

Inferring cell cycle phases from a temporal network of protein interactions

Maxime LUCAS

10th July 2021

Networks2021

🐦 maximelca

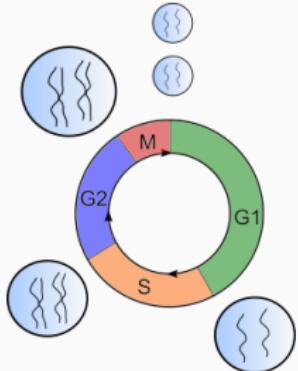
Aix-Marseille University

work with A. Barrat, B. Habermann, L. Tichit, A. Morris, and A. Townsend-Teague

bioRxiv:2021.03.26.437187

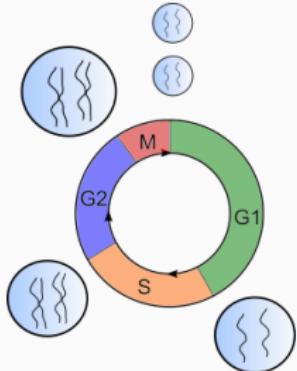
Biological systems go through different phases/states

Cell cycle phases



Biological systems go through different phases/states

Cell cycle phases

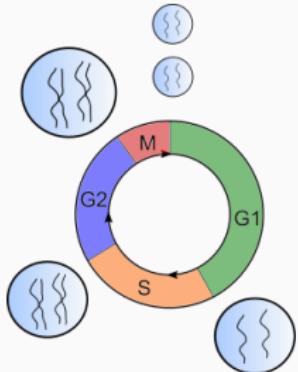


Sleep stages



Biological systems go through different phases/states

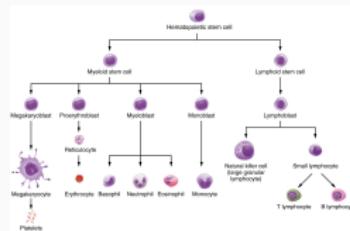
Cell cycle phases



Sleep stages

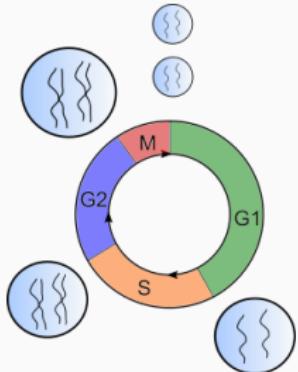


Cell differentiation



Biological systems go through different phases/states

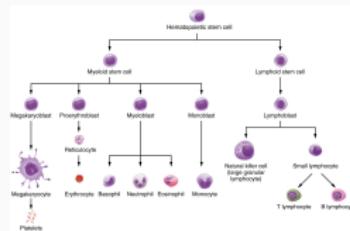
Cell cycle phases



Sleep stages



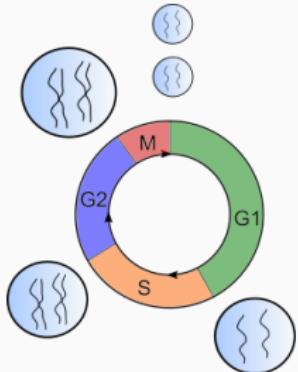
Cell differentiation



Temporal structure \leftrightarrow biological behaviour

Biological systems go through different phases/states

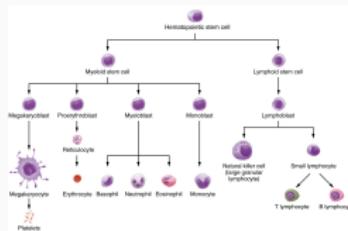
Cell cycle phases



Sleep stages



Cell differentiation



Temporal structure ↔ biological behaviour

active protein interactions
expressed genes

In this talk

A computational method, Phasik, to infer the temporal organisation of biological systems, over multiple temporal scales.

Goal : help biologist understand their system and design experiments better

In this talk

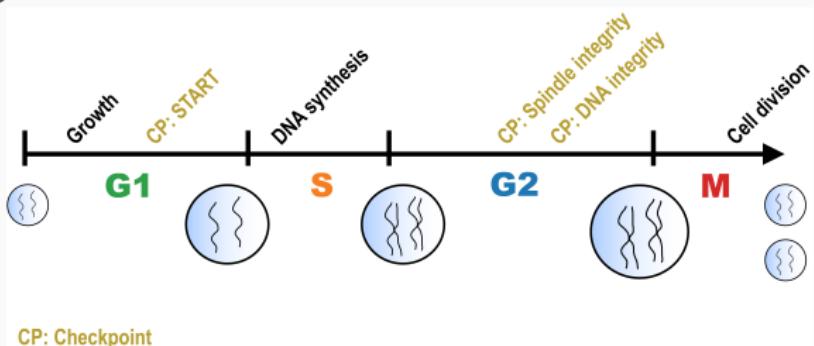
A computational method, Phasik, to infer the temporal organisation of biological systems, over multiple temporal scales.

Goal : help biologist understand their system and design experiments better

To validate our method: the **cell cycle**, because it is so well known.

Cell cycle: protein interactions that change over time

... leading to the cell division.



