Cell Syst. Author manuscript; available in PMC 2017 November 24.

Published in final edited form as:

Cell Syst. 2017 May 24; 4(5): 478–479. doi:10.1016/j.cels.2017.05.004.

Evolutionary context improves regulatory network predictions

Jonathan Gordon[1], **Brigida Gallone**[1,2], **Steven Maere**[2], and **Kevin J. Verstrepen**[1]

[1]Laboratory for Systems Biology, VIB Center for Microbiology, KU Leuven, Bio-Incubator, Gaston Geenslaan 1, 3001 Leuven, Belgium

[2]VIB Center for Plant Systems Biology, and Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Gent, Belgium

Struck by the seemingly low number of coding changes between chimpanzee and human DNA, Allan Wilson and Marie-Claire King pondered: “Is it possible, therefore, that species diversity results from molecular changes other than sequence differences in proteins?” (King and Wilson, 1975). A revolutionary idea at the time, it has since become clear that regulatory changes indeed play pivotal roles in evolution (Halligan et al., 2013). However, our understanding of how regulatory networks function and evolve is still limited and confounded by their sheer complexity, with a large number of regulatory proteins interacting with each other and/or with specific binding sequences that can be physically distant from their respective target genes. Moreover, unlike coding regions, binding sites are small, largely unstructured stretches of DNA with a variable base composition, making it difficult to even identify them based on their sequence alone, let alone predict how they are affected by mutations.

Accurately mapping transcriptional regulatory networks requires large transcriptome and interactome datasets from different conditions. As a result, high-quality networks with reasonable coverage only exist for a few model organisms. Networks for non-model organisms are often inferred from limited datasets, using the more comprehensive maps of phylogenetically related model species as a template (Thompson et al., 2015). On Pg. X of this issue, Koch and coworkers present a new algorithm for network inference, MRTLE. MRTLE relies on a probabilistic graphical modeling approach that increases the accuracy of regulatory network inference from gene expression data, by incorporating the simple yet powerful idea that regulatory network similarity between two species correlates with their genetic similarity (Figure 1). In addition to gene expression data for several species, MRTLE requires a phylogenetic tree as input, as well as orthogroups for both the regulators and the targets, and rate parameters for the gain and loss of regulatory interactions. Additionally, MRTLE can incorporate known binding site motifs as prior information. By comparing the networks inferred by MRTLE and other techniques from gene expression data for six yeast species to gold standard networks originating from ChIP-chip experiments (both on a larger scale in Saccharomyces cerevisiae and for selected well-studied transcription factors across multiple species), the authors convincingly show that their phylogenetically informed approach provides more accurate network inference than approaches that do not include phylogenetic information, and allows more accurate prediction of regulatory networks for non-model systems starting from only a limited dataset. For instance, for several (but not all) transcription factors tested, the TF targets predicted by MRTLE in Saccharomyce cerevisiae,

Kluyveromyces lactis and Candida albicans were found to exhibit higher fold-enrichment of known TF targets than the targets predicted by the similar but non-phylogenetically informed INDEP algorithm (Roy et al., 2011).

One of the major advantages of algorithms that apply an evolutionary framework to biological data is the inference of ancestral states. Given the inferred ancestral and measured contemporary states of a system, it becomes possible to trace the evolution of the system over time. Thus, apart from predicting network structures across related organisms that more accurately reflect their shared evolutionary history, algorithms like MRTLE can also offer a glimpse of how networks evolve over time by predicting how the ancestral regulatory networks might have looked. Some recent studies have been able to unravel how the regulation of a gene or gene cluster evolved and diverged between related species, but this has always involved a large amount of computational and wet lab work (see for example Pougach et al., 2014, doi: 10.1038/nature14613). A recurring theme in these studies is that regulatory changes often result from subtle and gradual changes in both transcription factors and their binding sites. Moreover, duplication of transcription factor genes is often found to be an important driver of regulatory evolution (Voordeckers et al., 2015).

Koch and co-workers demonstrate the utility of MRTLE to investigate the evolution of regulatory networks by tracing the evolution of the osmotic stress response network across ascomycete species. The inferred networks show conservation of repressors and activators acting as hubs among all the species or within groups of closely related species. The function of regulators as activators or repressors of stress-responsive gene expression is also generally conserved across species and stresses. For example, MSN2/4 and ribosomal biogenesis regulators were found to function as activators and repressors of stress-responsive genes, respectively, across all species and stresses. Exceptions to this trend are found relatively rarely, and mostly on a species-specific basis, suggesting that regulatory network rewiring is a gradual process. MRTLE was also used to examine evolutionary dynamics of edge gain and loss in ascomycete regulatory networks. The authors found that gene duplication generally increases the rate of both edge gain and loss, congruent with the fact that duplication of genes is thought to stimulate regulatory divergence (Pougach et al., 2014). More generally, the identification of regulators with low and high interaction turnover rates may be used to assess which processes are most conserved or diversifying in the species concerned, which may lead to better insight into the nature of the selection pressures acting on particular processes in particular species.

Despite the progress that MRTLE brings, regulatory network inference is still in its infancy. There is still ample room to improve the accuracy and coverage of transcriptional network predictions, both by generating more data for network inference algorithms, and by continuously adapting the inference algorithms to make use of newly emerging insights on the nature of transcriptional regulation processes and their evolution, often gleaned from detailed analyses on specific transcriptional systems. Ultimately, accurate and complete network reconstructions will allow the tracing of the evolution of expression dynamics for any given gene or pathway across a phylogeny. By incorporating an evolutionary dimension in the inference of transcriptional networks, MRTLE takes an important step towards this goal.

