

Reapplication of SetCoLa Specifications for Biological Networks

Jane Hoffswell, Alan Borning, and Jeffrey Heer

Paul G. Allen School of Computer Science & Engineering, University of Washington

In our paper, we showed how SetCoLa could be used to recreate Cerebral layouts for the TLR4 network. We demonstrate the reapplication of a similar SetCoLa specification across networks extracted from InnateDB. InnateDB [BFL*12] is a public database containing a large quantity of biological information and is integrated with analytic and visualization tools. For example, the database is integrated with Cerebral [BMGK08] to enable visualization of properties such as protein interactions or signaling pathways. The general specification used for each layout is shown in Figure 1. We selected four networks from InnateDB to recreate with the layout. For each one, we apply the same specification and show the result produced with SetCoLa as compared to the visualization produced by Cerebral.

To demonstrate the reapplication of a SetCoLa specification for similar graphs in the same domain, we used four examples from InnateDB: the TLR4 network (Figure 2, [Inn14c]), the DDX58 network (Figure 3, [Inn14a]), the NOD-like signaling pathway (Figure 4, [Inn18]), and the MAPK1 network (Figure 5, [Inn14b]). The full specification and data for each example is included in the other supplementary materials, along with the other specifications from the paper. For each figure, we added custom labels to match the labels produced by Cerebral in order to facilitate the comparison between the two layouts. These examples show that that layout can easily be reapplied across different graphs using the same SetCoLa specification.

While the SetCoLa specification works well for the TLR4 network, DDX58 network, and NOD-like signaling pathway, the specification produces an undesirable result for the MAPK1 network (Figure 5b). This graph produces a much more flattened version of the network because it has over twice the number of nodes as the other networks (e.g., 240 nodes as compared to about 100 in the smaller networks). In this case, the constants used in the specification are not particularly ideal for the larger network. Future work should explore better techniques for applying spacing relative to graph properties rather than constant values. An improved version of the MAPK1 network (with reduced spacing on the nodes in the “Nucleus” layer) is shown in Figure 5c. One of the other main differences between the SetCoLa and Cerebral specifications is the rendering style for the links, which uses a bundled routing style in Cerebral. Such a rendering style could be added to SetCoLa in the future to achieve this effect.

```
1 {
2   "nodes": [...],
3   "links": [...],
4   "guides": [
5     {"name": "left_guide", "x": 250},
6     {"name": "right_guide", "x": 475}
7   ],
8   "constraintDefinitions": [
9     {
10      "name": "boundary",
11      "forEach": [
12        {"constraint": "position", "position": "right", "of": "left_guide", "gap": 10},
13        {"constraint": "position", "position": "left", "of": "right_guide", "gap": 10}
14      ],
15      "sets": [{"partition": "Localization"}],
16      "forEach": [{"constraint": "padding", "amount": 16}]
17    },
18    {
19      "name": "layer",
20      "sets": [{"cellular location"}],
21      "forEach": [
22        {
23          "constraint": "order",
24          "axis": "y",
25          "by": "Localization",
26          "band": 125,
27          "gap": 30,
28          "order": ["Extracellular", "Cell surface", "Plasma membrane", "Cytoplasm", "Nucleus", "Unknown"]
29        }
30      ]
31    }
32  ]
33 }
34 ..
```

Figure 1: The specification reapplied across four biological networks: TLR4 interaction network (Figure 2), DDX58 interaction network (Figure 3), the NOD-like signaling pathway (Figure 4), and MAPK1 interaction network (Figure 5).