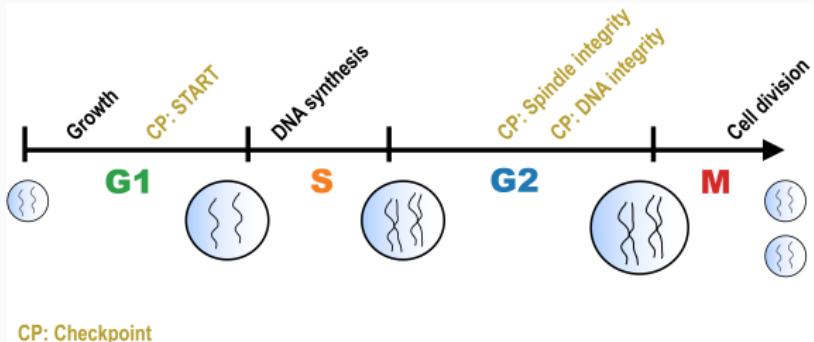
Multiple relevant timescales:

- Macro: 4 physiological phases (G1, S, G2, M)
- Meso: physiological subprocesses
- Micro: protein interactions that change over time

Time-ordering ensured by molecular **checkpoints**.

Cell cycle: protein interactions that change over time

... leading to the cell division.



Multiple relevant timescales:

- Macro: 4 physiological phases (G1, S, G2, M)
- Meso: physiological subprocesses
- Micro: protein interactions that change over time

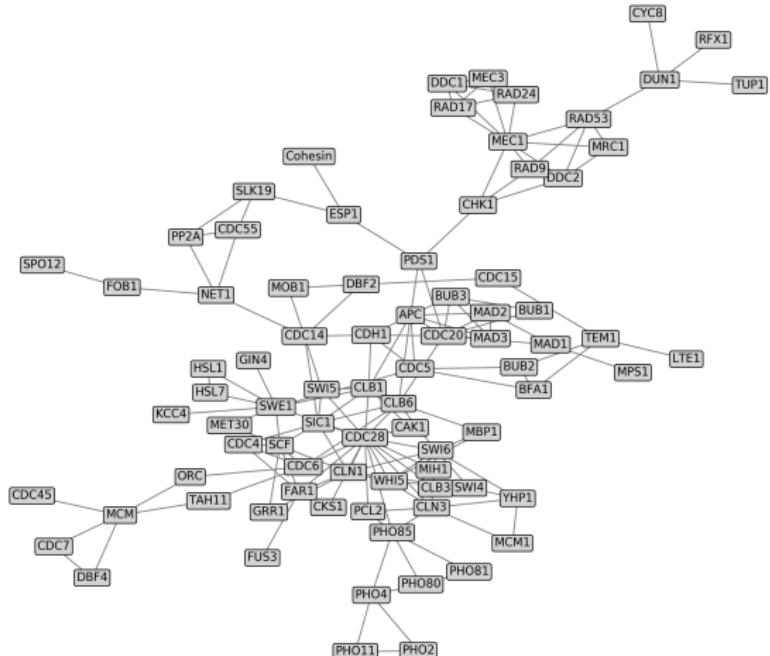
Time-ordering ensured by molecular **checkpoints**.

We focus on **budding yeast** because it is best known.

**Can we predict the phases
from the temporal protein
interactions?**

Static network of protein interactions represents the cell cycle

84 nodes (proteins) connected by 159 edges (protein interactions).



All temporal information is lost!

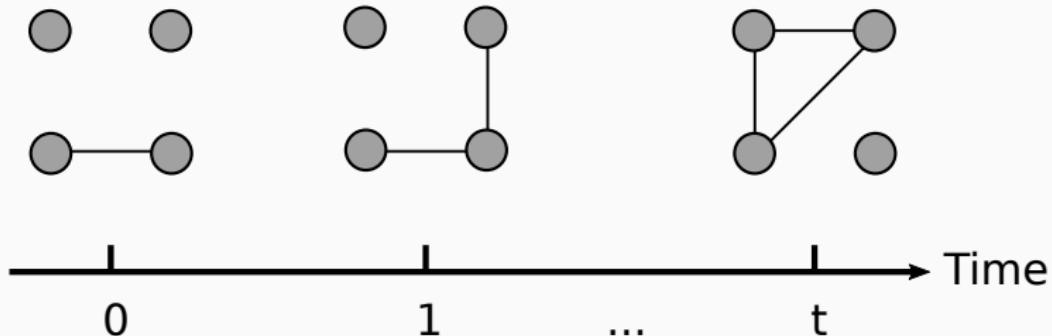
Interaction data: KEGG.

**All temporal information is lost..
we need temporal networks**

Temporal networks incorporate temporal information

The **edges** are now **time-varying**.

