

The procedure of evolution modelling of biochemical networks structure

Tatjana Rubina^{1*}

¹Department of Computer Systems, Latvia University of Agriculture, Liela iela 2, LV-3001, Jelgava, Latvia

*Corresponding author

tatjana.rubina@llu.lv

Received: 16 August 2013; accepted: 22 August 2013; published online: 29 August 2013.
This paper has no supplementary material.

Abstract: *The exploration of biochemical networks, such as gene regulation, metabolic, protein interaction and signal transduction networks helps to understand better cellular processes, properties and functions of biological system. An important task of biological systems investigation is exploration of biochemical networks evolution and dynamic changes of their structure under pressure of the mutations and natural selection that are mentioned as the main evolution forces. Proposed network growth models have been used to establish topological properties of biochemical networks, such as scale-free degree distribution, ultra-small-world property, centrality and modularity. But they consider network evolution implicitly, generally and ignore important properties of biological systems. To demonstrate and investigate the evolution course of biochemical networks structure caused by genetic mutations, chosen by natural selection and depending of the biological system properties, evolution models are needed what takes into account these features. In this paper evolution modelling procedure is introduced as well as algorithm of biochemical networks structure that occurs as a result of genetic alterations by pressure of natural selection and takes into account different importance levels of biochemical processes.*

Keywords: Biochemical network, structure evolution, evolution modelling procedure, evolution algorithm, mutation operators, process importance.

1. Introduction

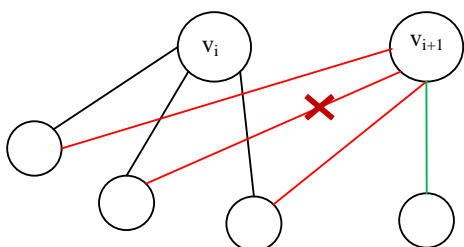
The number of known protein-protein interaction and metabolic data has increased with each completed genome sequencing project in last decade and availability of these data in several species enables comparative analysis to further functional and evolutionary understanding of molecular and cellular processes. The protein interaction and metabolic network data are derived from numerous experiments carried out using different techniques and parameters in various laboratories instead of single large-scale project (Yamada and Bork, 2009). Therefore it is very difficult and time-consuming through laboratory experiments derive all necessary data to build an understanding of biochemical network development, organization and evolution. For this reason the theoretical and computational models, modelling and simulations are needed.

In the last century researchers have begun to study topological properties of biochemical networks from the evolutionary aspects (Chen et al., 2009) with the aim to build an understanding of how networks evolve. That provides insights into the structure and function of biochemical networks. There have been several network growth models proposed to identify the topological features of biochemical networks such as scale-free degree distribution (Barabasi and Albert, 1999; Albert et al., 2000; Barabasi and Oltvai, 2004; Boccaletti, 2006), ultra-small-world property (Watts and Strogatz, 1998; Wagner and Fell, 2001; Albert et al., 2000; Cohen and Havlin, 2003), centrality (Jeong et al., 2001; Fell and Wagner, 2000; Chen et al., 2009), high-degree clustering and modularity (Jeong et al., 2000; Ravasz et al., 2002; Pastor-Satorras et al., 2003; Yook et al., 2004; Li et al., 2006)

assuming that the current topology of a network is formed through a series of network assembly and evolution events (Chen et al., 2009). For example, duplication-mutation (with complement) models (Solé et al., 2002; Vazquēz, 2003; Kwon and Cho, 2007) (Fig. 1), duplication-divergence models (Pastor-Satorras et al., 2003; Wagner, 2003; Farid and Christensen, 2006; Kim and Marcotte, 2008; Hase and Niimura, 2012), duplication-deletion-divergence models (Farid and Christensen, 2006), preferential attachment model (Barabasi and Albert, 1999; Albert et al., 2000; Solé et al., 2002; Wagner, 2003; Kim and Marcotte, 2008; Hase and Niimura, 2012) (Fig. 2), random growing network models (Callaway et al., 2000; Krapivsky et al., 2000; Kwon and Cho, 2007), small world network models (Watts and Strogatz, 1998; Barrat and Weigt, 2000; Kwon and Cho, 2007) and other models.

To explain network evolution, there have been invoked two kinds of processes (Sharan and Ideker, 2006; Yamada and Bork, 2009) that are taken for theoretical basis in developed network growth models: sequence mutations in a gene that modify a gene or its regulation, and gene duplication or deletion. Sequence mutations in a gene results in modifications of the interactions and corresponding protein may gain new connections or lose some of existing connections to other proteins. Sequence mutations can be point mutations, insertions or deletions, or mutations that affect the regulation of a gene, e.g. non-synonymous nucleotide substitution (Noort et al., 2004). Gene duplication implies the addition of a node with links identical to the original node, followed by the functional divergence of some of the redundant links between the two duplicate nodes. The gene deletion corresponds to the

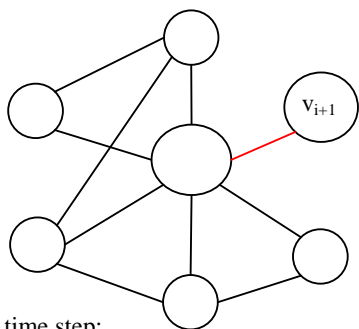
loss of node and all associated links to it. The addition or deletion of link affects the connectivity of a network. As a result of a combination of these processes, extensive network rewiring can be observed when comparing the interaction of homologous proteins (Yamada and Bork, 2009).



