SciencePG
Science Publishing Group

Review Article

Input Dataset Survey of *In-Silico* Tools for Inference and Visualization of Gene Regulatory Networks (GRN)

Taiwo Adigun[1, 3, *], Angela Makolo[2, 3], Segun Fatumo[4]

[1]Centre for Systems and Information Services, Covenant University, Ota, Nigeria

[2]Department of Computer Science, University of Ibadan, Ibadan, Nigeria

[3]Bioinformatics Research Group (UNIBReG), University of Ibadan, Ibadan, Nigeria

[4]H3Africa Bioinformatics Network (H3ABioNet) Node, National Biotechnology Development Agency (NABDA), Federal Ministry of Science and Technology (FMST), Abuja, Nigeria

Email address:
taiwofisayo2002@yahoo.com (T. Adigun), aumakolo@gmail.com (A. Makolo), segunfatumo@gmail.com (S. Fatumo)

Abstract: Understanding Gene Regulatory Network (GRN) is considered to be the fundamental approach to many biological questions, and the input dataset performs a crucial role in investigating and visualizing the gene regulatory network [5, 14, 17, 23, 34, 37, 40, 41, 44, 45]. Several software tools [2, 5, 7, 10, 11, 14, 21, 22, 25, 31-33, 37, 38, 40, 41, 44] have recently been developed for GRN inference, where some are designed for a particular dataset, an organism or a particular diseased cell. The questions that prompted this review are; what is (are) the kind of omic data needed to construct a GRN? Is there any peculiar property attached to a GRN of a particular data? And, could there be an integration of data from various omic experiments in form of a knowledge base? The input dataset for GRN are transcriptome information which is analyzed comprehensively including the two major technologies (sources) that produce them. We consider four omic datasets and two of their sources for the purpose of this review. The biological data source technologies are hybridization-based, and sequence-based. Dataset from microarray and ChIP-Chip experiments are hybridization-based while RNA-seq and ChIP-seq are sequence-based. Software tools published on Omic Tool website (http://omictools.com/gene-regulatory-networks-c435-p1.html) are analyzed for this review. However, the major disparity is whether the dataset is ChIP-X (ChIP-Chip and ChIP-seq) or expression (Microarray and RNA-seq) dataset not whether the source is from hybridization-based or sequence-based. Moreover, ChIP-X dataset gives more opportunity to investigate more biological problems. The importance of gene regulatory network suggests a GRN software template, which contains all the additional data from ChIP-X experiment and a knowledge base of biological prior knowledge, including integration of data from different omic datasets as a single knowledge base.

Keywords: In-Silico, Hybridization, Transcriptome, Microarray, ChIP-X, Epigenetics, Hybridization-Based, Next Generation Sequencing (NGS)

1. Introduction

Investigating gene regulatory network is an approach in bioinformatics to study the interactions of genetic materials (genes, proteins, enzymes, ligands etc), and also with the cellular components. Gene regulatory networks inference is a quickly evolving field, with new developments and algorithms being published almost daily. Modeling GRN with computational techniques involves development of a virtual biological cell that represents the dynamism of interactions and reactions among cellular components.

Development and functioning of organisms' cells emerge from interactions in genetic regulatory networks [16] and the regulation of gene expression is achieved through the interactions between DNA, RNA, proteins, and small molecules. This regulatory system can be described by the structure of network called genetic regulatory network (GRN) [35]. Several mathematical and computational models have been developed to analyze the gene regulatory networks and metabolic networks of different cells in different disease traits

especially [4, 6, 9, 12, 15, 16, 35, 43, 44,].