References

- [BFL*12] BREUER K., FOROUSHANI A. K., LAIRD M. R., CHEN C., SRIBNAIA A., LO R., WINSOR G. L., HANCOCK R. E., BRINKMAN F. S., LYNN D. J.: Innatedb: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic acids research* 41, D1 (2012), D1228–D1233.
- [BMGK08] BARSKY A., MUNZNER T., GARDY J., KINCAID R.: Cerebral: Visualizing multiple experimental conditions on a graph with biological context. *IEEE transactions on visualization and computer graphics* 14, 6 (2008), 1253–1260.
- [Inn14a] INNATEDB: Homo sapiens gene: Ddx58. <http://www.innatedb.com/getGeneCard.do?id=55854>, 2014. Accessed: 2018-03-12.
- [Inn14b] INNATEDB: Homo sapiens gene: Mapk1. <http://www.innatedb.com/getGeneCard.do?id=2147>, 2014. Accessed: 2018-03-12.
- [Inn14c] INNATEDB: Homo sapiens gene: Tlr4. <http://www.innatedb.com/getGeneCard.do?id=82738>, 2014. Accessed: 2018-03-12.
- [Inn18] INNATEDB: Nod-like receptor signaling pathway. <http://www.innatedb.com/interactionSearch.do?from=pw&exPathwayXref=8112&pathwayFilter=&pathwayXrefDB=&pathwayXref=&listType=interaction&coreInteractors=true>, 2018. Accessed: 2018-03-12.

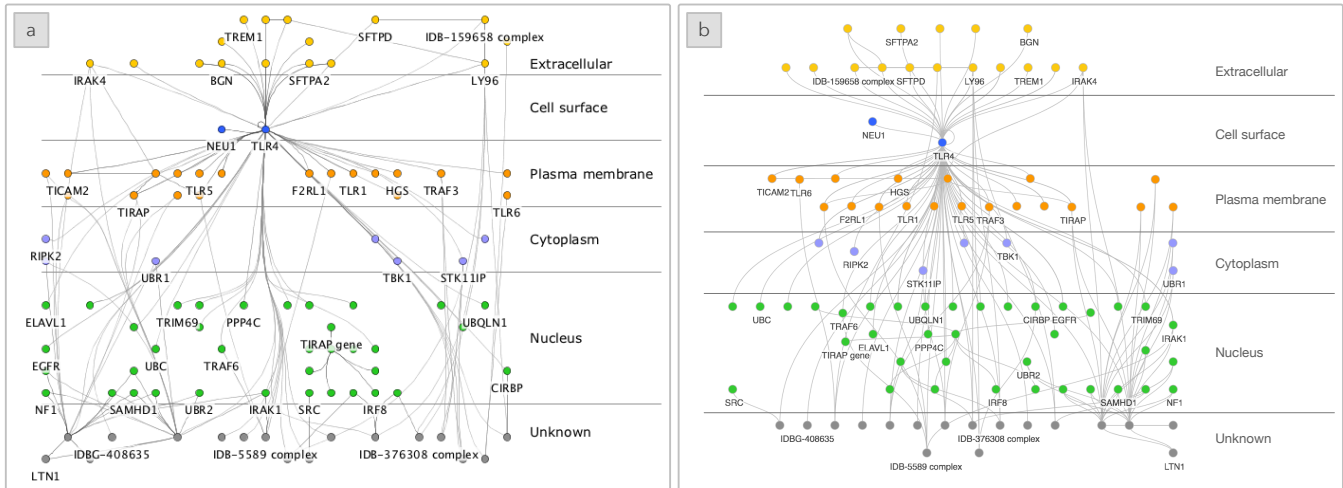


Figure 2: The layout for the TLR4 biological system produced using (a) the Cerebral visualization from InnateDB [Inn14c] as compared to (b) SetCoLa. The layers correspond to the location of the biomolecule within a cell.