At each time step:

- node v_j is chosen uniformly at random
- node $v_{i+1}=v_j'$ is added and connected to all the duplicated node v_j neighbours
- each link of new added node v_{i+1} is removed with probability q
- the new link is added with probability p between new v_{i+1} node and each node that is not a neighbour of the duplicated v_j node

Fig. 1. **Duplication-mutation models.**



At each time step:

- new node v_{i+1} is added
- new link is added between new v_{i+1} node and existing node in the network

Nodes with higher degree are likely to get a new link.

Fig. 2. **Scale-free or preferential attachment models.**

One of the main shortcomings of proposed network growth models is lack of underlying genome and simplified evolution implementation that arise generally in addition of nodes. The most frequent genetic events, such as whole-genome duplication, locally confined gene duplication and retrotransposition (Yamada and Bork, 2009) results in the gene duplication and lead to the node addition. However performed researches suggest that the evolution of links is coupled to evolution of nodes but is much more fine-tuned as links changes over time even if nodes are unaffected. Although quantification of link changes remains difficult while requires sufficient network data in several species which emerge slowly, there are plenty of genetic mechanisms that can easily lead to a link addition or deletion, such as point mutations, alternative splicing and domain accretion, inversion, shuffling and duplication (Yamada and Bork, 2009).

The developed network models consider network evolution implicitly, generally and ignore relevant properties of biological systems. One of such properties is process importance (Rubina and Stalidzans, 2012) which allows separating for systems viability and living quality essential processes, and such processes that cause inessential characteristics and features of biological system. Particular evolutionary constraints can be identified only when taking into account the background of the general properties of biological system processes. To demonstrate and investigate the evolution course of biochemical networks structure caused by genetic mutations, chosen by natural selection and depending on the biological system properties, evolution models are necessary that take into account these features. In this study an algorithm has been proposed to take into account processes of different importance and evolutionary changes in gene level. The algorithm is tested and results are presented.

2. Evolution modelling procedure of biochemical network structure

Various changes are introduced in several genes and a whole genome, in the properties, features and characteristics of a biological system, in an organization and a course of cellular processes. Evolutionary changes of biochemical networks occur in their structure as a result of genome level alterations, while genes define and regulate organization and operation of biochemical networks. It is possible to connect a gene to a network link accordingly to the central dogma if the catalysed enzyme and its expressed gene are known (Fig. 3). Therefore to explore structure evolution of biochemical network we should connect genome sequence to biochemical reactions (Rubina and Stalidzans, 2012).

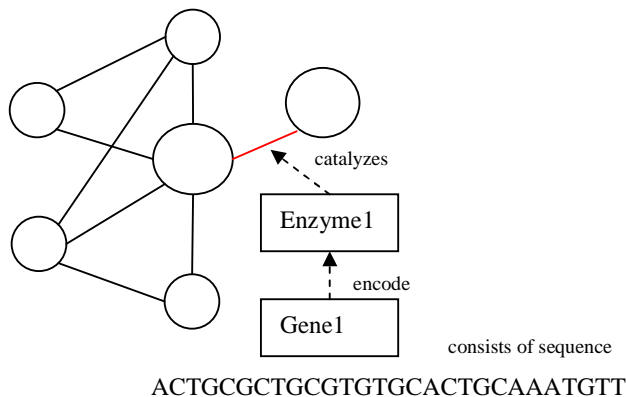


Fig. 3. **Genome and network level data relationship.**

To investigate the changes of biochemical network structure as a result of evolution, here is proposed evolution modelling procedure, which includes six consecutive stages (Fig. 4).

The first stage of procedure is definition of initial network structure data that includes the three main substages: 1) definition of network nodes and links, 2) definition of initial genome, 3) connection of genome to network links. Initial network can be defined manually, entering each network node and link data or can be loaded from several existing models, for example SBML that are available on public databases, such as BioCyc (<http://biocyc.org/>) (Karp et al., 2005), EcoCyc (<http://ecocyc.org/>) (Keseler et al., 2011), KEGG

(<http://www.genome.jp/kegg/>) (Kanehisa and Goto, 2000; Kanehisa et al., 2011), Reactome (<http://www.reactome.org>) (Croft et al., 2011). Initial genome should include gene sequences and information on current gene location in genome, i.e. a chromosome number. Initial artificial gene sequences and a chromosome number can be generated automatically.

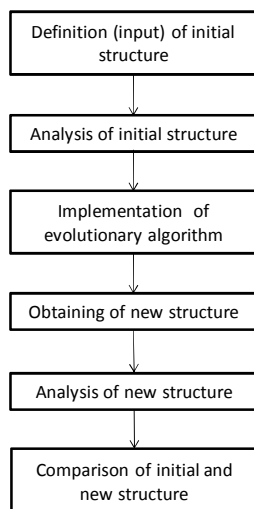


Fig. 4. Evolution modelling procedure.

