Modularization of Signal Transduction Pathways: detecting the trend of development among various species

Losiana Nayak and Rajat K. De
Machine Intelligence Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata-700108, West Bengal, India
{losiana_t,rajat}@isical.ac.in

Background

- Signal transduction pathways are very complex biochemical networks that contain large number of vital biomolecules and chemical compounds.
- They link, communicate, co-ordinate and cross-talk among themselves to bring forward important cellular functions.
- Dissecting such networks into meaningful partitions via partitioning algorithm(s) will be more appropriate for deciphering the underlying mechanisms that control physiological behavior of cell(s).
- Modularization algorithm is one among the partitioning algorithms that gives better partitions (modules) than the standard graph partitioning and community finding algorithms.

Modularization Algorithm

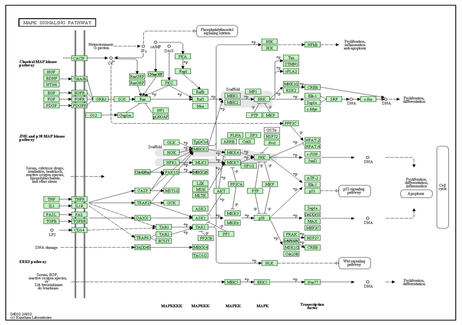
1. ENSURE $E \neq \phi$
2. Find start/central node
   - IF($T_n \leftarrow Max\{T_M\}$)
   - $n \leftarrow$ start point/central node
   - $M_P \leftarrow M_P \cup \{n\}, E \leftarrow E - \{n\}$
   - $k \leftarrow 0$
3. Extend module
   - FOR($k \leftarrow k + 1$)
   - select nodes from N_S and N_P of n and put in M^k
4. Check permanency of nodes
   - IF $N_S \cup N_P \subset M^k$ for a node n^k
   - $E \leftarrow E - \{n^k\}, M_P \leftarrow M_P \cup \{n^k\}$
5. Exclude node
   - IF[T_n^k– number of nodes in M^k related to n^k] $> c$
   - $M^k \leftarrow M^k - \{n^k\}$
   - FOR ($n^{(k-1)} \notin M_P$)
   - $T_{n^{(k-1)}} \leftarrow [(T_{n^{(k-1)}}) - 1]$
6. Build a complete module
   - REPEAT Step 2-4 UNTIL $M^k \subset M_P$
7. Create next module
   - REPEAT Step 1-5 UNTIL$E = \phi$

Algorithm Continued

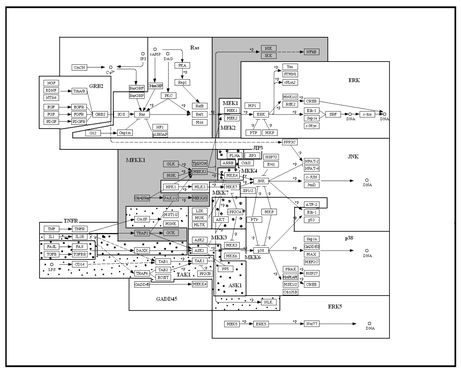
- E: Set of all nodes (representing gene products and chemical compounds) present in a network, where each node must have atleast one relation, i.e., isolated nodes are not included in this set
- M: Set of nodes present in a module (a part of network)
- k: *Extension* index (stage of inclusion of immediate neighbors of nodes in a module)
- c: Complexity level (a value fixed by the user determining the inclusion (exclusion) of nodes in (from) a module)
- M^k: Set of nodes present in a module after kth *extension*
- N_S: Set of succeeding nodes of a given node
- n_s: An individual member of N_S
- N_P: Set of preceding nodes of a given node
- n_p: An individual member of N_P
- r: type of interaction that exists between n_p and n_s; $r = a$ depicts a relation of activation, $r = b$ depicts a relation of binding or association, $r = i$ depicts a relation of inhibition, $r = d$ depicts a relation of indirect effect between nodes n_s and n_p
- The total number of relations with n as either a preceding or succeeding node is given by $T_n = R_{np} + R_{ns}$
- N_R: Set of relations (interactions that can exist between two nodes)
- $n_r = (n_p, n_s, r)$: An individual member of N_R
- R_{np}: Total number of relations that exist with n as the preceding node
- R_{ns}: Total number of relations that exist with n as the succeeding node
- M_P: Set of permanent nodes (nodes having all their relations inside a module)
- Max: A function that detects maximum value among all elements present in a given set

Results on MAPK Signaling Pathway

- 16 modules for $c = 3$.
- Module *GRB2*, *Ras* and *ERK* divide effectively the classic MAPK signaling pathway into 3 parts.
- JNK pathway is divided into 4 modules (*MEKK1*, *MKK4*, *MKK7* and *JNK*).
- Modules *p38*, *ASK1* and *TAK1* counter for p38 pathway except a few small modules.
- Four singleton and two small modules were found.
- The problem of over-splitting is better countered for than other c-values.
- A comparative view among 9 species-specific MAPK signaling pathways showed that the evolutionary trend among them from higher to lower sophistication is human, mouse - rat - cow - dog, pig, and chimpanzee.



