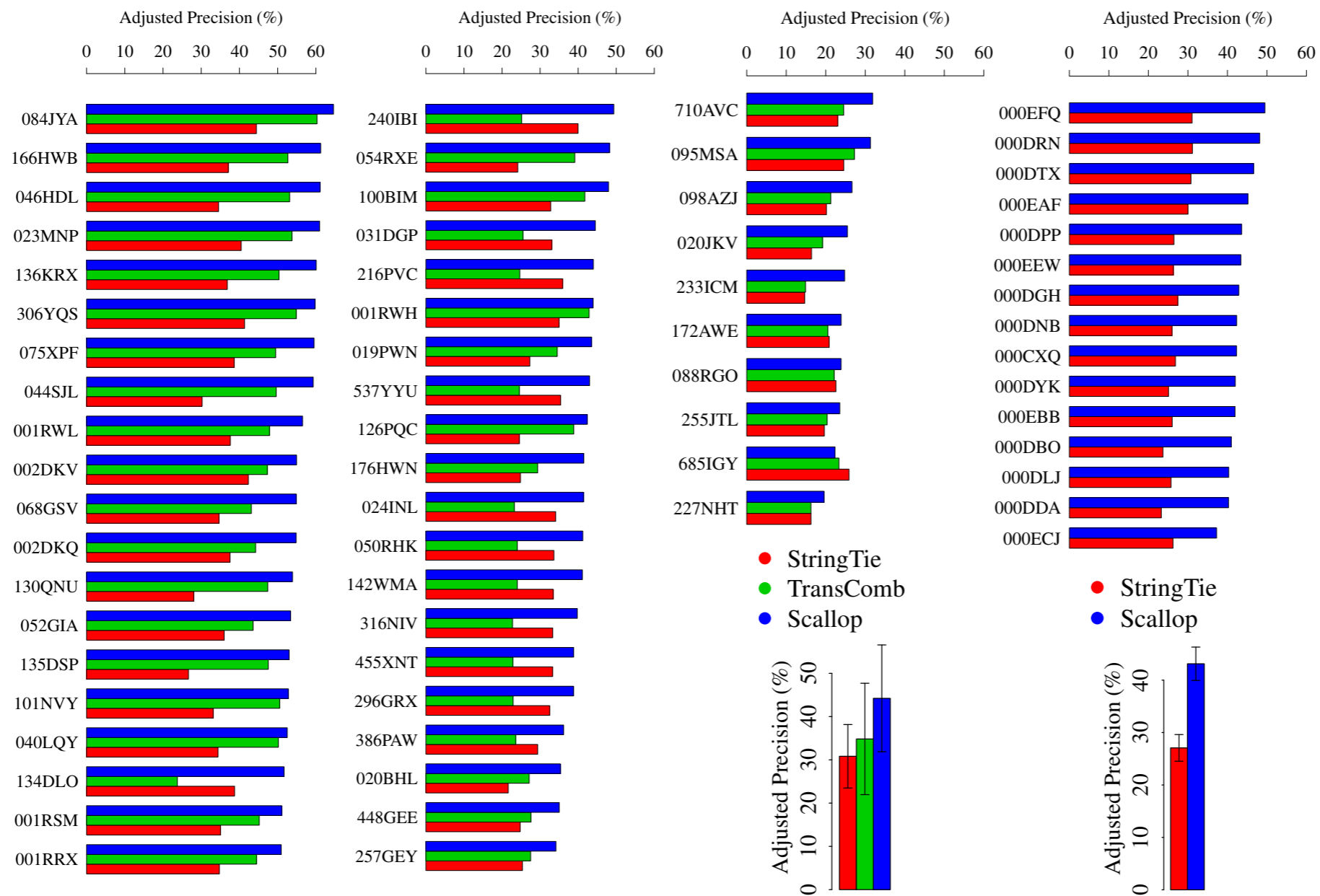
Components of Methods Paper

- **Experimental results**
 - Methods to compare
 - Data sets
 - Ground truth
 - Measurements



Components of Methods Paper

- More experimental results



Reproducibility

- **Goal:** provide codes / scripts / instructions so that users can easily reproduce the experimental comparison figures.
- Two GitHub repos:
 - <https://github.com/Kingsford-Group/scallop>
 - <https://github.com/Kingsford-Group/scalloptest>

Software Repo

- **Releases:** both source codes and binaries
- **Installation**
- **Usage:** describe parameters and how to set them

Reproducibility Repo

- **Programs:** all necessary tools
 - Softwares to compare
 - Simulation tools
 - Evaluation tools
- **Data:** used in the experimental comparison
- **Results:** scripts to run the experiments
- **Plots:** generate the figures in the paper

Summary

- A (separate) repo for reproducing experimental figures.
- Lower goal: provide all instructions so that the results can be produced.
- Higher goal: one command does all.

Summary

- A (separate) repo for reproducing experimental figures.
- Lower goal: provide all instructions so that the results can be produced.
- Higher goal: one command does all.

Thank you!