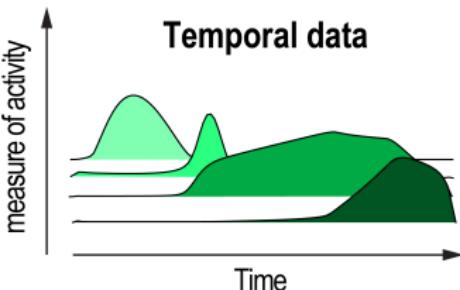
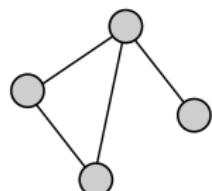
At each time corresponds a **snapshot** of the temporal network



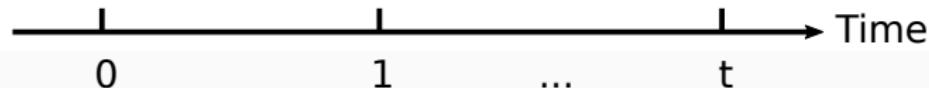
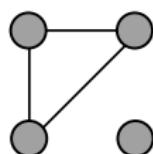
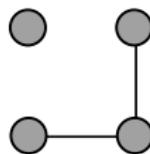
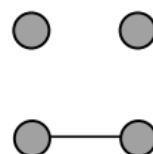
Phasik: 1. build a temporal network

Build temporal network by integrating time series

Static network

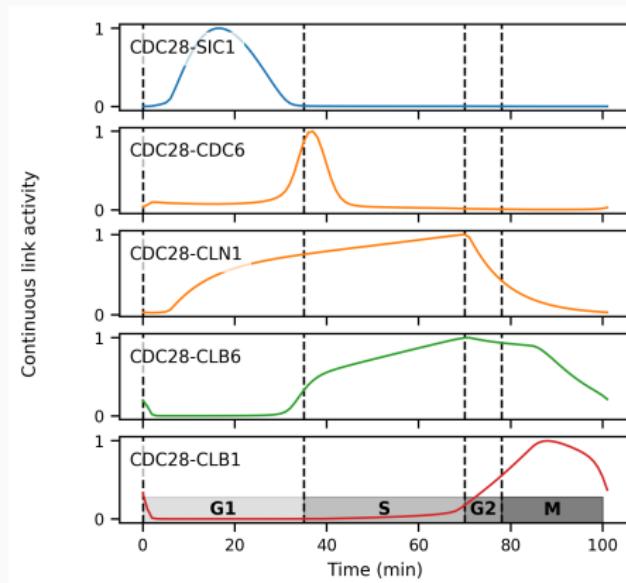


Temporal network



Inject time series data into the static network

Temporal data needed: mathematical model or RNA-seq.
Our network is **partially temporal** (34/159 edges).

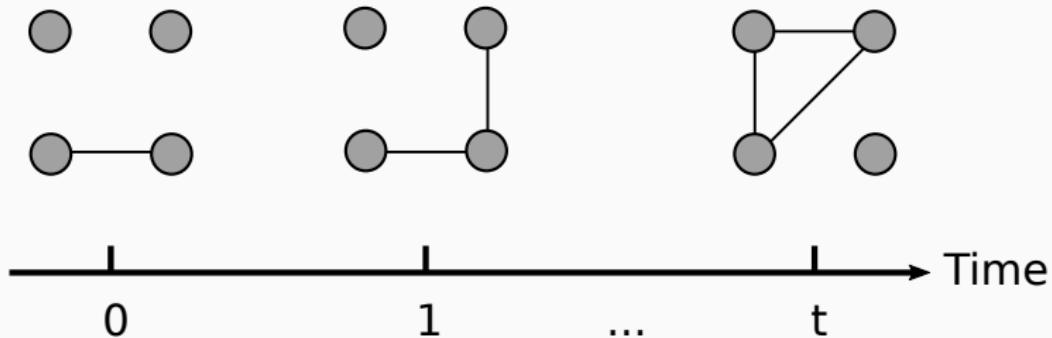


Temporal data: Chen 2014 and Kelliher 2016.

Phasik: 2. infer biological phases

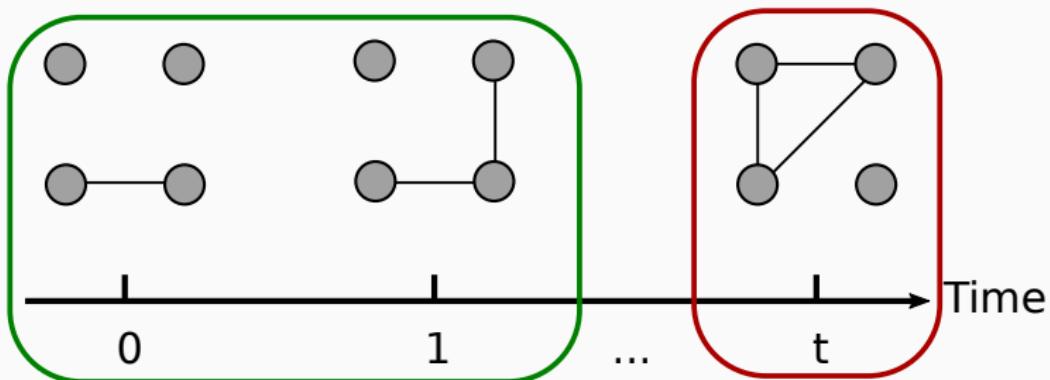
Inferring biological phases by clustering

Idea: the temporal network stays similar when it stays in the same “state” or phase, but changes a lot when it changes phase (Masuda and Holme, 2019).