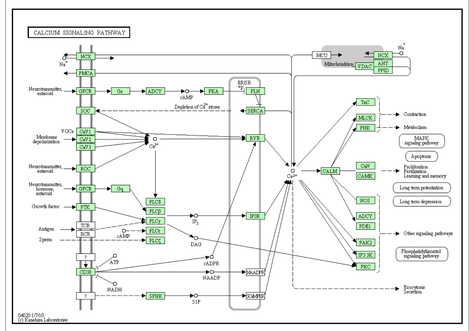
Human MAPK Signaling Pathway



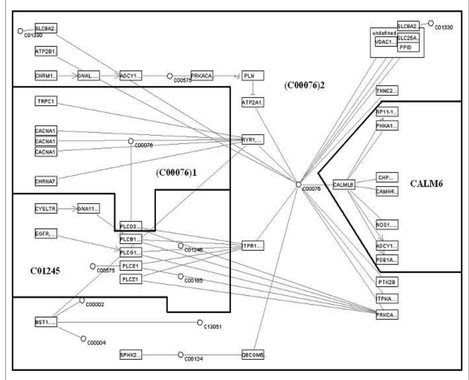
Modularized Human MAPK Signaling Pathway for $c = 3$

Results on Calcium Signaling Pathway

- The human Calcium signaling pathway got modularized into 4 modules ((C00076)2, CALML6, (C00076)1, C01245) for c=3.
- Analysis of the modules obtained from 7 different specie-specific pathways leads us to a conclusion that human, rat and mouse have highly developed calcium signaling mechanisms.
- Cow and pig lie as intermediates.
- Pathways of dog and chimpanzee are very much under developed.



Human Calcium Signaling Pathway



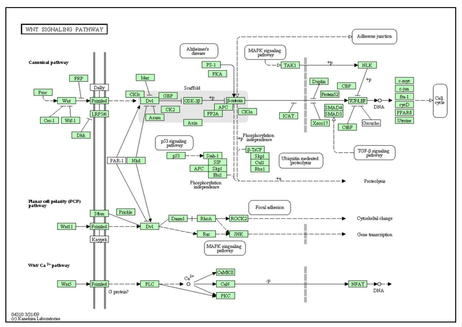
Modularized Human Calcium Signaling Pathway for $c = 3$

Results on Human Wnt Signaling Pathway

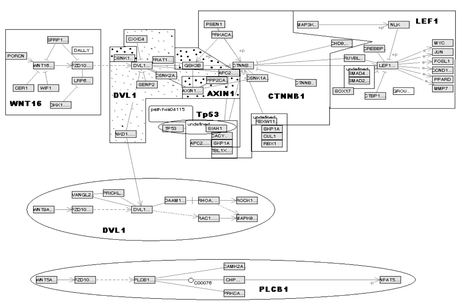
- 8 modules for $c = 3$.
- Module $WNT16$ contains signal initiation steps of the pathway.

Results on Human Wnt Signaling Pathway Continued...

- Modules $(DVL1)1$, $AXIN1$ and $CTNNB1$ account for steps around the PP2AC protein scaffold.
- Module $LEF1$ encloses intra-nuclear canonical Wnt signaling pathway.
- Modules $(DVL1)2$ and $PLCB1$ account for the Planar cell polarity and Wnt/Ca^{2+} pathways respectively.
- No biological significance found for very small module $Tp53$.
- Trend of evolution among 31 different species from higher to lower sophistication: human, mouse - rat - cow, dog, chimpanzee - rhesusmonkey - chicken - opossum - zebra fish - African clawed frog - purple sea urchin - western clawed frog - fruitfly - horse - sea anemone - honey bee - fruitfly related species (dpo) - red flour beetle - yellow fever mosquito - platypus - C. elegance - mosquito - pig - Florida lancelet - sea squirt - fruitfly related species (dan) - filaria, pea aphid - Trichoplax - C. briggsae



Human Wnt Signaling Pathway



Modularized Human Wnt Signaling Pathway for $c = 3$

Acknowledgement: Losiana Nayak is grateful to Council of Scientific and Industrial Research, India, for providing her a Senior Research Fellowship.