In Silico Gene Regulatory Network of the Maurer's Cleft Pathway in *Plasmodium falciparum*

Itunuoluwa Isewon[1,2], Jelili Oyelade[1,2], Benedikt Brors[3] and Ezekiel Adebiyi[1–3]

[1]Department of Computer and Information Sciences, Covenant University, Ota, Ogun State, Nigeria. [2]Covenant University Bioinformatics Research (CUBRe), Covenant University, Ota, Ogun State, Nigeria. [3]Department of Applied Bioinformatics, German Cancer Research Centre (DKFZ), Heidelberg, Germany.

ABSTRACT: The Maurer's clefts (MCs) are very important for the survival of *Plasmodium falciparum* within an infected cell as they are induced by the parasite itself in the erythrocyte for protein trafficking. The MCs form an interesting part of the parasite's biology as they shed more light on how the parasite remodels the erythrocyte leading to host pathogenesis and death. Here, we predicted and analyzed the genetic regulatory network of genes identified to belong to the MCs using regularized graphical Gaussian model. Our network shows four major activators, their corresponding target genes, and predicted binding sites. One of these master activators is the serine repeat antigen 5 (SERA5), predominantly expressed among the SERA multigene family of *P. falciparum*, which is one of the blood-stage malaria vaccine candidates. Our results provide more details about functional interactions and the regulation of the genes in the MCs' pathway of *P. falciparum*.

KEYWORDS: Maurer's cleft, gene regulatory network, *Plasmodium falciparum*, graphical Gaussian model

CITATION: Isewon et al. *In Silico* Gene Regulatory Network of the Maurer's Cleft Pathway in Plasmodium falciparum. *Evolutionary Bioinformatics* 2015:11 231– 238 doi: 10.4137/EBO.S25585.

TYPE: Original Research

RECEIVED: March 05, 2015. **RESUBMITTED:** July 28, 2015. **ACCEPTED FOR PUBLICATION:** August 03, 2015.

ACADEMIC EDITOR: Jike Cui, Associate Editor

PEER REVIEW: Three peer reviewers contributed to the peer review report. Reviewers' reports totaled 2,269 words, excluding any confidential comments to the academic editor.

FUNDING: This work was supported by study grants from the German Academic Exchange Service (DAAD), Covenant University, and German Cancer Research Centre (DKFZ). The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: itunu.isewon@covenantuniversity.edu.ng

Introduction

The existence of multiple-stained structures known to be Maurer's clefts (MCs) in the cytosol of erythrocytes derived from patients with *Plasmodium falciparum* malaria using light microscopy was first detected by George Maurer.[1–3] For over a century, the functions of the MCs remained obscure.[4] In the last decade, scientists have been interested in the mediatory role the MCs play in protein transportation from the cytoplasm of *P. falciparum* to the surface of the erythrocyte of the host, which they consider an unconventional and intriguing biological procedure.[5,6]

Upon evasion, *P. falciparum* totally recasts the infected human red blood cells to access nutrients and to elude the immune system.[7] The erythrocytes have no secretory systems that the parasite can subvert. As such, it evolves trafficking pathways for the export of macromolecules.[8] More than 10% of all the parasite's proteins (>400 proteins) are smuggled to the cytosol of its host.[7,9] The transported proteins induce extreme structural adjustment to the red blood cell, such as increasing the toughness of the cell of its host and the development of lumps. This in turn makes the affected cells adhere to the endothelium, hence leading to the development of malaria.[8,10,11] Reports also show that MCs are created at various portions in the life cycle of the parasite, and this affirms their relevance to parasite development and survival.[4]

Computationally, only about one-third of the transcription factors (TFs) expected of the genome of the size of *P. falciparum* are known, while the knowledge of the corresponding binding sites is lacking. A few works have experimentally validated motifs in the parasite, but their binding proteins are unknown. Others have identified some transcription-associated proteins, but very little is known about their binding sites.[12–15] This study is an attempt to identify novel transcriptional regulators (master activators/inhibitors) of the MC pathway in *P. falciparum* by predicting its gene regulatory network (GRN) from gene expression data.