The second stage of the procedure includes an analysis of initial network structure that should be performed to assess the quality and adequacy of a chosen biochemical network model. During a topological analysis of network structure, different topological parameters can be calculated (Rubina and Stalidzans, 2010) and network motifs can be determined. Analysis results should be stored for further comparison with an acquired structure as a result of evolution.

The third stage of the procedure is implementation of an evolution algorithm that includes two main parts: evolution of underlying genome by pressure of various mutations and structure evolution that depends on occurred alterations of underlying genome comparing it with an initial genome. An evolution algorithm is implemented as an iterative process, where an iteration number corresponds to a number of generations.

A new network structure arises as a result of an evolution algorithm implementation. Newly obtained structures are analysed at the next stage of the procedure to get its topological parameters, such as a number of connected and isolated elements, a number of reactions and links, an average degree, an average number of neighbours, an average path length, an average clustering coefficient.

At the last procedure stage, the new obtained network structure is compared with the initial one with the purpose to evaluate the acquired changes and draw a conclusion of the influence of applied mutations.

3. Evolution modelling algorithm

As a matter of fact, a proposed modelling algorithm for network evolution bases takes different types of genetic mutations and natural selection that is introduced in selection of the next generation offspring. The chosen genetic mutations influence on the underlying genome sequence, but do not change the length of a gene. The precondition of equal gene length is introduced to reduce similarity calculation time.

Biologically observed mutation operators are mapped to the network level changes according to the set of conditions. Implementation of this algorithm is based on the following assumptions and limitations:

- all genes are of equal size, i.e. nucleotide sequence length,
- genetic mutations act upon the underlying nucleotide sequence and do not always have one-to-one correspondence with network level changes,
- structure evolution emerge in links dynamic,
- processes or links in a biological system are not equally important and can be divided into three groups by importance level.

Each particular gene first should be connected to a biochemical network link before starting the evolution process to explore the structure evolution of a biochemical network.

To perform the evolution process of network structure, evolution parameters should be defined at the beginning (Fig. 5). Evolution process will execute N times generating genomes of N generations. In each generation 10 genome copies are created and are subjected to the mutation and natural selection processes. There are selected candidates for the offspring of the next generation from 10 mutated genome copies. From n possible candidates only one genome is selected as an offspring accordingly to the principles of natural selection. The number of genome copies can be changed, but smaller number of copies decreases the number of potential candidates of the next generation offspring and it will have an impact on the evolution results. Than it is less candidates, than is less opportunity to choose the best one.

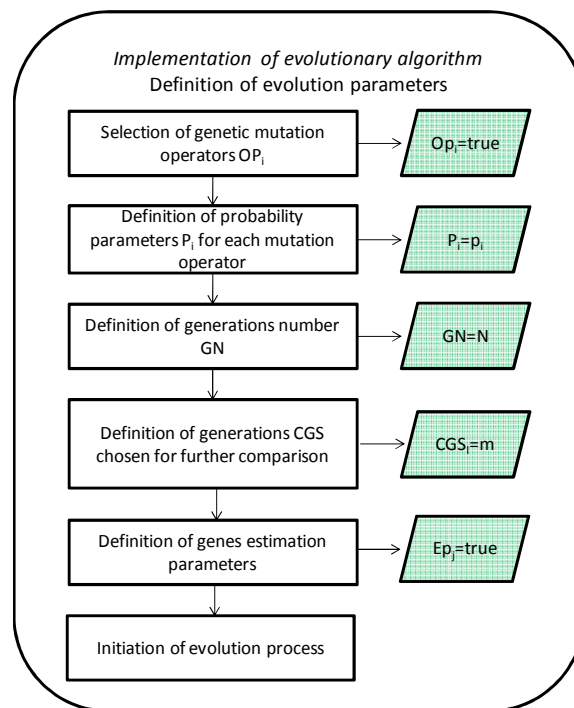


Fig. 5. Definition of evolution parameters.

An evolution algorithm consists of two main stages: genome evolution and structure evolution. Evolution process (Fig. 6) repeats N times, i.e. generations. At each generation 10 copies of the previous generation or initial genome are created

and are subjected to a mutation process. Genome evolution process can be provoked by following mutation operators which rates are set by the user:

- point mutation – an alteration of a single nucleotide in the gene sequence;
- nucleotide inversion – a reinsertion of the gene segment in the same place but in reverse direction;
- duplication – a duplication of gene sequence that results in a link addition;
- deletion – a deletion of the gene sequence that results in a link removing;
- inversion – a reinsertion of the gene sequence in reverse direction;
- translocation – an exchange of a chromosomal segments between two nonhomologous chromosomes.