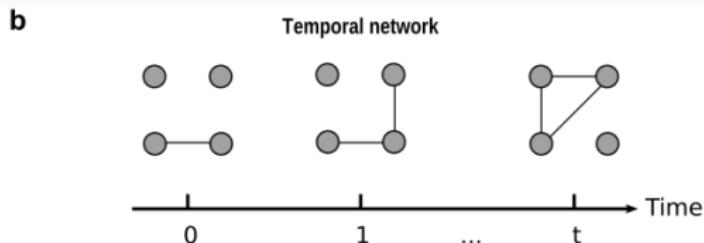
Inferring biological phases by clustering

Idea: the temporal network stays similar when it stays in the same “state” or phase, but changes a lot when it changes phase (Masuda and Holme, 2019).



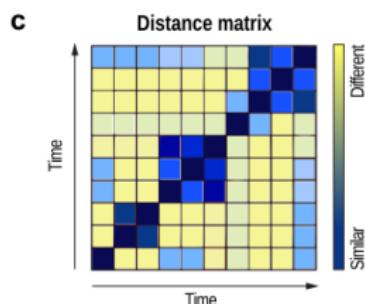
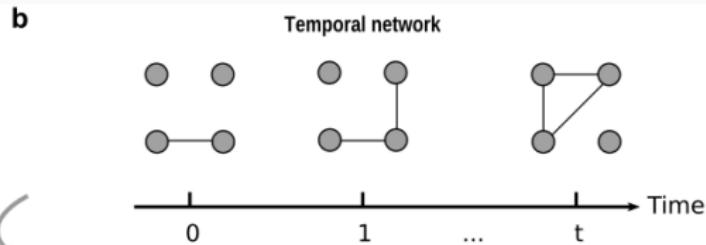
In more details

We apply hierarchical clustering to the snapshots.



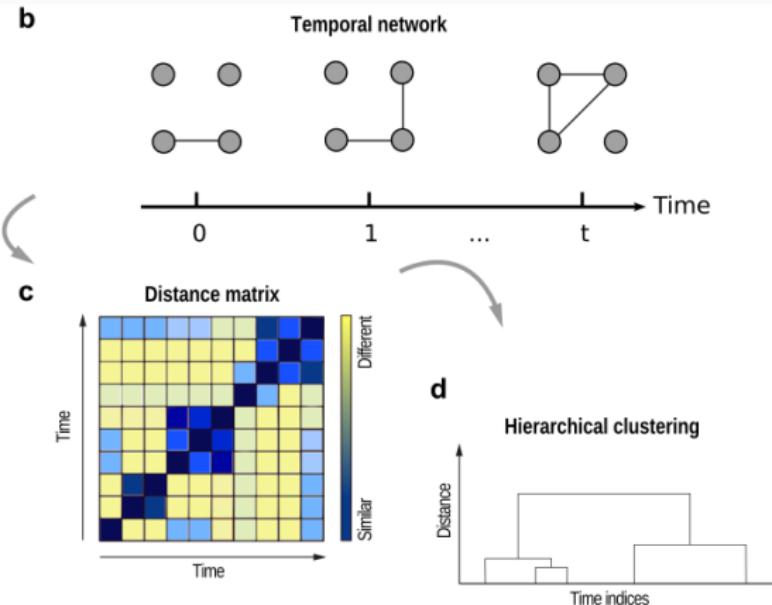
In more details

We apply hierarchical clustering to the snapshots.



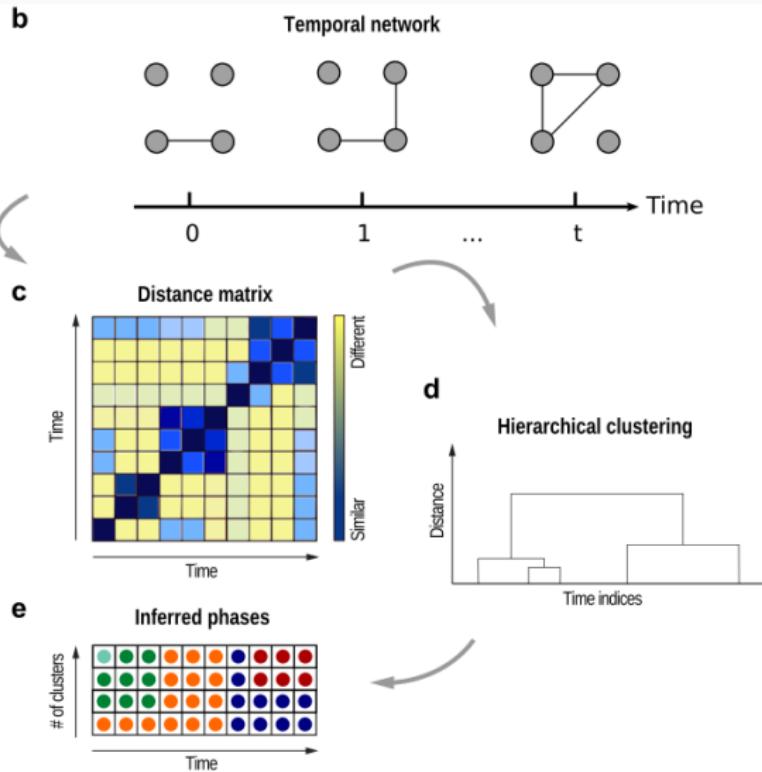
In more details

We apply hierarchical clustering to the snapshots.