Gene regulatory networks. Gene networks depict transcriptional regulation, ie, how genes respond to signals.[16] GRNs simply show how and when genes in a cell interact with each other. GRNs expressly depict the causality of developmental processes.[17] Several methods exist for the modeling and simulation of GRNs from high throughput experimental data, such as expression profiles, genomic sequences, and TF-binding site assays. These include, but are not limited to, Boolean models,[18,19] Petri net models,[20,21] neural network models,[22,23] ordinary differential equation models,[24,25] Bayesian network (BN) models,[24] relevance/association networks,[26] and graphical Gaussian models (GGMs).[27–29]

A commonly used computational method in modeling GRNs is to compute the Pearson correlation coefficients (r) between all possible combinations of the genes of interest. Any of such combinations having Pearson correlation coefficients above a predefined threshold is considered to be significant, ie, exhibit useful relationships, influence, or dependence.[27] Such networks are referred to as relevance/association networks. It is important to note that they may lead to vague results, particularly if it is a dense system of connections.[30]

BNs are classified as probabilistic graphical models.[31] BNs are also a type of directed acyclic graphs that are other types of graphical model structures. BNs and its extensions have been used extensively to predict regulatory networks.[32–38] However, there are drawbacks with using BNs for modeling GRNs. Some drawbacks include the difficulty in dealing with time series data, ie, higher number of genes compared to sample size, and inability to capture feedback regulation/loops that are normal in real-life gene networks and are nondeterministic polynomial-time hard learning problems.[39–41] The implementation of dynamic BNs also has some drawbacks, including excessive computing time, result precision largely dependent on quantity of genes, and requirement of more prior information of the transcription regulation.[5,41–43]

GGM is a preferred alternative. It uses partial correlation (pcor) to ensure a rich appraisal of a direct relationship between any combinations of genes.[44,45] Unlike Pearson correlation, pcor of two genes calculates the amount of correlation left after eliminating the influence of other genes. Previous works have established that GGMs are very effective in inferring conditional dependency and modeling interactions among genes.[27,45–49] Earlier, predicting network interactions using GGM had been limited to data sets with fewer genes due to the fact that gene expression experiments had few samples (n), which are usually very less compared to the amount of genes (P).[45,46] This scenario is referred to as small sampling problem, ie, $P \gg n$ cannot be accommodated by standard GGM.[49,50] One possibility of resolving this problem is by using the regularized GGM approach proposed by Schäfer and Strimmer.[48] This approach is referred to as the shrinkage approach that is suited to cases where the number of genes is slightly larger than the number of samples.[51,52] In order to introduce directionality into the network, as GGMs are essentially undirected networks displaying only direct linear relationships, Opgen-Rhein and Strimmer[53] introduced an extension of the shrinkage approach to graphical Gaussian modeling, which they referred to as causation networks. We have used this regularized directed GGM, as implemented in the GeneNet R package, to build a gene network for MC pathway of *P. falciparum* based on gene expression profiles.

Materials and Methods

Microarray data. Gene expression data with accession number GSE24416 were downloaded from the Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo/, accessed 11/02/2015). The experiment was done by Foth et al.[54] The 24 time point samples were harvested from a tightly synchronous 6.5-L biofermenter culture of *P. falciparum* (Dd2) at two-hour intervals during one entire (48 hours) intraerythrocytic developmental cycle and compared against a 3D7 RNA reference pool.