To represent the influence of various types of mutations that occur in nature, here is offered to establish concordance coefficients for each gene that determine on which cases network level changes should be introduced. A concordance coefficient of each separate gene is calculated comparing the acquired gene sequence by nucleotide triplets with the initial benchmark-genome corresponding to gene sequence (concordance coefficient Rgk_i comparing to the initial gene sequence) and all other genes sequences (concordance

coefficients Qgk_{ij} comparing to all the other genes). Concordance coefficient can take values between 0 and 1 including these thresholds. In case when some of $Qgk_{ij} > Rgk_i$, and one of below accounted conditions are true, it is assumed that i th gene has begun to function as a j -th gene that results in i th link deletion and j -th link addition in the network structure.

The higher is essentiality level of process, the stronger requirements are defined for gene sequence changes and concordance coefficients values. We assume that regulating genes of vital processes (1.essentiality level processes) can mutate nor more than for 30%, it means, to differ from the initial gene sequence, otherwise, an adjustable link in the network structure is deleted. Regulating genes of quality processes (2.essentiality level) can mutate to 50%. In a case as regulating gene sequence changes from 30 to 50%, than intensity parameter of adjustable link or node is reduced. If the regulating gene sequence mutates more than for 50%, the adjustable link is deleted. Insignificant genes may change up to 80%, otherwise an adjustable link is deleted. In a case as regulating gene sequence changes from 50 to 80%, than an intensity parameter of an adjustable link or a node is reduced. The mutation limits of regulating genes of different essentiality level processes can be adjusted to a particular researcher's opinion accordingly to the research purpose.

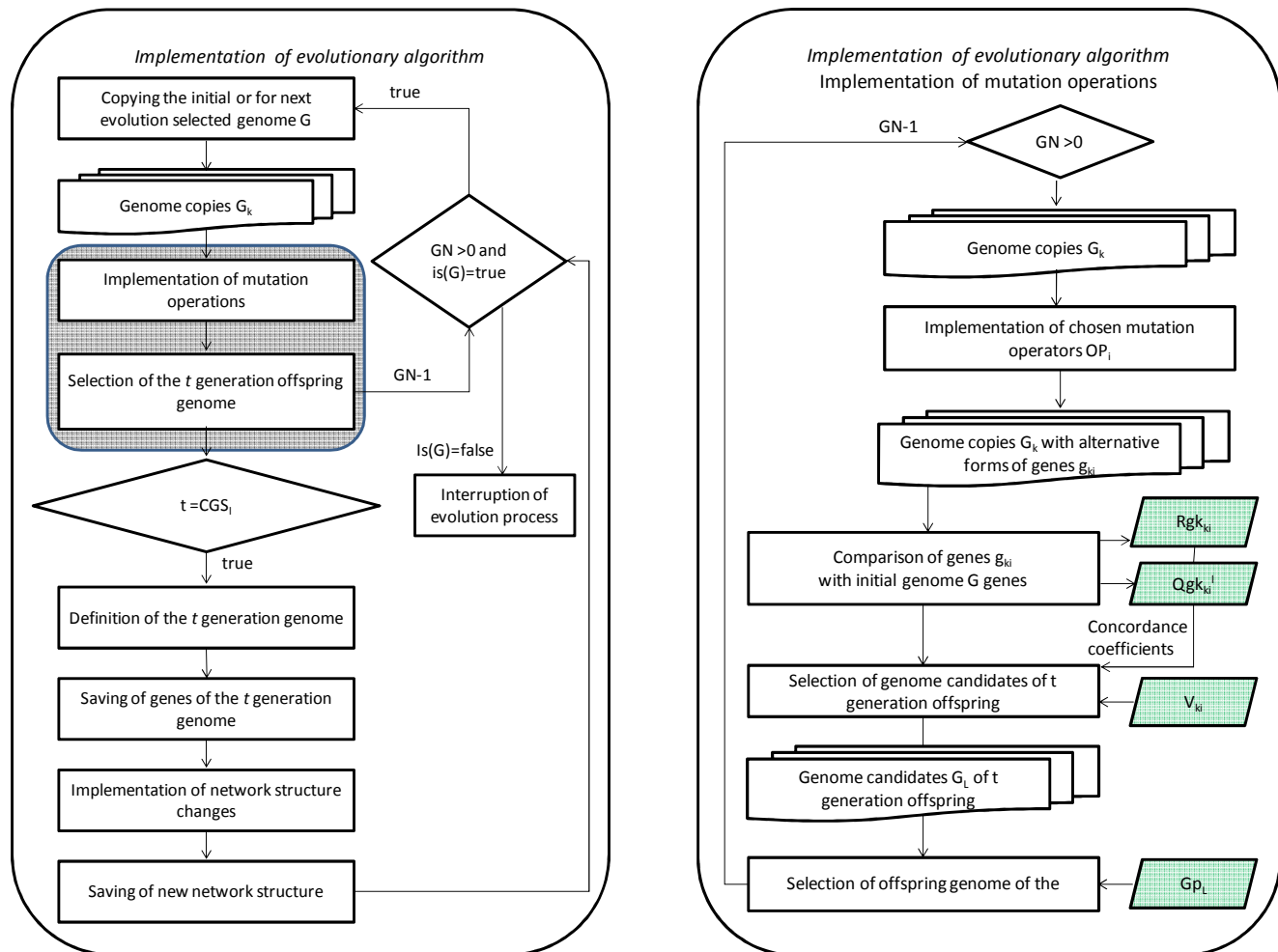


Fig. 6. Implementation of evolution algorithm.