In more details

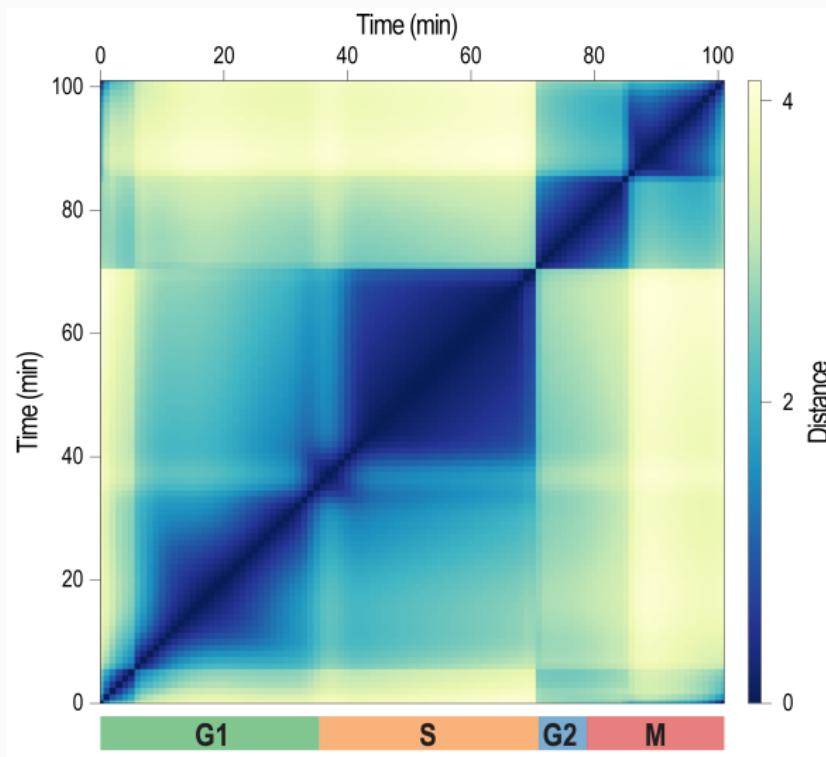
We apply hierarchical clustering to the snapshots.



Results

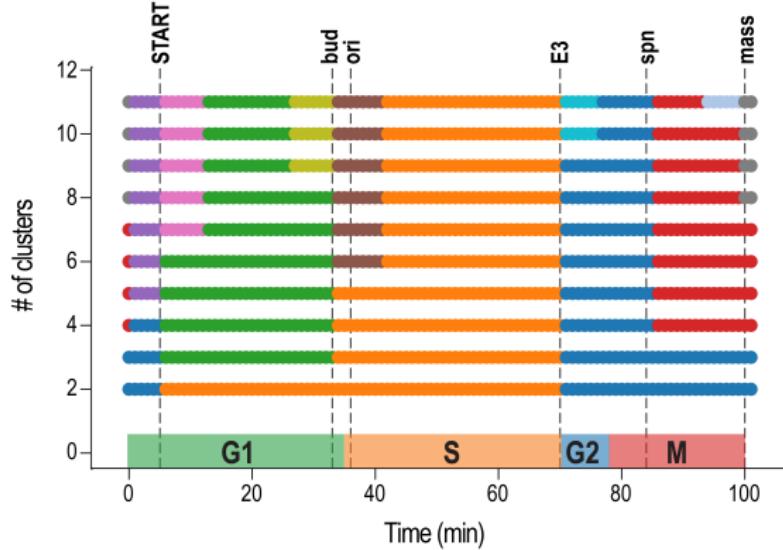
We recover known cell cycle phases

Distance matrix:



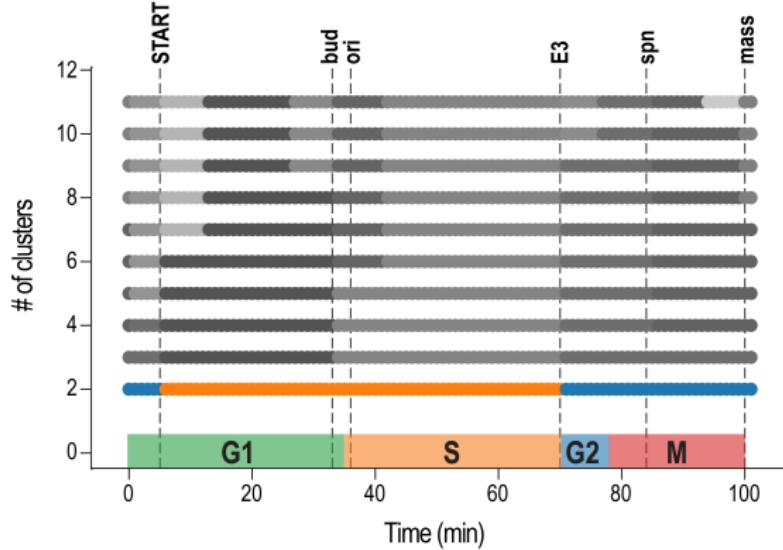
We recover known cell cycle phases

Few clusters → coarse-grained view of the cell cycle.