MC pathway proteins of *P. falciparum*. Genes that belong to the MCs' pathway were downloaded from PlasmoDB version 13.0 (http://plasmodb.org, accessed 11/02/2015). Nine pathways concerned with the MCs, based on the classification from Hagai Ginsburg's Malaria Parasite Metabolic Pathways database (at http://mpmp.huji.ac.il/), were used for this study. They include "Biogenesis of Maurer's clefts" – 3 genes, "Characteristics of *Plasmodium falciparum* export proteins that remodel infected erythrocytes" – 41 genes, "Established and putative Maurer's clefts proteins" – 134 genes, "Exported parasite proteins associated with Maurer's clefts" – 11 genes, "Genes coding for protein traffic related proteins" – 61 genes, "Maurer's clefts" – 7 genes, "Properties of proteins exported to erythrocyte" – 39 genes, "Subcellular location of exported proteins" – 43 genes, and "Subcellular localization of proteins involved in invasion" – 81 genes. Expression values were found for 319 of them from the gene expression data set used. The list of genes used for this study with their gene IDs is provided in Supplementary File S01.

Network construction. All analyses were done in R version 3.1.1 (www.R-project.org), Bioconductor version 2.14,[55] and RStudio version 0.98.978.[56] The following R packages were used for the network modeling: GEOquery version 2.30.1 to load the gene expression data from GEO database directly into R[57]; GeneNet version 1.2.12 to model the GRN[58]; Graphviz version 2.38,[59] dnet version 1.0.7,[60] venerable version 3.0,[61] and igraph version 0.7.1[62] for network visualization; and BCRANK version 1.26.0.[63] The detailed computational pipeline used for the analyses is presented in Supplementary File S02. The R source code for the analyses is provided in Supplementary File S03.

Results

The dynamic shrinkage approach as implemented in the GeneNet[58] R package was used to estimate the pcor matrix of all possible edges. The resulting pcor network had 50,721 edges connecting 319 nodes with pcor values ranging from −0.20 to +0.213. The density function, distribution function, and local false discovery rate (FDR) values are shown in Figure 1.

Our approach favored the edges with lower pcor values. This is because pcors having the lowest absolute values mean they have the largest amount of effects after other genes have been removed. We extracted edges whose |pcor| is ≥0.04. This resulted in 1160 edges out of 50,721 edges. The Pearson correlation coefficient (r) was estimated for these edges and used to further prune the network by eliminating edges with $-0.60 \leq r \leq 0.60$ as any r close to 0 indicates independence.