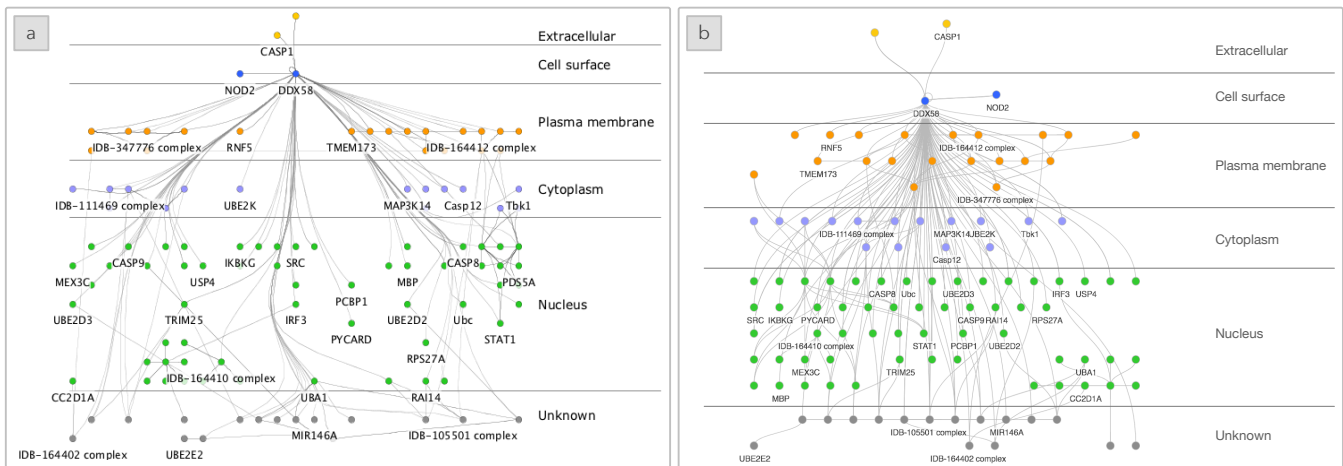


Figure 3: The layout for the DDX58 biological system produced using (a) the Cerebral visualization from InnateDB [Inn14a] as compared to (b) SetCoLa. The layers correspond to the location of the biomolecule within a cell.

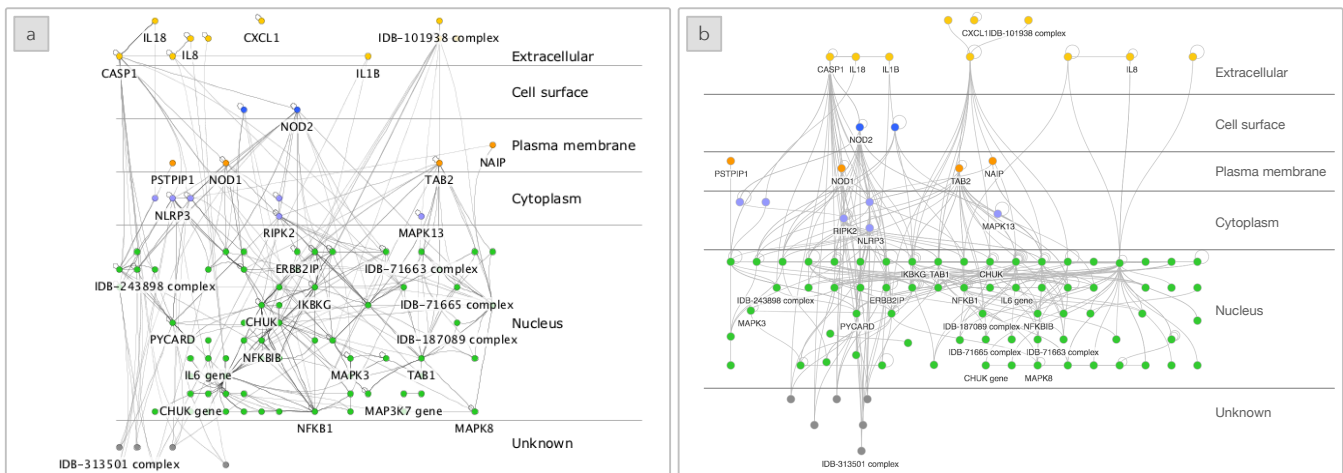


Figure 4: The layout for the NOD-like signaling pathway produced using (a) the Cerebral visualization from InnateDB [Inn18] as compared to (b) SetCoLa. The layers correspond to the location of the biomolecule within a cell.