Various mathematical and computational approaches have been used to infer gene regulatory network, and this has resulted into development of different *In-Silico* tools for the reconstruction and visualization of gene regulatory networks [10, 15, 20, 26, 35, 43, 45].Classifications of these computational techniques have been reviewed in several articles based on different criteria. In [16], general review of existing models was performed based on a number of dimensions. The approaches were compared on whether the models were discrete or continuous, static or dynamic, deterministic or stochastic, and qualitative or quantitative. [36] classified the computational models of gene regulatory networks into four classes, (i) logical models, (ii) continuous models, (iii) single molecule models and (iv) hybrid models. In yet another review by [27], network inference algorithms were categorized based on their major properties and they are: (i) the underlying method, (ii) the result, (iii) the directionality of interactions, (iv) the consideration of dynamics (v) the integration of prior knowledge (PK) (vi) non-linearity or linearity, (vii) the explicit consideration of stimulation, (vii) the consideration of stochastics and application of probabilities, (viii) the network size, (ix) the number of required data and (x) the availability as a software tool.

Some of the tools for the purpose of investigating and visualizing gene regulatory networks were presented in [24] which include LegumeGRN, geWorkbench, GENeVis, Cytoscape, NetBioV, FastMEDUSA etc. The weaknesses and strengths of these tools were discussed based on the study performed on their features. In this paper, we perform a comprehensive analysis of GRN software tools based on their various input datasets. We are able to analyze various kinds of input datasets to visualize gene regulatory networks, and investigate peculiar properties evolving from a GRN by using a particular datasets gotten from a particular source. We also analyze tools that are using additional parameters apart from the major dataset, and the added features of the GRN as a result of these additional parameters.

The paper is arranged as follow: section 2 describes the various input datasets to visualize GRN. The GRN visualization and inference tools are analyzed in details in section 3 to reveal the build-up of each tool around a particular dataset. Section 4 presents the conclusion while suggestions are outlined in the last section.

2. Grn Input Data

The analysis of large datasets of information derived from various biological experiments plays a vital role in functional genomics, and a good inference of gene regulatory network is one of these analyses. The availability of high-throughput technologies that allow measuring simultaneously expression of thousands of genes, have given rise to different kinds of genomic datasets used to visualize GRN [22], and majority of GRN software tools were built around a particular dataset. These datasets used for the *in-silico* inference and visualization of gene regulatory networks are transcriptome

information, and the transcripts are acquired and quantified through two major technologies; (i) hybridization array and (ii) sequence-based approaches. Transcriptome gives the complete set of transcripts in a cell and their quantity to analyze the functional constituent of the cell [39]. The transcriptome information produced by these technologies are as follow;

a) mRNA

Messenger RNA is the major RNA molecule produced from the transcription stage of gene expression, that is, conversion of DNA to RNA molecules. It carries genetic information from DNA to the ribosome, which specifies the amino acid sequence of the corresponding protein. mRNAs are arranged into codons consisting of three bases each, and each codon encodes for a specific amino acid, except the stop codons, which terminate protein synthesis. mRNA is a sequence of nucleotides [28].

b) microRNA (miRNA)

It is a conserved class of small noncoding RNAs, whose function is RNA silencing and post-transcriptional regulation of gene expression [30]. MicroRNAs (miRNAs) are about 22nt long that are processed by Dicer from precursors with a characteristic hairpin secondary structure [3].

c) tRNA

This RNA molecules called a transfer RNA is typically 76 to 90 nucleotides in length, and serves as the carrier of amino acids to the ribosomes, so that the ribosomes can put this amino acid on the protein that is being synthesized as an elongating chain of amino acid residues, using the information on the mRNA to determine which amino acid should be put on next. Each type of amino acid has its own type of tRNA, which performs the binding and transporting the amino acid to the growing end of a polypeptide chain if the next code word on mRNA calls for it [28].

d) rRNA

The ribosomal RNA is an important part of ribosome in the cell. It contains about 60% of ribosome complexes whereby the rest 40% is protein molecules. A Ribosome is divided into a large and small subunit, each of which contains its own rRNA molecule or molecules [28]. The nucleotide sequence of rRNA is highly complex depending on whether it is eukaryotic rRNA or prokaryotic rRNA, large subunit or small subunit. The basic function of ribosomes is translation of mRNA into protein by linking amino acids together. The small subunit of rRNA reads the order of amino acids while linking of the amino acids together is the function of the rRNA in the large subunit of the ribosome.