All the mutated genome copies are estimated by each gene concordance coefficient comparing it to the initial benchmark-genome and a set of conditions. Comparing genes of mutated genome copies the emergence of the missense and nonsense mutations can be checked:

- missense mutation – checking for single nucleotide change that results in a codon which codes for a different amino acid;
- nonsense mutation – checking for single nucleotide change that results in a stop codon. If nonsense mutation occurs than the concordance coefficient of corresponding gene is reduced by half, ie. $\frac{Rgk_i}{2}$.

From 10 mutated genome copies, there are selected n candidates which probability ratio being chosen for the offspring of the next generation depends on the concordance coefficient of separate genome $gRgk$:

$$gRgk = \frac{\sum_{i=1}^k Rgk_i}{k} \tag{1}$$

where Rgk_i – a concordance coefficient of an i th gene comparing it with the initial gene sequence,
 k – a number of genes in a genome.

A mutated genome copy cannot be selected as candidate of the next generation offspring, if:

- some of the vital genes are mutated more than for 30% in case of vital and mixed network and/or
- all the quality genes are mutated more than for 50% in case of quality or mixed network.

For the offspring of the next generation there is only one candidate chosen from n selected candidates. According to the concordance coefficient of each individual gene of the selected offspring genome, the network structure changes are generated (Fig. 7) that can be the following: link addition, link deletion or decreasing/increasing of the link intensity.

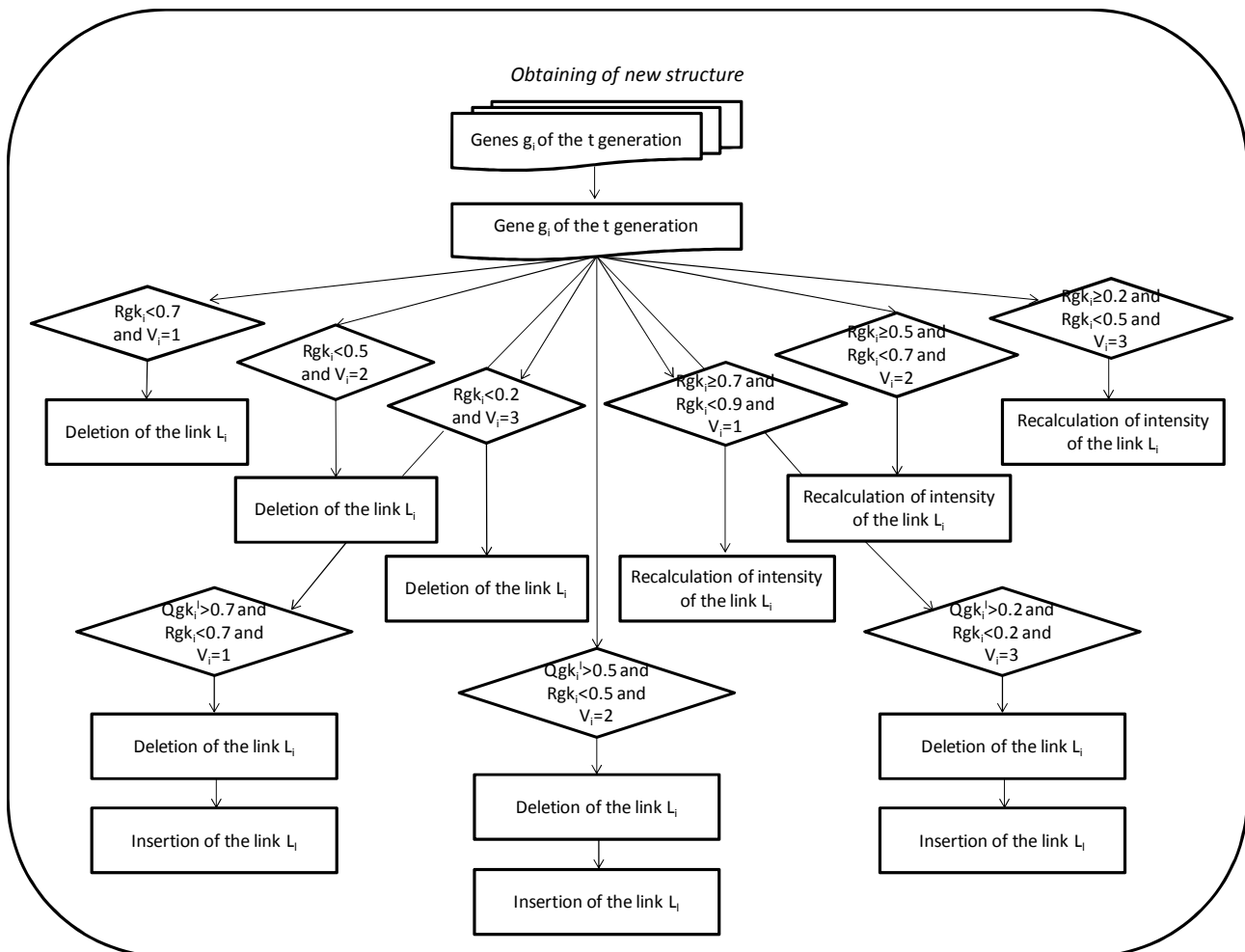


Fig. 7. Structure changes.

