Multiplex methods provide effective integration of multi-omic data in genome-scale models

Structures called networks are often used as a way of simplifying complex problems. For instance, you've probably heard of social networks, where we simplify the complexity of human interactions by categorizing each pair of people as either "friends", or "not friends".

A multiplex network is a slightly more complex way of looking at things, where we think about multiple possible types of links being possible. To continue the social network example, we might have both a "colleague" connection, and a "friend" connection with another person. Each person constitutes a *node* of the network. The same pair of people could be linked in either way, both, or neither, but there might be interesting patterns in how these connections relate to each other. Each of the two types of connections has its own layer. There will be therefore a layer showing all “friends” connections, and a layer showing all “colleagues” connections. Both layers will be part of the same network, called *multiplex network*.

Multiplex networks are very useful in biology, since this often involves situations where the same entities could have multiple types of interactions. Imagine now each node represents a growth condition (for instance the temperature and how much food it has available) for a bacterium, *Escherichia coli*. Conditions may be linked by their similarity, which could be measured in a number of different ways. In our paper, we use two methods of measuring similarity: how the bacteria respond genetically to these different conditions and how we predict their genetic reactions would affect their behaviour. These two measures of similarity give us two network layers, in the same way as measuring people by how close their are as friends or colleagues gives two social network layers.

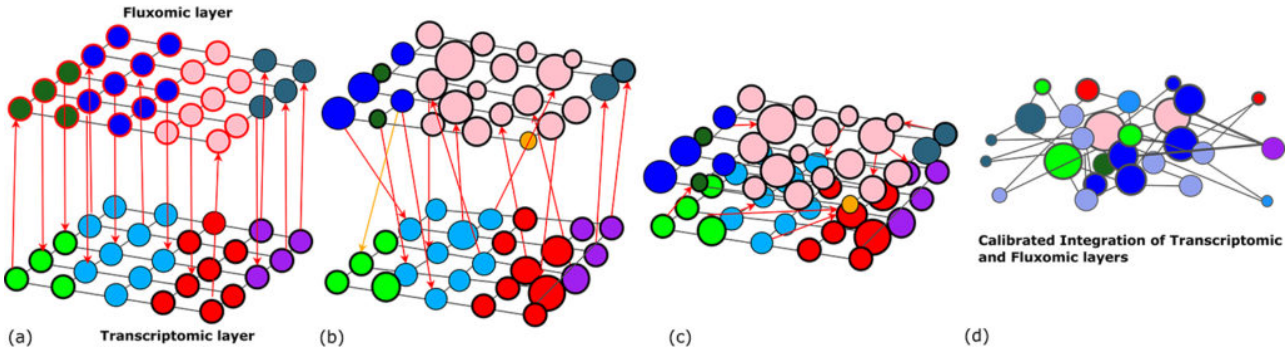

Fig. 1.

We then fuse these layers into one network, which allows us to see a combined picture of how the different conditions affect the bacteria in both genetic and behavioural ways at once. We can then

see clusters in the bacterial responses. To continue the social network analogy, this is a bit like combining the ‘friends’ and ‘colleagues’ networks and finding that, for instance, people in the same town are more likely to be friends and colleagues (Fig. 1).

Our method can be used to evaluate and cross-compare different collections of conditions among different species. The advantage is that acquiring multi-layer information on bacterial growth conditions makes it possible to build condition-specific models even when experimental data is not directly available. Furthermore, it is possible to computationally predict conditions in which the bacterial growth can be controlled effectively.

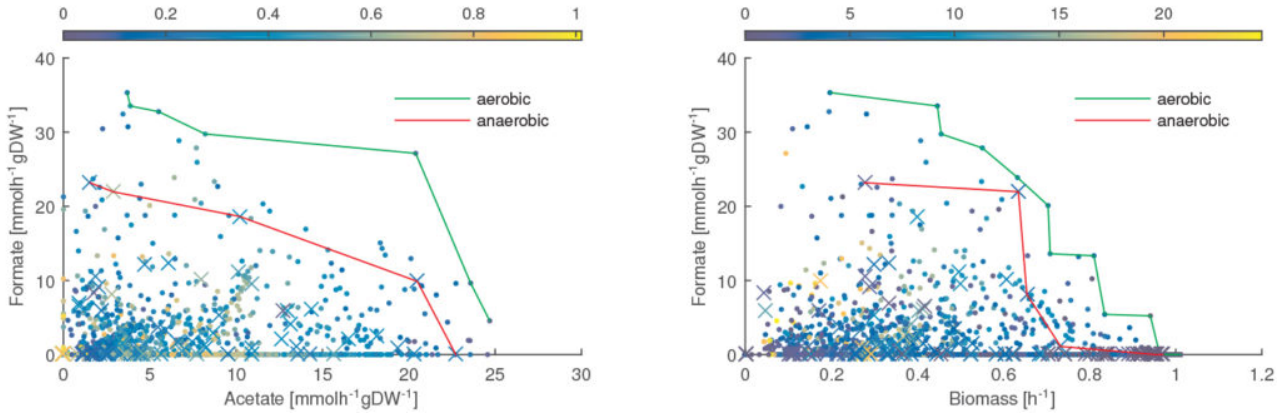

Fig. 2.

More broadly, this kind of computational approach reduces the number of real world experiments that need to be performed, therefore guiding experimental biologists and significantly reducing the cost of biological and medical experiments. For instance, within our method, we are able to predict the growth rate, acetate and formate that *E. coli* will produce in over 4000 different conditions. These are shown in the Figure below, one point per condition, separated into aerobic (with oxygen) and anaerobic (without oxygen) growth (Fig. 2).

Angione C [1], ***Conway M*** [2], ***Lió P*** [2]
[1]*School of Computing - Teesside University, Middlesbrough, UK*
[2]*Computer Laboratory - University of Cambridge, Cambridge, UK*

Publication

Multiplex methods provide effective integration of multi-omic data in genome-scale models.
Angione C, Conway M, Lió P

BMC Bioinformatics. 2016 Mar 2