e) other non-coding RNAs

These include snoRNAs, microRNAs, siRNAs, snRNAs, exRNAs, piRNAs and the long ncRNAs

These datasets are the major input dataset to infer and visualize GRNs in various literatures [10, 11, 13, 14, 19, 21, 22, 25, 26, 29, 31, 34, 37, 38, 40, 41], which are generally known as gene expression data. One common attribute of these datasets is their constituent, which is nucleotide bases but they are of different lengths and different functions they perform. So, it is crucial to know in advance the structure and

function(s) of a particular data before using such in inferring and visualizing GRN. In addition with these major datasets are other datasets that act as additional input parameters in inferring GRNs like Transcriptional Start Sites (TSSs), Transcriptional Factors (TFs), promoter elements and binding signals. The use and combination of these datasets are analyzed in the following section, including the influence of the input dataset on the outcome. A number of organisms and disease traits like cancer and cardiovascular diseases have been profiled and are available in different public databases like Gene Expression Omnibus (GEO) repository.

A particular input dataset for *in-silico* investigation of regulatory networks has a source, and the sources have been categorized into two, the hybridization-based and the sequence-based technologies [29]. The primary aim of these two technologies is to get the transcriptome information of different genes at the same time. Hybridization-Based

Technology is a technique for identifying, among a sample of many different DNA fragments, the fragment(s) containing a particular nucleotide sequence. This is achieved by combining two complementary single-stranded DNA molecules and allowing them to form a single double-stranded molecule through base pairing. The hybridization-based techniques widely used for GRN analysis and visualization are DNA microarray and ChIP (Chromatin immunoprecipitation) Microarrays (ChIP-chip). Whereas, sequence-based technologies that are widely used for GRN analysis and visualization are RNA-seq and ChIP-seq (Chromatin Immunoprecipitation Sequencing). These are parts of Next Generation Sequencing techniques. Table 1 shows the properties of these techniques [17, 34, 18]. Having full understanding of any dataset is paramount in inferring gene regulatory network and if necessary, integrating multiple datasets from distinct sources is challenging [40].

Table 1. Some of Transcriptome Profiling Techniques.

	Hybridization-Based		Sequence-Based	
Features	DNA Microarray	ChIP-Microarray (ChIP-chip)	RNA-seq	ChIP-seq
Transcriptome Profiling	Yes	Yes	Yes	Yes
DNA-Protein Interaction	No	Yes	No	Yes
Sensitivity	Low	Low	High	High
Specificity	Low	Low	High	High
Range of expression	Low	Low	High	High
Signal-to-Noise Ratio	High	High	Low	Low
Cost	Less Expensive	Less Expensive	Costly (Cost coming down)	Costly (Cost coming down)
Discovery-based experiments	No	No	Yes	Yes
Histone Modification Analysis (Epigenetics)	No	Yes	No	Yes
DNA methylation Analysis (Epigenetics)	No	Yes	No	Yes
Functional Analysis of Transcription Factor	Not Feasible	Possible	Not Feasible	Possible
Transcription Factor Binding Sites (TFBS) Mapping	Not Feasible	Possible	Not Feasible	Possible

3. Grn Visualization Tools

We are able to perform a comprehensive analysis of several software tools for investigating and visualizing gene regulatory network, which are obtained from publicly available biological tools on Omic Tool website (http://omictools.com/gene-regulatory-networks-c435-p1.html). Their various properties in relation to their input dataset are reported in several literatures [1, 2, 5, 7, 8, 10, 11, 13, 14, 19-22, 25, 26, 29, 31, 34, 37, 38, 40, 41]. We present the summarized analysis of the tools in Table 2, taking cognizance of the input dataset among other properties.

a) ARACNE (Algorithm for the Reconstruction of Accurate Cellular Network)

ARACNE is a powerful network inference tool designed to scale up to the complexity of regulatory networks in mammalian cells. It uses microarray dataset from human B cells, both a realistic and synthetic datasets. The microarray dataset makes it possible to measure statistical interactions and dependencies using mutual information, which does not

require discretization of the expression level [29]. However, ARACNE is unable to infer edge directionality because it does not use temporal data.

b) ARTIVA (Auto Regressive TIme VArying Model)

ARTIVA uses time course gene expression data to perform a gene-gene analysis, and to infer the topology of the network and how it changes over time using two microarray datasets. The data related to the developmental stages of Drosophila Melanogaster and Benomyl data are the major input data into the software, but including the gene ontology information to perform the knock-out procedure and functional annotations and transcription factor binding information to access the biological relevance of the result [26].