Under the influence of mutations, the protein amount that is introduced by a gene can be reduced. It can influence a rate of chemical reaction to, in which protein takes part. For this reason the intensity parameter for network links is entered. The intensity parameter in case of different importance processes can change value by different statistical laws:

- vital processes or links intensity parameter changes can be described by a power law,
- quality processes or links intensity parameter changes - by a linear law, and
- insignificant processes or links intensity parameter changes can be described by a polynomial law.

The link addition and deletion displays only the marginal network structure states. Hence the introduction of links intensity parameter helps to bring out and mark out the intermediate states of the network.

4. Application of the algorithm

Proposed evolution algorithm of biochemical network structure is successfully implemented in modelling and analysis tool BINESA (Biochemical NETwork Structure Analyser) and all the stages of proposed procedure of biochemical network structure can be performed in this tool.

To demonstrate results of the network structure evolution, that are obtained using software tool BINESA, a medium-scale model of *Z.mobilis* central metabolism (Pentjuss et al., 2013) is used. The model includes 81 metabolites, 96 reactions (187 links) and has average degree - 3.68, average in-degree - 1.83 and out-degree - 1.85, average number of neighbours is 5.08 and average clustering coefficient is 0.11. Next example demonstrates the influence of inversion on the structure of mixed network.

In this example, the influence of the inversions on the *Z.mobilis* network structure is considered, which includes 16 vital, 18 qualitative and 62 insignificant reactions. The importance of reactions was defined accordingly to the produced ATP quantity.

To get the evolution results, there were 13 simulations conducted with 10 experiments at each set. The parameters of a simulation were: point mutation and nucleotide inversion probability 10^{-7} , inversion probability rate within [0.04, 1], missense and nonsense mutations were checked. The evolution process in all the experiments was interrupted when there was no candidate for the next generation offspring, i.e. all the mutated genome copies contained at least one defective vital gene or no qualitative genes which mutated for more than 30 % or 50% accordingly. When the inversion probability increases the average concordance coefficient values decrease and a greater number of genes is damaged.

At the Fig. 8 are plotted average viability duration (2) values and its standard deviation (3). The viability duration increases when the inversion probability decreases by a power law, but its standard deviation values increase vastly to.

$$aVD = \frac{\sum_{i=1}^{10} VD_i}{10} \quad (2)$$

$$Std = \sqrt{\frac{\sum_{i=1}^{10} (VD_i - aVD)^2}{10 - 1}} \quad (3)$$

where VD_i – a viability duration, i.e. the number of the last surviving generation,
i - the experiment number in the experiments set.

Fig. 9 demonstrates changes of the average number of reactions where the number of insignificant and quality reactions decrease (corresponding genes were deleted or corrupted), when the inversion probability rate is less than or greater than 20%. But the number of vital reactions remains increasing (corresponding genes mainly were not affected), when the inversion probability decreases.

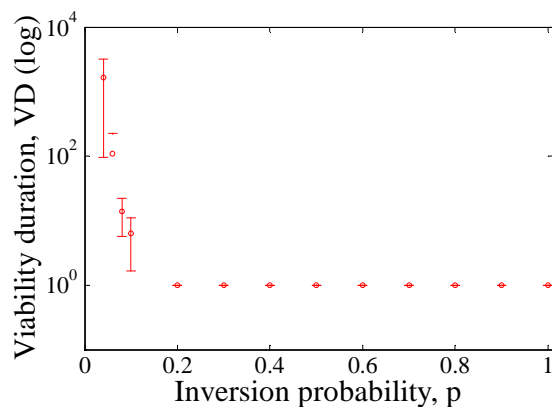


Fig. 8. Average viability duration dependence on inversion probability rate.

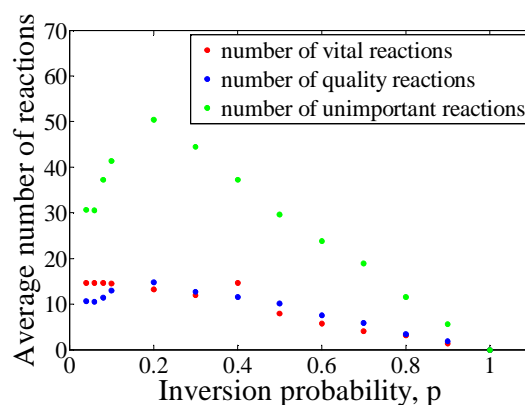


Fig. 9. Average number of reactions dependence on inversion probability rate.