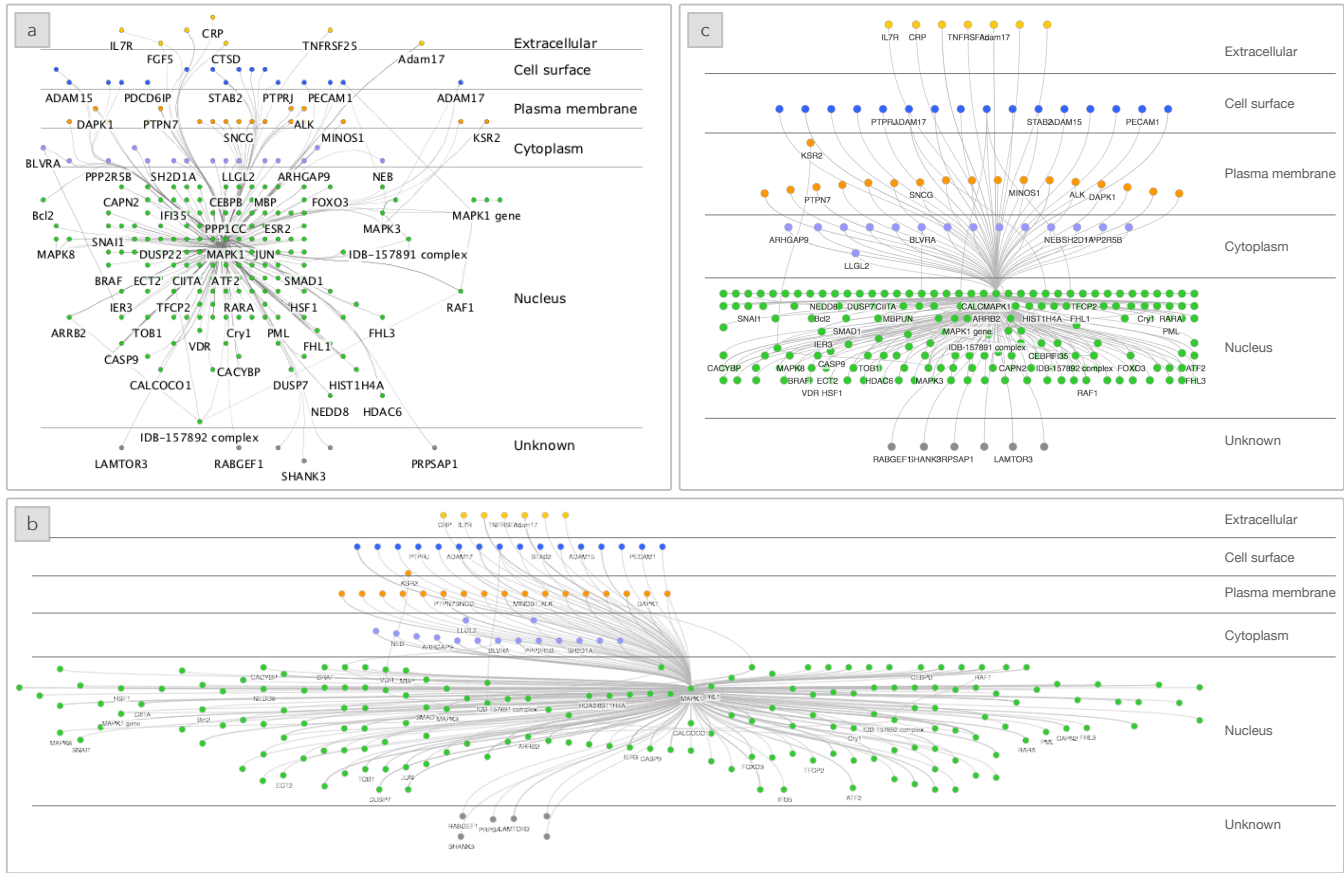


Figure 5: The layout for the MAPK1 biological system produced using (a) the Cerebral visualization from InnateDB [Inn14b] as compared to (b) SetCoLa. The layers correspond to the location of the biomolecule within a cell. The excessive spacing and constraints on the size of each layer produces an undesirable layout for the larger graph as compared to the smaller alternatives (Figures 2, 3, 4). (c) By adding an additional constraint to reduce the spacing, we can achieve a more desirable layout for this graph.