ARTIVA provides powerful and evolving mechanism in inferring gene regulatory network because the nature of incoming input is investigated, works with a continuous datasets and no threshold is needed to define up and down regulated groups of genes. It is thereby concluded that ARTIVA will be able to incorporate data originating from different sources like ChIP-Chip or ChIP-seq experiment.

c) AtmiRNET

AtmiRNET is another software tool for biological network inference to explore mechanisms of transcriptional regulation and microRNA functions in Arabidopsis Thaliana. The gene regulatory networks gives users an intuitive insight into the pivotal roles of Arabidopsis miRNAs through the crosstalk between miRNA transcriptional regulation (upstream) and miRNA mediate (downstream) gene circuit [5].

The data input include miRNA of Arabidopsis, core promoter element and high confidence transcription factors. It was observed that the plant miRNAs are primarily encoded in intergenic regions and that they have their own promoters unlike animals' miRNAs.

d) CMGRN

Constructing Multilevel Gene Regulatory Networks (CMGRN) performs several instigative biological functions and gene regulatory network inference. It uses ChIP-seq count, binding data, gene expression profile, miRNA and targets as input dataset and according to [14], the use of ChIP-seq data allows high-fidelity mapping of different regulators; binding data to identify regulatory modules and reconstruct GRN; gene count to infer the causal relationship between Transcription Factors (TFs) epigenetic modification; gene expression data and miRNA/regulatory signal of TFs are used to construct the GRN by multi-level factors.

In summary, CMGRN generates hierarchical regulatory network structures controlled by the interacting factors at transcriptional, post-transcriptional and epigenetic layers.

e) ChIP-Array

It is a web server biological network visualization tool that integrates ChIP-X (ChIP-Chip, ChIP-seq, etc) and gene expression data from human, mouse, yeast fruit-fly and Arabidopsis to analyze both the ChIP-X and expression data together. It requires binding locations from ChIP-X data, differential expression data from gene expression profile and other parameters to construct GRN [34]. The direct and indirect target genes are detected, which is regulated by a Transcriptional Factor (TF) of interest.

This ultimately aids the characterization of function(s) of the TF.

f) ChIP-Array2

ChIP-Array2 is an enhanced version of ChIP-Array, which accommodates additional type of omics data from rat and worm to investigate a more comprehensive gene regulatory network involving diverse regulatory components. ChIP-Array2 can be used to detect the direct and indirect target genes separately; this is as a result of independence of both ChIP-X and expression data. It can run without either ChIP-X or expression data, direct target genes are detected with only expression data while indirect target genes are detected with only ChIP-X data [37].