5. Conclusion

An algorithm is offered for evolutionary modelling of biochemical network structure that provides possibility to connect a genome to the network links and to generate network level changes based on alterations that occur during the evolution process, implementing genetic mutation operators at the genome sequence. The limitation of algorithm is that only mutations that do not change the length of gene are implemented. The algorithm takes into account different importance of biochemical processes and impact of their damage by mutations on the viability. The proposed algorithm can be used also for dynamic exploration of biochemical network during the evolution.

The produced experiments on the software BINESA where the algorithm is implemented demonstrate dependence of the number of network reactions, the viability duration of the network, the genome concordance coefficient and the topological parameters of the network structure on the probability of different types of mutations in a *Zymomonas mobilis* central metabolism network.

Acknowledgements

This work and academic study is funded by a financial support from the European Social Fund within the projects “Support for doctoral studies in LUA” /2009/0180/1DP/1.1.2.1.2/ 09/IPIA/VIAA/017” agreement Nr. 04.4-08/EF2.PD.68”.

References

- Albert, R., Jeong, H., Barabasi, A.L. (2000), "Error and attack tolerance of complex networks," *Nature*, Vol. 406, pp.378-382. <http://dx.doi.org/10.1038/35019019>
- Barabasi, A.L. and Albert, R. (1999), "Emergence of scaling in random networks," *Science*, Vol. 286, pp.509-512. <http://dx.doi.org/10.1126/science.286.5439.509>
- Barabási, A.L. and Oltvai, Z.N. (2004), "Network biology: understanding the cell's functional organization," *Nature reviews. Genetics*, Vol. 5 No. 2, pp. 101–13. <http://dx.doi.org/10.1038/nrg1272>
- Barrat, A. and Weigt, M. (2000), "On the properties of small-world network models," *The European Physical Journal B*, Vol. 13, pp.547-560. <http://dx.doi.org/10.1007/s100510050067>
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D. (2006), "Complex networks: Structure and dynamics," *Physics Reports*, Vol. 424 No. 4-5, pp. 175–308. <http://dx.doi.org/10.1016/j.physrep.2005.10.009>
- Callaway, D.S., Hopcroft, J.E., Kleinberg, J.M., Newman, E.J., Strogatz, S.H. (2001), "Are randomly grown graphs really random?," *Physical Review E*, Vol. 64(4), pp.1-7. <http://dx.doi.org/10.1103/PhysRevE.64.041902>
- Chen, L., Wang, R.-S., Zhang, X.-S. (2009), "Biomolecular networks: Methods and Applications in System Biology," *Jersey, John Wiley & Sons; Inc.*; Hoboken; New.
- Cohen, R. and Havlin, S. (2003), "Scale-free networks are ultra small," *Physical Review Letters*, Vol. 90, No. 5, 058701, p.1-4.
- Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., Caudy, M. Et al (2011), "Reactome: a database of reactions, pathways and biological processes," *Nucleic Acids Research*, Vol. 39 (Database issue), D691-D697. <http://dx.doi.org/10.1093/nar/gkq1018>
- Farid, N. and Christensen, K. (2006), "Evolving networks through deletion and duplication," *New Journal of Physics*, Vol. 8, No. 9, p.1.17. [10.1088/1367-2630/8/9/212](http://dx.doi.org/10.1088/1367-2630/8/9/212)
- Fell, D.A. and Wagner, A. (2000), "The small world of metabolism," *Nature Biotechnology*, Vol. 18, pp.1121-1122. <http://dx.doi.org/10.1038/81025>
- Hase, T. and Niimura, Y. (2012), "Protein-Protein Interaction Networks: Structures, Evolution, and Application to Drug Design," *Protein-Protein Interactions – Computational and Experimental Tools*, pp.405-426.
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabasi, A.L. (2000), "The large-scale organization of metabolic networks," *Nature*, Vol. 407, pp.651-654. <http://dx.doi.org/10.1038/35036627>
- Jeong, H., Mason, S.P., Barabasi, A.L., Oltvai, Z.N., (2001), "Lethality and centrality in protein networks," *Nature*, Vol. 411, pp.41-42. <http://dx.doi.org/10.1038/35075138>
- Kanehisa, M. and Goto, S. (2000), "KEGG: Kyoto Encyclopedia of Genes and Genomes," *Nucleic Acids Research*, Vol. 28, pp.27–30. <http://dx.doi.org/10.1093/nar/28.1.27>
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M. (2011), "KEGG for integration and interpretation of large-scale molecular data sets," *Nucleic Acids Research*, Vol. 40 (Database issue), D109–D114.
