

Circadian Rhythm Cell Signalling Pathways in Fungi

Austin Bassett
Test123@hotmail.com
University of Victoria

Helen Dun
Test123@hotmail.com
University of Victoria

ABSTRACT

Fill In Later.

KEYWORDS

fungi, orthologous, circadian rhythm, cellular signalling pathway, protein-protein interactions (ppi)

ACM Reference Format:

Austin Bassett and Helen Dun. 2021. Circadian Rhythm Cell Signalling Pathways in Fungi. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

1.1 Circadian Rhythms

Even with varying bedtimes and wake up times, people typically live on a 24-hour cycle. This process is known as the circadian rhythm. The circadian rhythm is an internal and natural process which repeats roughly every 24 hours and is known for regulating the sleep-wake cycle [1]. Many organisms have some form of circadian rhythm and have been observed in animals, plants and even fungi and cyanobacteria [1]. Circadian rhythms are adjusted by external cues called zeitgebers, and these include light, temperature, and redox cycles [1]. The driving force of the circadian rhythm is the circadian clock. The circadian clock is a biochemical oscillator which cycles using a stable phase while also being synchronized with solar time [2]. Circadian clocks consist of three major components, a central biochemical oscillator, input pathways to the central oscillator, and output pathways tied to distinct phases of the oscillator [2]. These are all put together using cell signalling.

1.2 Cell Signalling Pathways

Cell signalling is how the cells of an organism signal to each other what to do. Signals are usually propagated using different types of proteins, although other molecules are also used. The way the signal is propagated from a protein to another protein to yet another protein can be mapped to a directional pathway. Pathways are not always linear and can be very complex as sometimes multiple different proteins are needed at the same time to propagate signals and signals can be propagated at the same time to different proteins. Visually mapping the pathways is very important for helping scientists and doctors understand how biological systems

work at the cellular level and how different medicines affect the pathways at a chemical level.

1.3 Overview

Due to the complexity of cellular signalling pathways and the interestingness of circadian rhythms, for our humble project we compare the circadian rhythm pathways of fungi. Specifically, we compared what proteins are and aren't orthologous between the pathways and graphed this visually so we could see which parts of the circadian rhythm pathways were significantly different or similar between fungi. Finally, we analyzed why sections of the pathway were similar, simpler or more complex and what the benefits of each were.

2 PROGRAM

To achieve our goal, we developed a Python script that would read in the data sets of different databases and compare the data. There were 3 different databases used for this:

- Circadian Gene Database (CGD)
- Orthologous MAtrix Project (OMA)
- STRING Database

CGD was used to find the lists of circadian rhythm proteins. OMA was used to find orthology between the proteins of organisms. STRING was used to graph the output of the program for better visualization and analysis.

For 2 organisms A and B, the program outputs 8 different sets of proteins as text files: Imperfect Union (A,B), Perfect Union (A,B), Imperfect Except (A,B), Perfect Except (A,B), Imperfect Union (B,A), Perfect Union (B,A), Imperfect Except (B,A) and Perfect Except (B,A). The next few sections will explain, first, the reason why the Union sets are ordered, then what Union and Except sets are, and, finally, the difference between Imperfect and Perfect sets.

2.1 Why Order Matters for Union Sets

Some proteins of organism A can be orthologous to many proteins in organism B and vice-versa. Therefore, for the union of orthologous proteins in A and B the program would have to create a node that combines all proteins orthologous with each other. This was infeasible as for the graph part STRING only takes a list of proteins so we would need to create or use some other graphing software for the project. Simplifying the graph of the pathway may also be destructive to information about the proteins. Instead, Union (A,B) is the set of proteins in A that are orthologous to some protein in organism B and Union (B,A) is the set of proteins in B that are orthologous to some protein in organism A.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference'17, July 2017, Washington, DC, USA
© 2021 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

2.2 What are Union and Except Sets

There are 2 set operations this program performs on the proteins of the 2 selected organisms: Union (U) and Except (-). As mentioned in the previous section, the Union operation is slightly different from the well-known set operation as it does not consolidate the 2 input sets. The Union operation takes the sets of proteins for 2 organisms, A and B, and the orthology between them and returns the subset of proteins of organism A that are orthologous to some protein of organism B. The Except operation takes the same input but does the opposite of the Union function: it returns the subset of proteins of organism A that are not orthologous to any protein of organism B. Protein A from Organism X is orthologous to a set of proteins P of Organism Y. For some protein Q in P, Q may or may not be a protein of the cellular signalling pathway being studied in Organism Y. Suppose Q is not. Do we keep that A is orthologous to Q? The Imperfect set keeps the orthology between A and Q and the Perfect set does not.