Table 2. *Dataset Summary of GRN Inference Tools.*

Tool	Expression Data (Microarry, RNA-seq)	Regulatory Data (ChIP-X)	Combined Expression data and ChIP-X data	Applied Genome
ARACNE	Yes	No	No	Human B cells
ARTIVA	Yes	No	No	Drosophila Melanogaster and Benomyl
AtmiRNET	Yes	No	No	Arabidopsis
biRte	Yes	No	No	E.Coli, Pancreas and Ovarian Cancer
C3NET	Yes	No	No	E.coli
CASCADE	Yes	No	No	Lymphocytes B-cells and Murine Lymphocytes T-cells
CGC-2SPR	Yes	No	No	Yeast Metabolic Cycle(YMC)
ChIP-Array	Yes	Yes	Yes	Human, Mouse, Yeast, Fruit-Fly and Arabidopsis
ChIP-Array2	Yes	Yes	Yes	Human, Mouse, Rat, Yeast, Fruit-Fly, Worm and Arabidopsis
CL	Yes	No	No	Cancer
CLR	Yes	No	No	E.coli
CMGRN	Yes	Yes	Yes	-
CompareSVM	Yes	No	No	-
CoREGNET	Yes	Yes	Yes	-
DDGni	Yes	No	No	Yeast and Caenorhabditis Elegan
ESCAPE	Yes	Yes	Yes	mouse Embryonic Stem Cells(mESCs)
GENIE3	Yes	No	No	-
INFERELATOR	Yes	No	No	-
INFERENCE SNAPSHOT	Yes	No	No	Hematopoietic Stem Cells
iRegulon	Yes	Yes	Yes	Human, Mouse and Drosophila
Jump3	Yes	No	No	-
LOREGIC	Yes	Yes	Yes	Human Leukemia
ModEnt	Yes	No	No	-
netClass	Yes	No	No	-
Netmes	Yes	No	No	-

g) iRegulon

This is another powerful network inference tool designed to identify master regulators and detect target genes in human, mouse and drosophila genes. It is used for motif discovery, which provides access to many cancer-related TF-target subnetworks/regulon [21].

The input datasets are TF binding data (from ENCODE ChIP-seq), co-expressed genes downstream of a TF perturbations, miRNA and genes involved in the same signaling pathway.

Other important gene regulatory network software tools are PANDA, PathWave, PTHGRN, ROBNCA, SIRENE, SMARTS, TFmiR and TIGRESS.

Finally, we observed that large percentage of the tools if not all has additional input parameters like TSSs, TFs, protein-protein interaction, functional annotations, binding information and other gene ontologies information to unravel the interaction and relationship complexity of GRN.

4. Conclusion

Analysis of input dataset especially based on the technology of source data provides insight for future researchers on the kind of data to use for a specific investigation. We have been able to dissect GRN input dataset and discovered that the datasets can be expression data or regulatory data (ChIP-X) aside the technology of the source of the data. Though, Next Generation Sequencing (NGS) allows the elucidation and demarcation of complex transcriptional regulatory networks [18], the major disparity is whether the dataset is ChIP-X or expression dataset not whether the source is from hybridization-based or sequence-based.

Gene expression dataset is from microarray or RNAseq while ChIP-X dataset is either from ChIP-Chip or ChIP-seq experiment. We observed from this analysis that there is no one universal method suitable for inference of GRN for all biological conditions [10], the suitability of inference method depends on the kind of datasets employed and the important features of each dataset can suggest other hypothesis in investigating organism or diseased cells.

Notwithstanding the presence of expression data in all the tools analyzed above, the ChIP-X dataset has been observed to help elucidate the genetic, epigenetic and environmental states of a cell and help in a great extent to determine the phenotype of the cell. In fact, [41] submitted that, no true causal relationships can be represented with any pure expression driven method, and that the problem can be solved by using ChIP-Seq binding data. Also, [13] observed that with ChIP-X dataset, visualizing GRN and modeling of metabolic and signal networks can be combined to model the global operation of cells with unprecedented completeness and accuracy. Besides, with ChIP-X dataset, relevant subnetworks that underlie observed genetic interactions can be reconstructed [21], and cooperations among other regulatory elements can be studied such as splicing factors, long non-coding RNAs, etc. [38].

Finally, the resulting networks of inference tools with more input parameters are more accurate than those produced using individual dataset, and E.coli dataset is believed to be the benchmark of biological dataset [19].