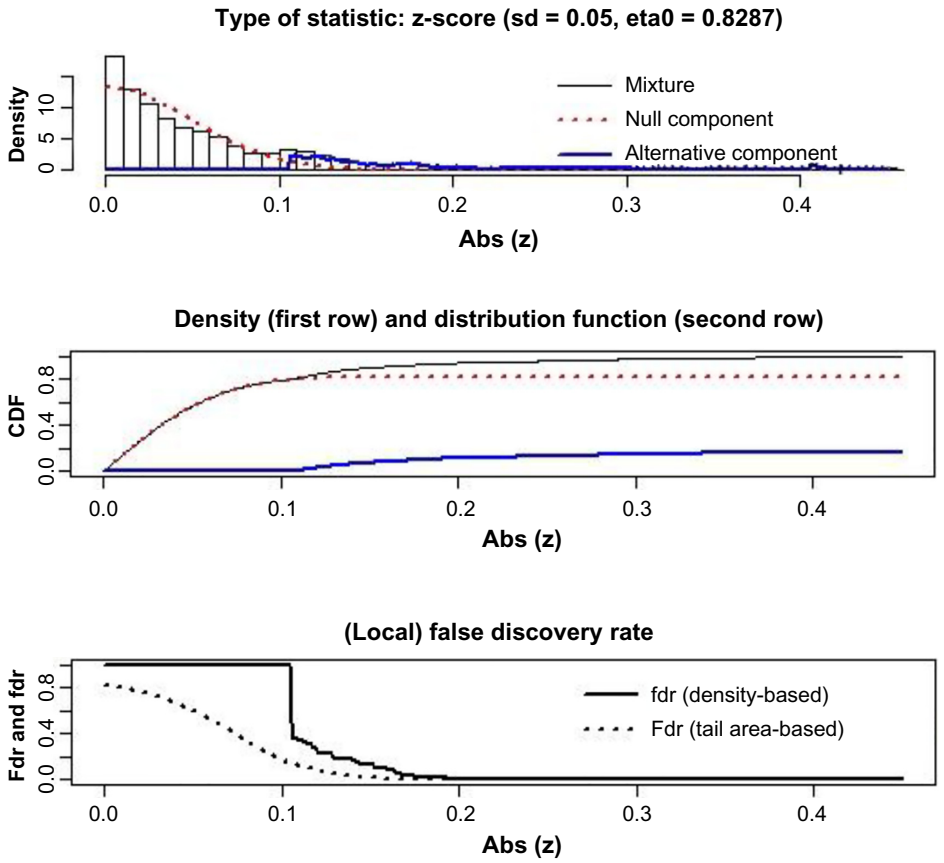

Figure 1. Plot of the density function, distribution function, and local FDR values. The null distribution is depicted by the dashed line; it follows a normal distribution with 0 mean and a standard deviation of 0.05. The solid line signifies the alternative distribution. The empirical distribution (indicated by the histogram) is composed of the null distribution and the alternative distribution.

This left us with 178 edges. We then selected the top 100 edges (sorted according to the absolute strength of the correlation from the strongest to the weakest), representing ~0.2% of all possible edges as significant edges. The resulting GRN of the MCs' pathway is shown in Figure 2. The complete significant edge list with their pcor coefficients, Pearson correlation coefficients, *P*-values, *q*-values, and FDR values is presented in Supplementary File S04.

From this network, we estimated the out-degree of the participating nodes to identify hubs. The hubs are major players in the network (master activation/master inhibitors). Our hypothesis here is that these genes are possibly playing a key role in the regulation of the pathway, ie, the protein trafficking activities of the parasite, and should be further investigated. From our analysis, four genes had the highest number of direct connections in the network: PFL0780w, PFB0340c, MAL7P1.92, and PF08_0002. We extracted the subnetworks of the genes. Our hypothesis is that these genes are the possible target genes of the respective hub genes. The subnetworks are shown in Figure 3A–D.

Predicting binding sites. We extracted 1000 bp upstream of transcription start site for the target genes identified for each master activator from PlasmoDB in FASTA format. The sequences in each group were ranked (arranged) according to the strength of correlation with their master activator. We then used the methods implemented in the R package BCRANK to find the overrepresented motifs in the upstream sequences. BCRANK is useful for predicting binding site consensus for ranked DNA sequences. See the package manual for more details on the BCRANK algorithm.[52] A number of predictions resulted from the running of this algorithm, and the first motif was selected as the candidate motif as it is the most statistically significant motif. Details on the resulting motifs for each set are provided in Supplementary File S05.

Discussion

Our results show four major players that we hypothesized to play important roles in the regulation of activities in the MCs. The first is PFL0780w; from our network, we predicted the following target genes: PFD0955w, PFE1605w, MAL7P1.87, PFI1745c, and PF11_0351. The first two have the motif ACATGTAA overrepresented in their upstream sequence that we predicted to be the binding site of the TF. PFL0780w was reported by the ontology-based pattern identification database to be involved in translational regulator activity and tRNA metabolic process, which is in line with our predictions. From life cycle expression data (combined