2.3 Output

Now that that has all been explained, let us go back to the overview of the algorithm. So the program outputs to the ./output/text directory the 8 different sets of proteins as text files. Each protein in the text file is represented by its OMA id and the proteins are separated by newlines. For Union set text files, each protein also lists the set of proteins orthologous to it using OMA ids and with orthology scores. The program also outputs to the ./output/graphs directory the 8 different sets of proteins and the full set of proteins for each organism, but this time it skips the extra Union set stuff and outputs STRING ids instead of OMA ids. These text files can be directly input to the STRING database to get a graph of the set of proteins, which there are some examples of in the next section. Along with all of that, the program also outputs how many proteins are in each set.

We used the program to compare and analyze the phylogeny of the circadian rhythm pathways of the fungi, *Neurospora crassa* and *Aspergillus flavus*.

2.4 Imperfect vs. Perfect

The difference between Imperfect and Perfect sets is that Imperfect sets have proteins that are orthologous to proteins that may not be part of the circadian rhythm pathway in the other organism. If that did not make sense, here is the explanation again in logical form.

3 RESULTS

- Talk about the *Neurospora* *Aspergillus* circadian rhythm pathway
- Full pathway graphs
- Talk about the orthologous proteins / Number results for the graphs
- Analysis of results
- Note why not all the nodes are connected

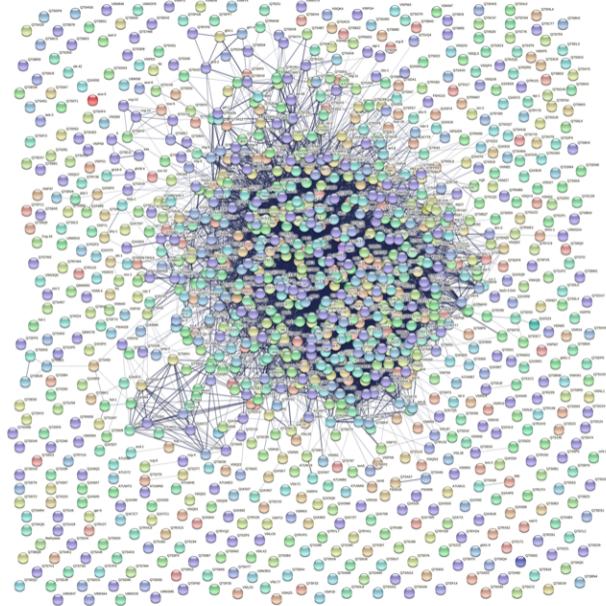


Figure 1: *Neurospora crassa*

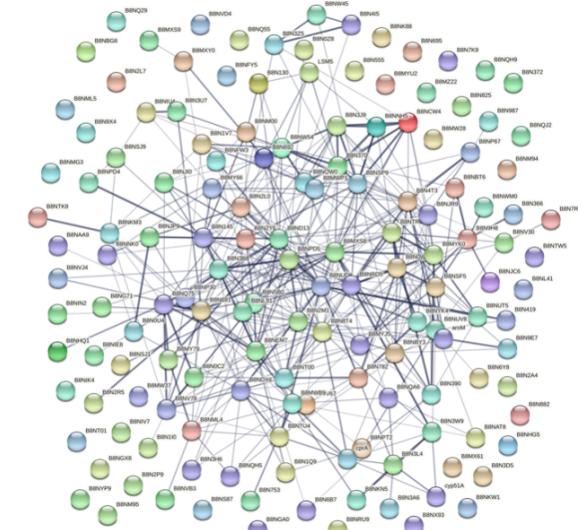
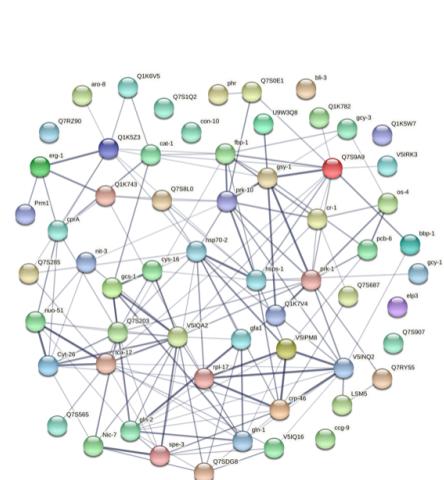