- Karp, P. D., Ouzounis, C. a, Moore-Kochlacs, C., Goldovsky, L., Kaipa, P., Ahrén, D., Tsoka, S., et al. (2005), "Expansion of the BioCyc collection of pathway/genome databases to 160 genomes.," *Nucleic acids research*, Vol. 33, pp.6083–6089. <http://dx.doi.org/10.1093/nar/gki892>
- Keseler, I. M., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muñiz-Rascado, L., Bonavides-Martinez, C., et al. (2011), "EcoCyc: a comprehensive database of Escherichia coli biology," *Nucleic acids research*, Vol. 39 (Database issue), D583–D590. <http://dx.doi.org/10.1093/nar/gkq1143>
- Kim, W.K. and Marcotte, E.M. (2008), "Age-dependent evolution of the yeast protein interaction network suggests a limited role of gene duplication and divergence," *PLoS Computational Biology*, Vol. 4, Issue 11, e1000232, pp.1.10. [10.1371/journal.pcbi.1000232](http://dx.doi.org/10.1371/journal.pcbi.1000232)
- Krapivsky, P.L., Redner, S., Leyvraz, F. (2000), "Connectivity of Growing Random Networks," *Physical Review Letters*, Vol. 85, Issue 21, pp.4629-4632. <http://dx.doi.org/10.1103/PhysRevLett.85.4629>
- Kwon, Y.-K. and Cho, K.-H. (2007), "Analysis of feedback loops and robustness in network evolution based on Boolean models," *BMC Bioinformatics*, Vol. 8, p. 430. <http://dx.doi.org/10.1186/1471-2105-8-430>
- Li, D., Li, J., Quyang, S., Wang, J., Wu, S., Wan, P., Zhu, Y. et al (2006), "Protein interaction networks of Saccharomyces cerevisia, Caenorhabditis elegance and Drosophila melanogaster: Large-scale organization and robustness," *Proteomics*, Vol. 6, pp.456-461. <http://dx.doi.org/10.1002/pmic.200500228>
- Von Noort, V., Snel, B., Huynen, M.A. (2004), "The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model," *EMBO Reports*, Vol. 5, pp.280-284. <http://dx.doi.org/10.1038/sj.embor.7400090>
- Pastor-Satorras, R., Smith, E., Solé, R.V. (2003), "Evolving protein interaction networks through gene duplication," *Journal of Theoretical Biology*, Vol. 222, p.199-210. [http://dx.doi.org/10.1016/S0022-5193\(03\)00028-6](http://dx.doi.org/10.1016/S0022-5193(03)00028-6)
- Pentjuss A., Odzina I., Kostromins A., Fell D.A., Stalidzans E., Kalnenieks U. (2013), "Biotechnological potential of respiring Zymomonas mobilis: A stoichiometric analysis of its central metabolism," *Journal of Biotechnology*, Vol. 165, pp.1-10. <http://dx.doi.org/10.1016/j.jbiotec.2013.02.014>
- Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N., Barabasi, A.-L. (2002), "Hierarchical organization of modularity in metabolic networks," *Science*, Vol. 297, pp.1551-1555. <http://dx.doi.org/10.1126/science.1073374>
- Rubina, T. and Stalidzans, E. (2012), "Evolution modeling algorithm of biochemical networks. 10-th Industrial Simulation Conference," *ISC'2012. A publication of EUROSIS*, ISBN 978-90-77381.71.7, 4-6.06.2012, Czech Republic, Brno, p.24-30.
- Rubina, T. and Stalidzans, E. (2010), "Topological features and parameters of biochemical network structure," *ISC'2010, 8-th Industrial Simulation Conference. A publication of EUROSIS*, ISBN 978-90-77381.5-57, Hungary, Budapest, p.228-236.
- Sharan, R. and Ideker, T. (2006) "Modeling cellular machinery through biological network comparison," *Nature Biotechnology*, Vol. 24, pp.427-433. <http://dx.doi.org/10.1038/nbt1196>
- Solé, R.V., Pastor-Satorras, R., Smith, E., Kepler, T. (2002), "A model of large-scale proteome evolution," *Adv. Complex. Systems*, Vol. 5, pp.43–54. <http://dx.doi.org/10.1142/S021952590200047X>
- Vazquēz, A. (2003), "Growing network with local rules: preferential attachment, clustering hierarchy, and degree correlations," *Physical Review E* 67, 056104-1 - 056104-15.
- Wagner, A. (2003), "How the global structure of protein interaction networks evolves," *Proceedings of the Royal Society B*, Vol. 270, p.457-466. <http://dx.doi.org/10.1098/rspb.2002.2269>
- Wagner, A. and Fell, D. (2001), "The small world inside large metabolic networks," *Proceedings R. Soc. Lond B*, Vol. 268, No. 1478, p.1803-1810. <http://dx.doi.org/10.1098/rspb.2001.1711>
- Watts, D.J and Strogatz, S.H. (1998), "Collective dynamics of 'small-world' networks," *Nature*, Vol. 393, pp. 440–442. <http://dx.doi.org/10.1038/30918>
- Yamada, T. and Bork, P. (2009), "Evolution of biomolecular networks: lessons from metabolic and protein interactions," *Nature reviews. Molecular cell biology*, Vol. 10 No. 11, pp. 791–803. <http://dx.doi.org/10.1038/nrm2787>
- Yook, S.-H., Oltvai, Z.N., Barabasi, A.L. (2004), "Functional and topological characterization of protein interaction networks," *Proteomics*, Vol., pp.928-942.