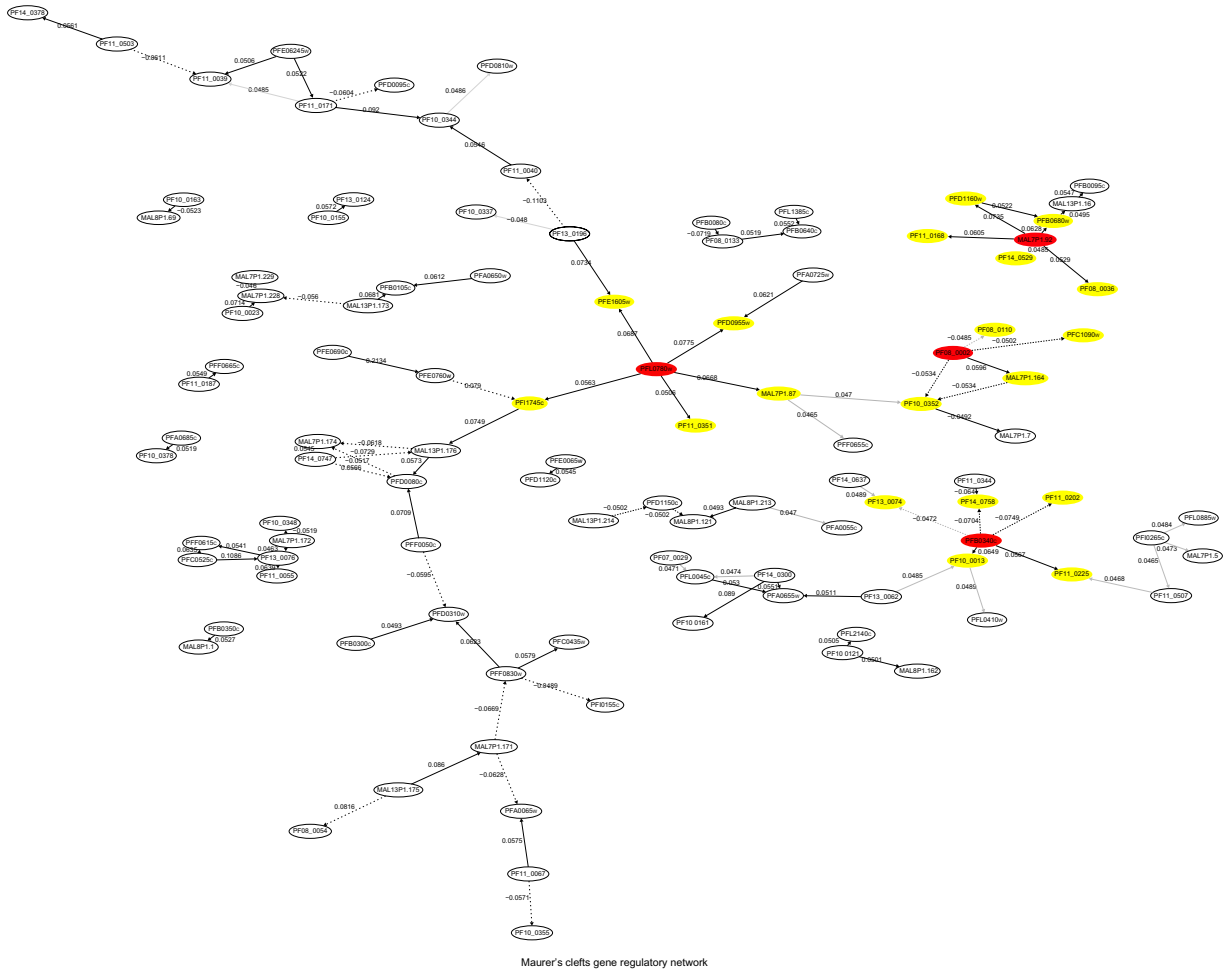

Figure 2. The resulting GRN of the MCs' pathway in *P. falciparum*. The straight lines represent activation relationship (positive correlation), while the dotted lines represent inhibitory relationship (negative correlation). The intensity of the lines represents the amount of regulation. The red nodes are the hubs – master activators/repressors – while the yellow nodes are their predicted target genes.