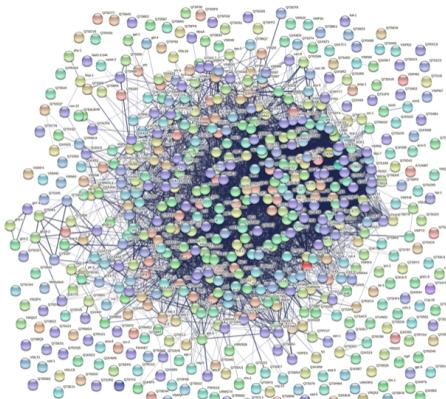
Figure 2: *Aspergillus flavus*

3.1 Order: (*Neurospora crassa*, *Aspergillus flavus*)

While many of the proteins of the *Neurospora crassa* are orthologous to proteins in *Aspergillus flavus*, not many are orthologous to proteins in *Aspergillus flavus*'s circadian rhythm (52.5% of proteins in Imperfect Union → 5.8% of proteins in Perfect Union). This is either because the proteins of the imperfect set but not of the perfect



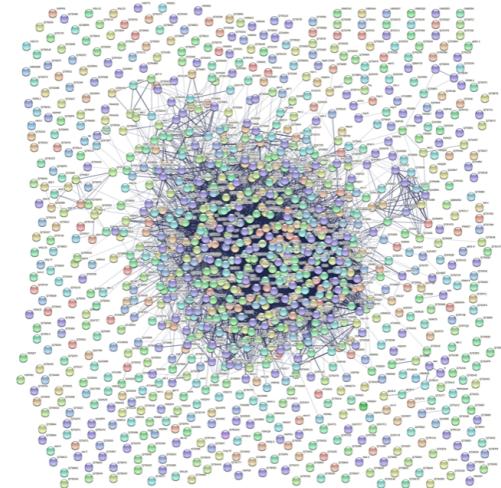
(a) Perfect Union



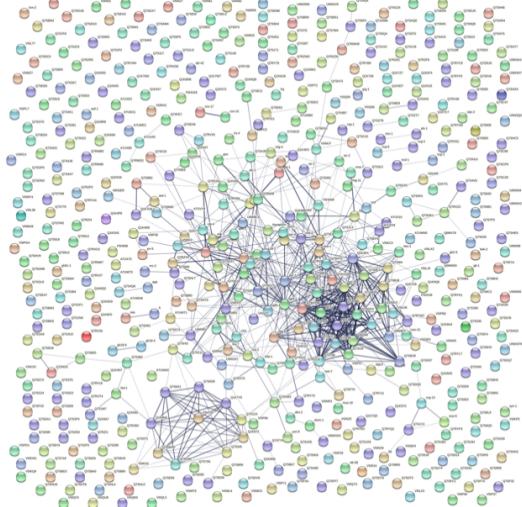
(b) Imperfect Union

Figure 3: Union of *Neurospora crassa*, *Aspergillus flavus*

set are loosely connected to the circadian rhythm and propagate input and output signals, and would evolve to or from proteins in other pathways. Or this is because the list of circadian rhythm proteins used is unfortunately out-of-date. If the first is true, then it means more pathways or evolved, more complex pathways use the circadian rhythm in *Neurospora Crassa* than in *Aspergillus flavus*.