Europe PMC Funders Author Manuscripts

References

Halligan DL, Kousathanas A, Ness RW, Harr B, Eöry L, Keane TM, Adams DJ, Keightley PD. Contributions of Protein-Coding and Regulatory Change to Adaptive Molecular Evolution in Murid Rodents. PLOS Genet. 2013; 9:e1003995. [PubMed: 24339797]

King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. Science. 1975; 188:107–116. [PubMed: 1090005]

Pougach K, Voet A, Kondrashov FA, Voordeckers K, Christiaens JF, Baying B, Benes V, Sakai R, Aerts J, Zhu B, et al. Duplication of a promiscuous transcription factor drives the emergence of a new regulatory network. Nat Commun. 2014; 5:4868. [PubMed: 25204769]

Roy S, Werner-Washburne M, Lane T. A multiple network learning approach to capture system-wide condition-specific responses. Bioinformatics. 2011; 27:1832–1838. [PubMed: 21551143]

Thompson D, Regev A, Roy S. Comparative analysis of gene regulatory networks: from network reconstruction to evolution. Annu Rev Cell Dev Biol. 2015; 31:399–428. [PubMed: 26355593]

Voordeckers K, Pougach K, Verstrepen KJ. How do regulatory networks evolve and expand throughout evolution? Curr Opin Biotechnol. 2015; 34:180–188. [PubMed: 25723843]

Europe PMC Funders Author Manuscripts

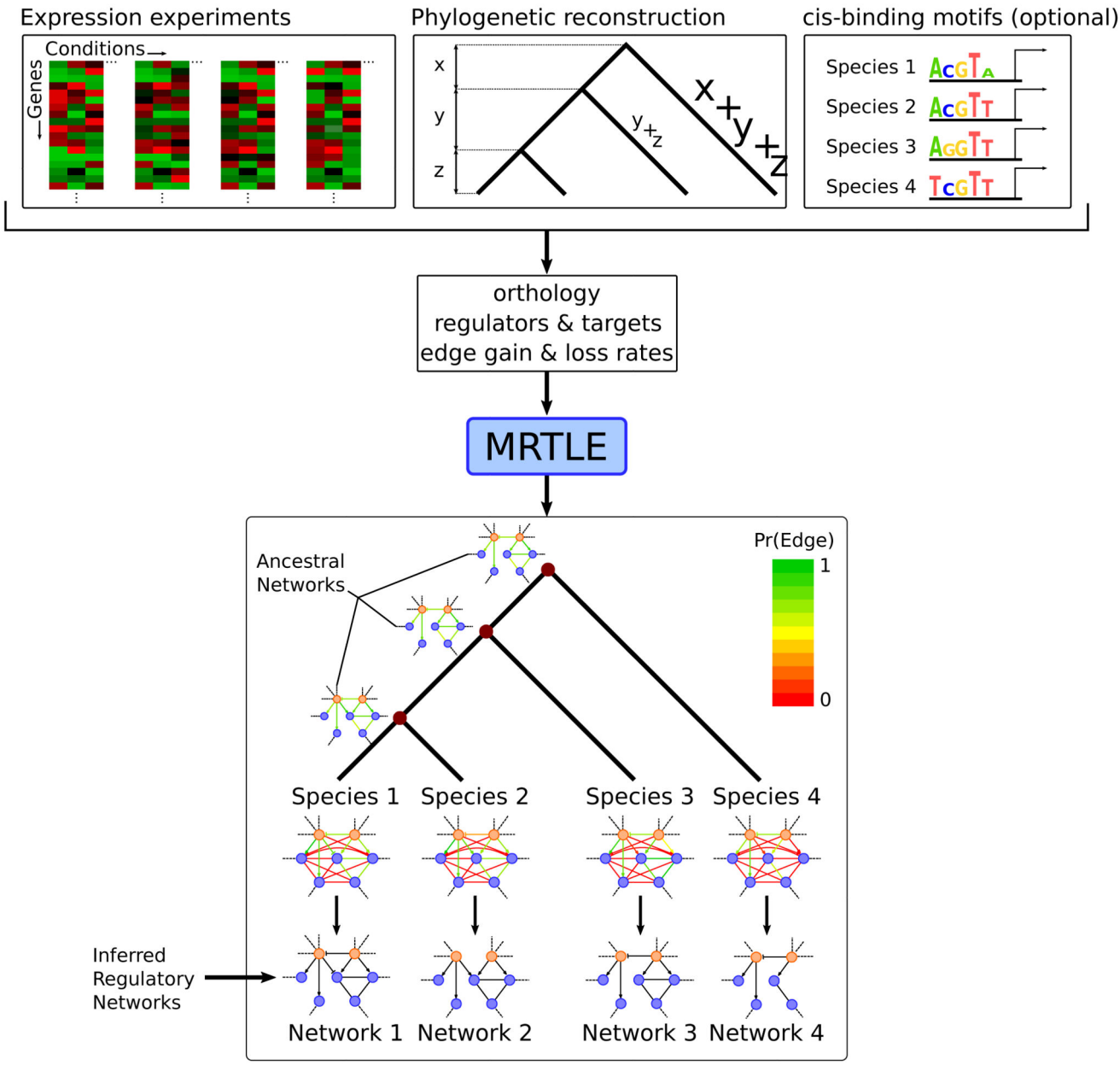

Figure 1.

Schematic representation of the MRTLE framework. MRTLE combines data from largescale expression experiments, orthology, assignment of regulators and targets, phylogenetic distances, and optional cis-regulatory motifs across several species to estimate the probability of the edges in a regulatory network graph at the extant and ancestral nodes of the phylogeny.