5. Suggestions

Taking this comprehensive analysis into consideration, and looking at importance of gene regulatory networks as the fundamental solution to majority of biological questions, the following suggestion should be taken care of.

a) A concerted effort should be made in designing a GRN software template that will contain all the additional data from ChIP-X experiment, which makes it different from expression data. This will make expression dataset to be useful to carry out further investigations like histone modification analysis, DNA methylation analysis, functional analysis of transcription factors and Transcription Factor Binding Sites (TFBS) mapping.

b) A knowledge base of prior knowledge should be built, and it should be available to every GRN inference developer. It is observed that all the tools require information like gene ontology, functional annotation, protein-protein interaction information, binding information etc. to successfully reconstruct the GRN. It can also help to achieve above suggestion by integrating data from different omic datasets as a single knowledge base.

c) Both the expression and ChIP-X datasets of E.coli should be obtained and made available to be used as the simulated dataset for all GRN inference tools because majority of the authors in the papers reviewed used E.coli data successively.

References

[1] Altay, G. and Emmert-Streib, F. (2010) Inferring the Conservative Causal Core of Gene Regulatory Networks. BMC Systems Biology, 4:132.

[2] Altay, G., Kurt, Z., Dehmer, M., and Emmert-Streib, F. (2014). Netmes: Assessing Gene Network Inference Algorithms by Network-Based Measures. Evolutionary Bioinformatics: 1–9 doi: 10.4137/EBO.S13481.

[3] Ambros, V., Bartel, B., Bartel, D.P. et al. (2003) A uniform system for micro RNA annotation. RNA 9, 277–279.

[4] Bower, J. (2001) Computational Modelling of Genetic and Biochemical Networks. MIT Press, Cambridge.

[5] Chien, C., Chiang-Hsieh, Y., Chen, Y., Chow, C., Wu, N., Hou, P., and Chang, W.(2015). Atmi RNET: A Web-Based Resource for Reconstructing Regulatory Networks of Arabidopsis Micro RNAs. Database, 2015, 1–11 doi: 10.1093/database/bav042.

[6] Ching, W., Fung, E., Ng, M. and Akustu, T. (2005) On Construction of Stochastic Genetic Networks Based on Gene Expression Sequences. International Journal of Neural Systems, 15(4), 297-310.

[7] Cun, Y. and Frö̈hlich, H. (2014) .net Class: An R-Package for Network Based, Integrative Biomarker Signature Discovery. BIOINFORMATICS, Vol. 30, no. 9, pages 1325–1326 doi:10.1093/bioinformatics/btu025.

[8] Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, et al. (2007) Large-Scale Mapping and Validation of Escherichia Coli Transcriptional Regulation From a Compendium of Expression Profiles. PLoS Biol 5(1): e8. doi:10.1371/journal.pbio.0050008.

[9] Fatumo, S., Plaimas, K., Mallm, J.P., Schramm, G., Adebiyi, E., Oswald, M., Eils, R., König, R., 2009. Estimating novel potential drug targets of Plasmodium falciparum by analysing the metabolic network of knock-out strains in silico. Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 9, 351–358.

[10] Gillani, Z., Akash, M., Rahaman, MD. and Chen, M. (2014). Compare SVM: Supervised, Support Vector Machine (SVM) Inference of Gene Regularity Networks. BMC Bioinformatics, 15:395.

[11] Glass, K., Quackenbush, J., Spentzos, D., Haibe-Kains, B., and Yuan, G. (2015). A Network Model for Angiogenesis in Ovarian Cancer. BMC Bioinformatics 16:115 DOI 10.1186/s12859-015-0551-y.

[12] Goutsias, H. and Lee, N.H.(2007) Computational and Experimental Approaches for Modeling Gene Regulatory Networks. Curr. Pharm. Design, 13(14):1415–1436.

[13] Greenfield, A., Hafemeister, C. and Bonneau, R. (2013). Robust Data-Driven Incorporation of Prior Knowledge into the Inference of Dynamic Regulatory Networks. BIOINFORMATICS, Vol. 29 no. 8 2013, pages 1060–1067, doi:10.1093/bioinformatics/ btt099.

[14] Guan, D., Shao, J., Deng, Y., Wang, P. Zhao, Z., Liang, Y., Wang, J. and Yan, B.(2014).CMGRN: A Web Server For Constructing Multilevel Gene Regulatory Networks using Chip-Seq and Gene Expression Data. BIOINFORMATICS, Vol. 30 no. 8, pages 1190–1192, doi:10.1093/bioinformatics/btt761.