from several microarray experiments available in PlasmoDB version 13.0), its expression pattern is increasing from the early ring to early trophozoite stage (ie, from 0 to ~25 hpi) and then it reduces; this explains the fact that it is truly an important protein in the survival strategy mechanisms of the parasite. Apart from the MCs, PFL0780w participates in other pathways, such as glycolysis, fatty acid synthesis, and mitochondria electron flow. The second is PFB0340c with the following predicted target genes: PF11_0202, PF14_0758, PF10_0013, PF11_0225, and PF13_0074. The first two genes have the motif TCTTATTTT as the predicted binding site. PFB0340c also known as SERA5 is involved in the regulation of immune response, proteolysis, and immunoglobulin production. SERA5 is predominantly expressed among the SERA multigene family of *P. falciparum* (which is one of the blood stage malaria vaccine candidates), and the acquired antibody titers correlate with the serum inhibition of the parasite growth.[64] Life cycle expression data show that PFB0340c is expressed late in the erythrocytic cycle, ie, from late trophozoite to early schizont stage (~30–35 hpi). The third is MAL7P1.92, and we predicted PFD1160w, PFB0680w, PF11_0168, PF08_0036, and PF14_0529 as its target genes. The last two genes have the motif ATGCCTTTTT overrepresented in their upstream sequence. MAL7P1.92 also known as cysteine repeat modular protein 2 (CRMP2) is localized in parasite plasma membrane during schizont stage and in MC during early trophozoite stage. The fourth is PF08_0002; MAL7P1.164, PF10_0352, PFC1090w, and PF08_0110 are its predicted target genes. The last two genes have the motif TAATAATTTA as their predicted binding site. PF08_0002 is surface-associated interspersed protein 8.2 (SURFIN 8.2). It is upregulated in the early trophozoite stage. Apart from the MC pathway, PF08_0002 is also involved in the interactions between modified host cell membrane and endothelial cell pathway. Table 1 presents more details about each of the four predicted TFs. For each predicted TF, its predicted target genes are listed with their gene names, product description, annotated gene ontology (GO) function term, and annotated GO process term. The enrichment analysis map[65,66] for these predicted TFs and their target genes is presented in Figure 4. The following GO categories were found to be significantly overrepresented: protein transport,