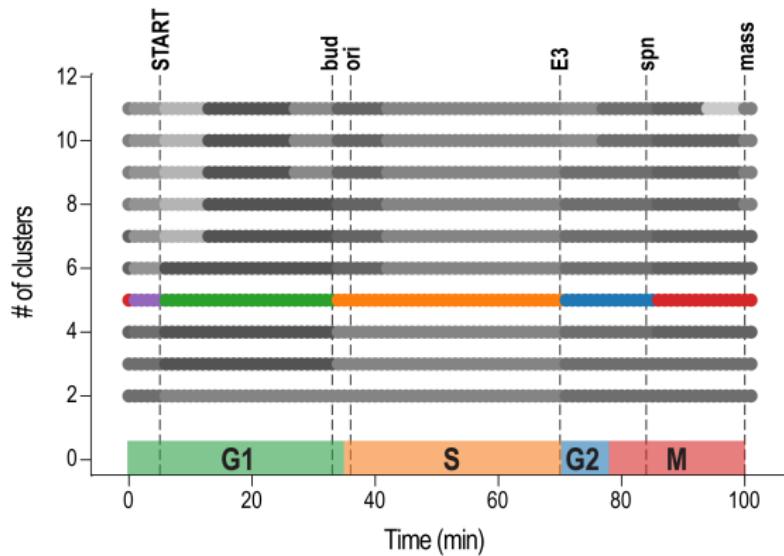
We recover known cell cycle phases

Few clusters → coarse-grained view of the cell cycle.



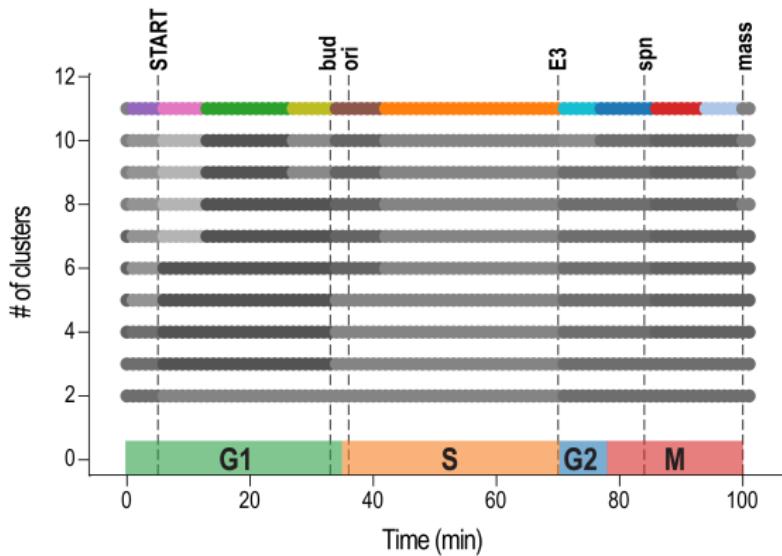
We recover known cell cycle phases

Few clusters → coarse-grained view of the cell cycle.



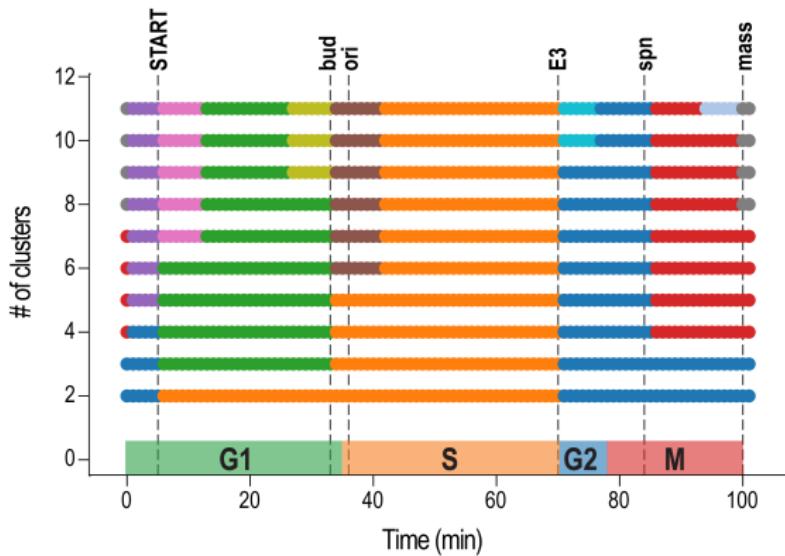
We recover known cell cycle phases

Few clusters → coarse-grained view of the cell cycle.



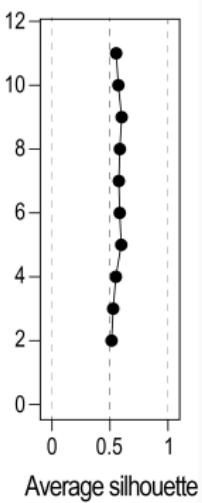
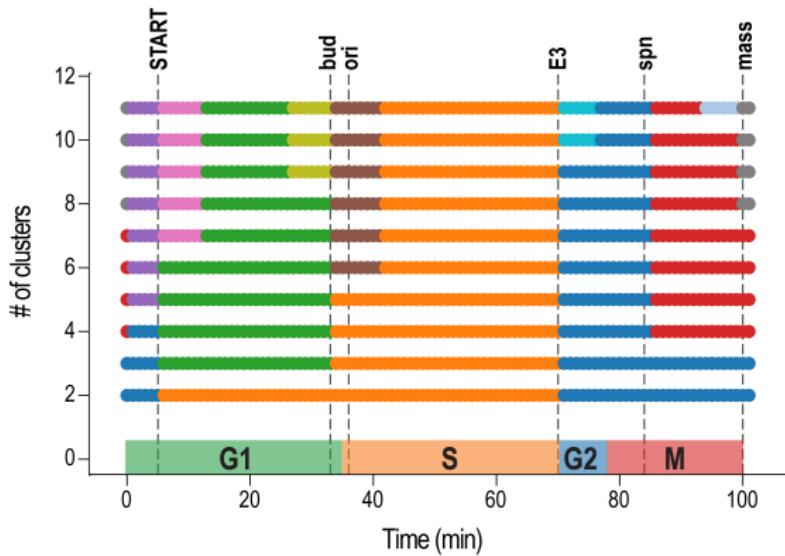
We recover known cell cycle phases