[15] Hartemink, A., Gifford, D., Jaakkola, T., et al. (2002) Bayesian Methods for Elucidating Genetic Regulatory Networks. IEEE Intelligent Systems, 17(2), 37-43.

[16] Hidde, D.J. (2002) Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. Journal of Computational Biology, 9(9), 67-103.

[17] Ho, J., Bishop, E., Karchenko, P., Nègre, N., White, K. and Park, P. (2011).ChIP-chip versus ChIP-seq: Lessons for Experimental Design and Data Analysis. BMC Genomics 2011, 12:134.

[18] Hurd, P. J. and Nelson, C. J. (2009). Advantages of Next-Generation Sequencing Versus the Microarray in Epigenetic Research. Briefings in Functional Genomics and Proteomics. Vol 8. No 3, p174 -183, doi:10.1093/bfgp/elp013.

[19] Huynh-Thu, V., Irrthum, A., Wehenkel, L. and Geurts, P.(2010). Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. PLoS ONE, Volume 5, Issue 9, e12776.

[20] Huynh-Thu, V. and Sanguinetti, G. (2015). Combining Tree-Based and Dynamical Systems for the Inference of Gene Regulatory Networks. Bioinformatics, 1614–1622 doi:10.1093/bioinformatics/btu863.

[21] Janky, R., Verfaillie, A., Imrichova, H., Sande, B. et al. (2014).

iRegulon: From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections. PLOS Computational Biology, Volume 10, Issue 7, e1003731.

[22] Jung, N., Bertrand, F., Bahram, S., Vallat, L. and Maumy-Bertrand, M. (2014). Cascade: A R Package to Study, Predict and Simulate the Diffusion of a Signal through a Temporal Gene Network. Bioinformatics Applications Note, Vol. 30 no. 4, pages 571–573, doi:10.1093/bioinformatics/btt705.

[23] Karlebach, G. and Shamir, R. (2012). Constructing logical models of gene regulatory networks by integrating transcription factor-DNA interactions with expression data: an entropy-based approach. Journal of Computational Biology.: 30-41.doi: 10.1089/ cmb.2011.0100.

[24] Kharumnuid G. and Roy S. (2015) Tools for in-silico Reconstruction and Visualization of Gene Regulatory Networks (GRN). Second International Conference on Advances in Computing and Communication Engineering. IEEE Computer Society. 978-1-4799-1734-1/15, p421-426.

[25] Kling, T., Johansson, P., Sanchez, J., Marinescu, V., Jornsten, R. and Nelander, S. (2015). Efficient Exploration of Pan-Cancer Networks by Generalized Covariance Selection and Interactive Web Content. Nucleic Acids Research, doi: 10.1093/nar/gkv413.

[26] Lèbre, S., Becq, J., Devaux, F., Stumpf, M., Lelandais, G. (2010) Statistical Inference of the Time-Varying Structure of Gene-Regulation Networks. BMC Systems Biology, 4: 130.

[27] Linde, J., Schulze, S., Henkel, S., Guthke, R. (2015) Data- And Knowledge-Based Modeling of Gene Regulatory Networks: An Update. EXCLI Journal; 14: 346-378 – ISSN 1611-2156.

[28] Lodish, H., Berk, A., Zipursky, S., Matsudaira, P., Kaiser C., Krieger M., Scott M., Darnell J.(2000). Molecular Cell Biology. 4th edition. Section 4.4. New York: W. H. Freeman.

[29] Margolin, A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Favera, R., and Califano, A. (2006) ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. BMC Bioinformatics, 7(Suppl 1): S7 doi: 10.1186/1471-2105-7-S1-S7.

[30] Nelson P., Kiriakidou M., Sharma A., Maniataki E. and Mourelatos Z. (2013). The micro RNA World: Small is Mighty. TRENDS in Biochemical Sciences Vol.28 No.10, p534-540.