(a) Perfect Except

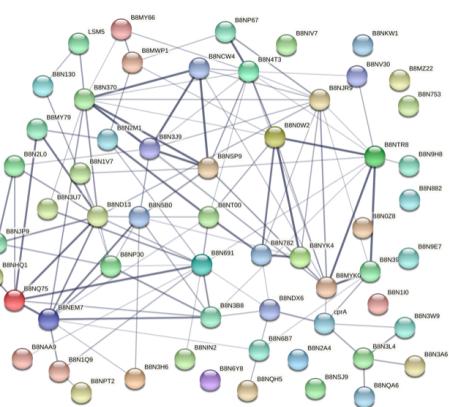


(b) Imperfect Except

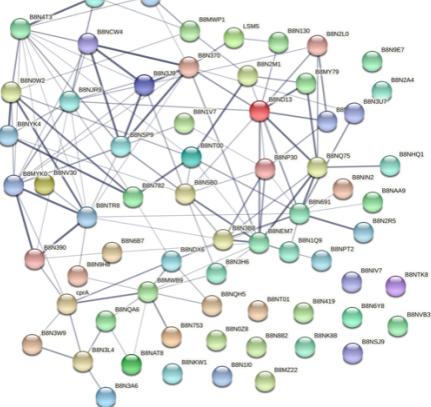
Figure 4: Union of *Neurospora crassa*, *Aspergillus flavus*

3.2 Order: (*Aspergillus flavus*, *Neurospora crassa*)

In comparison to the huge disparity between the Imperfect and Perfect Union sets of *Neurospora crassa*, *Aspergillus flavus*'s Imperfect and Perfect Union sets are not very different at all (44.6% of proteins in Imperfect Union → 39.2% of proteins in Perfect Union). This disparity seems to be evidence of a lack of up-to-date knowledge about circadian rhythm pathways by the CGD, or at least an inconsistency in the way the CGD considers some proteins of the circadian rhythm pathway and others not. Thus the previous idea that other pathways have evolved from a few output-signal-propagating proteins is inconclusive for these graphs, but it may still pertain to future, more accurate graphs.



(a) Perfect Union



(b) Imperfect Union

Figure 5: Union of *Apergillus flavus*, *Neurospora crassa*,

4 DISCUSSION

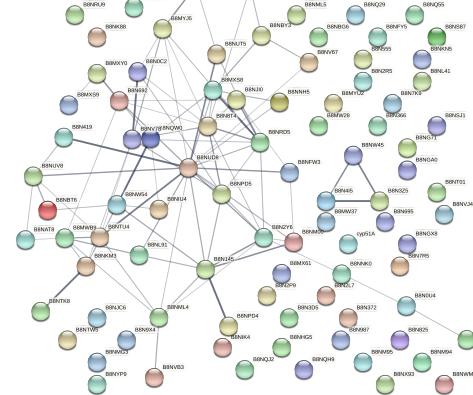
- Issues with mapping from database to database
- Future improvements
 - full graph with nodes colored green and red
 - interactable graph that links to protein's interactions

5 CONCLUSION

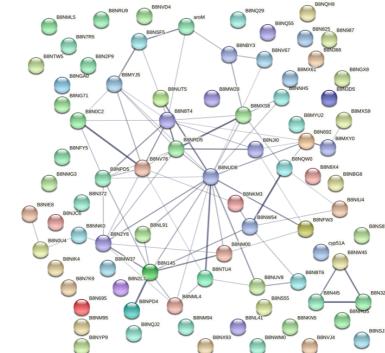
- Reiteration of why this is important
- How this can be used for later

REFERENCES

- [1] "Circadian rhythm," Wikipedia.[Online]. Available: https://en.wikipedia.org/wiki/Circadian_rhythm. [Accessed: 26-Nov-2021.]



(a) Perfect Except



(b) Imperfect Except

Figure 6: Except of *Apergillus flavus*, *Neurospora crassa*,

[2] "Circadian clock," Wikipedia.[Online]. Available: https://en.wikipedia.org/wiki/Circadian_clock. [Accessed: 26-Nov-2021.]