Few clusters → coarse-grained view of the cell cycle.



Multiple scales are relevant

Quality of clusterings (average silhouette) is constant across scales



Is the method, Phasik, robust?

Yes, we made many robustness tests

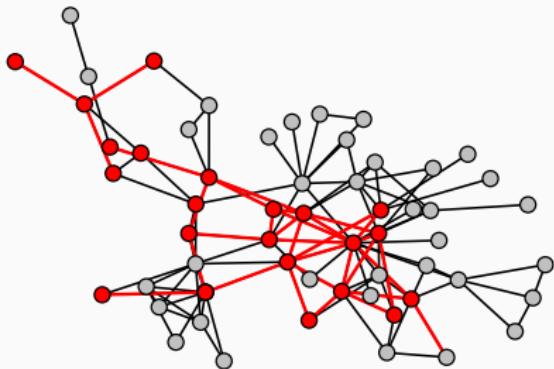
Phasik and its results are **robust against**:

- changes in clustering method
- changes in distance metric
- measurement noise in time series
- downsampling of time series

**How little temporal information do
we need?**

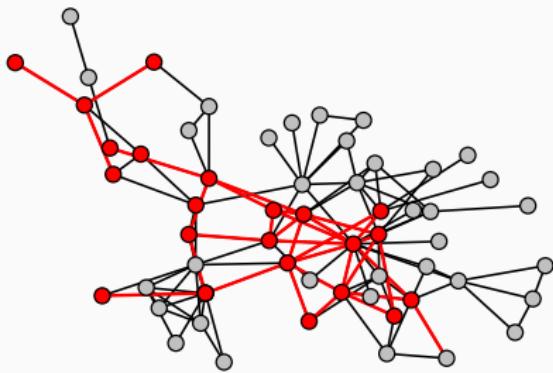
Imagine we have access to only CDC28's interactions

Original: 34 / 159 edges with
temporal information

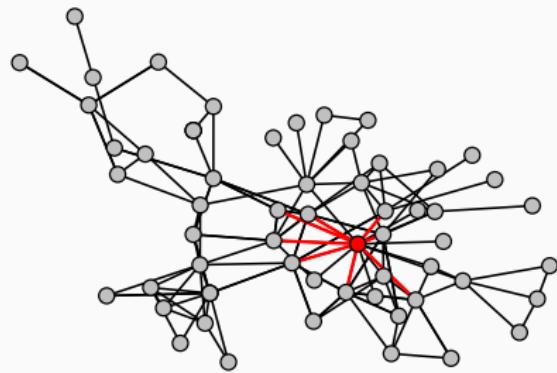


Imagine we have access to only CDC28's interactions

Original: 34 / 159 edges with temporal information



Now: 8 / 159 edges with temporal information

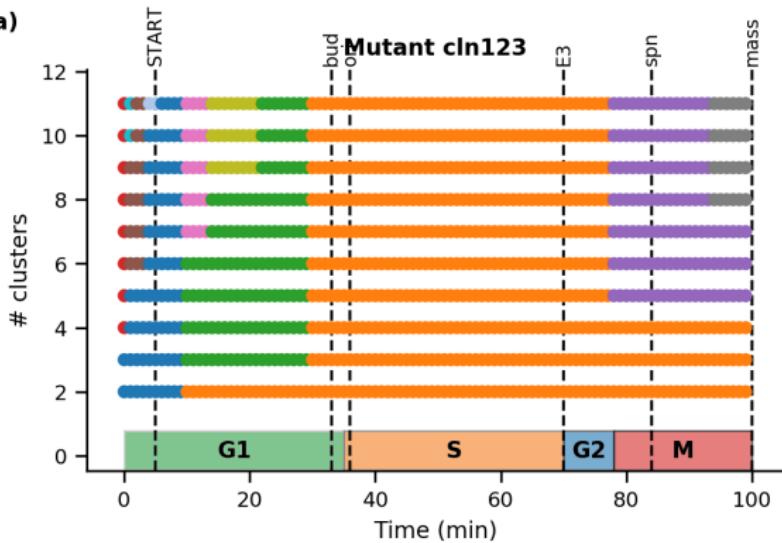


Edges with no temporal information are shown in grey.