[31] Nicolle, R., Radvanyi, F., and Elati, M. (2015).CoREGNET : Reconstruction and Integrated Analysis of Co-Regulatory Networks. Bioinformatics, 1–3, doi: 10.1093/bioinformatics/btv305.

[32] Ocone, A., Haghverdi, L., Mueller, N.S. and Theis, F.J. (2015). Reconstructing Gene Regulatory Dynamics from High-Dimensional Single-Cell Snapshot Data. Bioinformatics, ISMB/ECCB 2015, i89–i96, doi: 10.1093/bioinformatics/btv25.

[33] Petralia, F., Wang, P., Yang, J. and Tu, J. (2015). Integrative Random Forest for Gene Regulatory Network Inference. Bioinformatics, ISMB/ECCB 2015, i197–i205, doi: 10.1093/bioinformatics/btv268.

[34] Qin, J., Li, M., Wang, P., Zhang, M. and Wang, J. (2011). ChIP-Array: Combinatory Analysis of Chip-Seq/Chip and Microarray Gene Expression Data to Discover Direct/Indirect Targets of a Transcription Factor. Nucleic Acids Research, Vol. 39, Web Server issue W430 – W436.

[35] Rui-Rui Ji, Ding Liu, Wen Zhang (2010) The Application of Hidden Markov Model in Building Genetic Regulatory Network. Biomedical Science and Engineering, J. Biomedical Science and Engineering, 3, 633-637.

[36] Vijesh, N., Chakrabarti, S., Sreekumar, J. (2013) Modelling of Gene Regulatory Networks: A Review. Biomedical Science and Engineering, 6, 223-231.

[37] Wang, P., Qin, J., Qin, Y., Zhu, Y., Wang, L.Y., Li, M., Zhang, M. and Wang, J. (2015).ChIP-Array 2: Integrating Multiple Omics Data to Construct Gene Regulatory Networks. Nucleic Acids Research, Vol. 43, Web Server issue, W264–W269.

[38] Wang, D., Yan, K., Sisu, C., Cheng, C., Rozowsky, J., Meyerson, W. and Gerstein, M.B. (2015). Loregic: A Method to Characterize the Cooperative Logic of Regulatory Factors. PLOS Computational Biology DOI: 10.1371/journal. pcbi. 1004132.

[39] Wang Z., Gerstein M., and Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet.10(1): 57–63. doi: 10.1038/nrg2484.

[40] Xu, H., Ang, Y., Sevilla, A., Lemischka, I.R. and Ma'ayan, A. (2014).Construction and Validation of a Regulatory Network for Pluripotency and Self-Renewal of Mouse Embryonic Stem Cells. PLOS Computational Biology, Volume 10, Issue 8, e1003777.

[41] Yalamanchili, H.K., Yan B., Li M.J., Qin, J., Zhao, Z., Chin, F. and Wang, J. (2014). DDGni: Dynamic Delay Gene-Network Inference from High-Temporal Data using Gapped Local Alignment. BIOINFORMATICS, Vol. 30 no. 3, pages 377–383 doi:10.1093/bioinformatics/btt692.

[42] Yao, S., Yoo, S. and Yu, D. (2015). Prior Knowledge Driven Granger Causality Analysis on Gene Regulatory Network Discovery. BMC Bioinformatics 16:273, DOI 10.1186/s12859-015-0710-1.

[43] Zhang, S.-Q., Ching, W.-K. and Yue, J. (2008) Construction and Control of Genetic Regulatory Networks: A Multivariate Markov Chain Approach. Journal of Biomedical Science and Engineering, 1, 15-21.

[44] Zhang, Z.-F. (2004) Constructing and Predicting Gene Regulatory Network Using Micro Array Data. National Central University, Taiwan.

[45] Zacher, B., Abnaof, K., Gade, S., Younesi, E., Tresch, A., and Fröhlich, H. (2012) Joint Bayesian Inference of Condition-Specific miRNA and Transcription Factor Activities from Combined Gene and micro RNA Expression Data. Vol. 28 no. 13, pages 1714–1720 doi: 10.1093/bioinformatics/bts257.