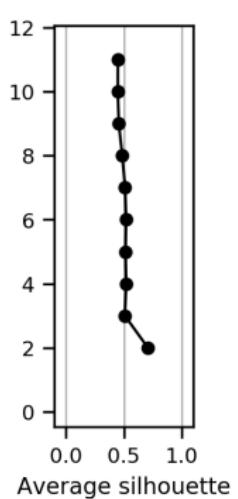
**Can Phasik detect modified phases
in mutants?**

Mutant phases: G1 arrest in Δ CLN1/2/3

(a)



(b)



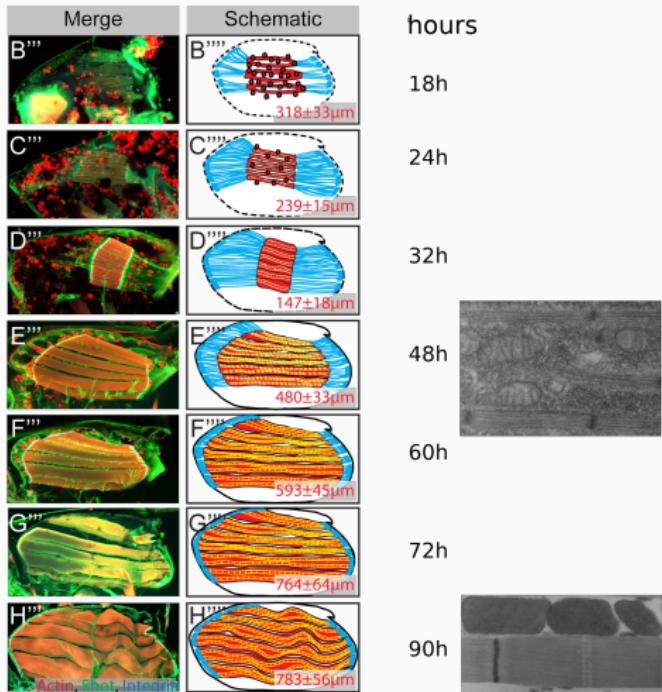
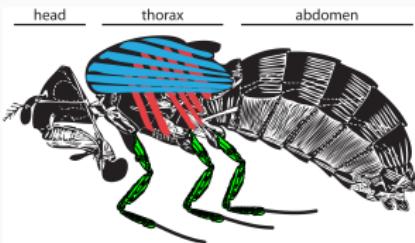
**Phasik can be used with gene
expression data too!**

**Can the method be used on other
biological systems?**

Yes!

Flight muscle development in *Drosophila*

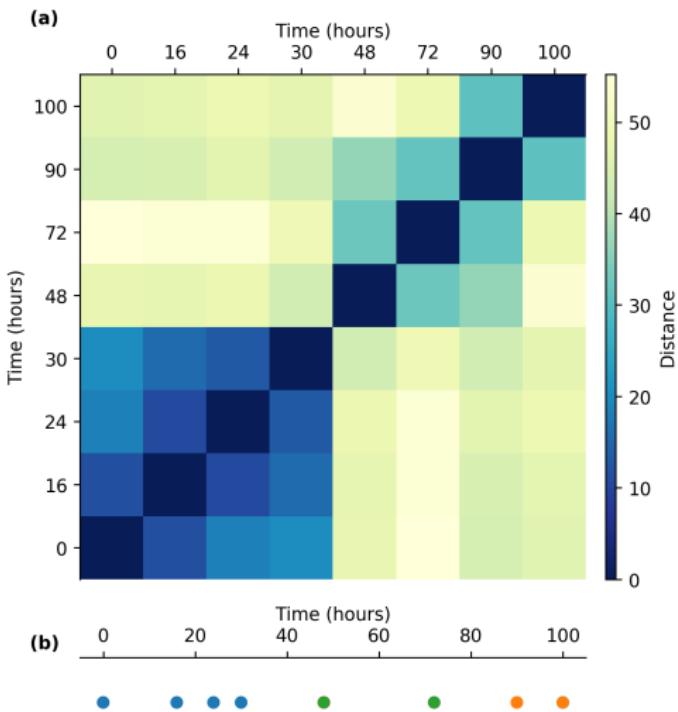
Flight muscles (blue) have densely packed, cristae-rich mitochondria



RNA-seq measurement of the genes at each time.

Flight muscle development in *Drosophila*

We have 8 time points between 0 and 100 hours.



Phasik can be used on a wide range of biological systems

All you need is:

- **time series data** about biological units (e.g. proteins/genes/..) or their interactions
- **interaction data** (e.g. static PPI network)

Use our code on your data!



- Our user-friendly **code is available online**:
https://gitlab.com/habermann_lab/phasik
- Functions for each step of the pipeline: temporal networks building, and phase inference.
- **Online documentation**:
<https://phasik.readthedocs.io/en/latest/>
- Available as a **Python package**. **Install it** in the terminal:
`pip install phasik`

Conclusions

- We represented the **cell cycle** as a **temporal network** of protein interactions
- From that, we **inferred biological phases** of the cell cycle by clustering snapshots
- We investigated **how much**, and what, **temporal data** is necessary.
- We applied the method to cell cycle mutants and flight muscle development in *Drosophila*

Next steps: Apply this method to other less well-known biological systems. Let us know about yours!

Thank you for your attention!

Thanks to the people I work with: Alain Barrat, Bianca Habermann, Laurent Tichit, and everyone in their teams!

✉ maxime.lucas.1@univ-amu.fr

🐦 maximelca

Any questions?