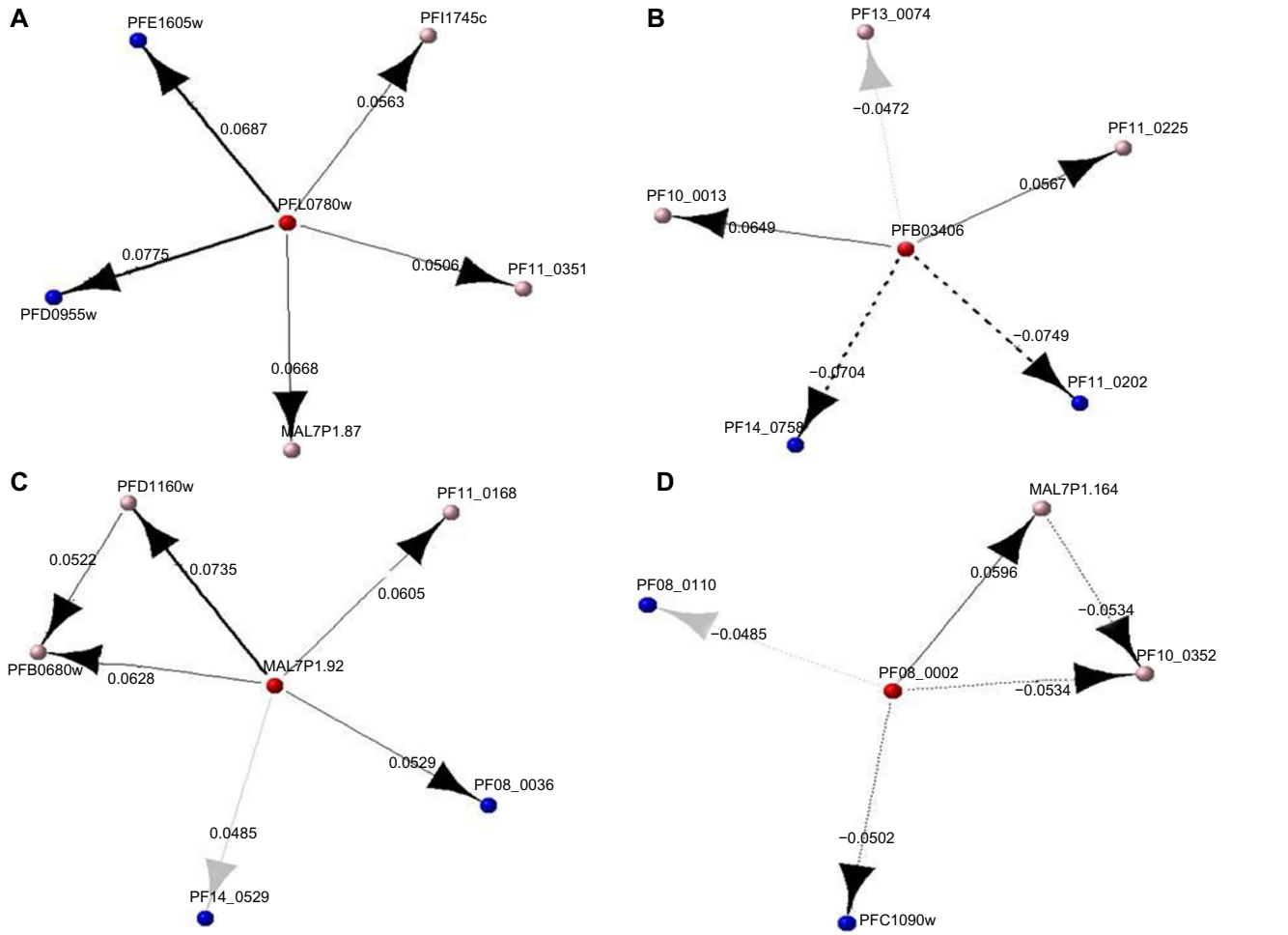

Figure 3. Subnet for the master activators/repressors with their target genes. The red nodes are the predicted TFs, while the blue nodes are the predicted target genes that share an overrepresented motif in their promoter region (predicted binding site) and the pink nodes are predicted target genes that do not share the motif but have a direct interaction from the predicted regulatory network in Figure 2: (**A**) subnetwork for PFL0780w, (**B**) subnetwork for PFB0340c, (**C**) subnetwork for MAL7P1.92, and (**D**) subnetwork for PF08_0002.

Table 1. Full details of the predicted TFs and their target genes.

TRANSCRIPTION FACTOR	TARGET GENES	GENE NAME	PRODUCT DESCRIPTION	ANNOTATED GO FUNCTION TERM	ANNOTATED GO PROCESS TERM
PFL0780w			Glycerol-3-phosphate dehydrogenase, putative	NAD binding, glycerol-3-phosphate dehydrogenase [NAD+] activity, protein homodimerization activity	Carbohydrate metabolic process, glycerol-3-phosphate catabolic process, glycerol-3-phosphate metabolic process, oxidation-reduction process
	PFD0955w	*Pf34*	Apical merozoite protein (Pf34)		
	PFE1605w	*LyMP*	Lysine-rich membrane-associated PHISTb protein (LyMP)	Protein binding	
	MAL7P1.87		Conserved Plasmodium protein, unknown function	ATP binding, actin binding, calmodulin binding, motor activity	
	PFI1745c		Early transcribed membrane protein		
	PF11_0351	*HSP70–3*	Heat shock protein 70 (HSP70–3)	ATP binding, unfolded protein binding	Protein folding, response to heat, response to unfolded protein
PFB0340c		*SERA5*	Serine repeat antigen 5 (SERA5)	Cysteine-type peptidase activity	Immunoglobulin production, proteolysis, regulation of immune response

(*Continued*)

Table 1. (*Continued*)

TRANSCRIPTION FACTOR	TARGET GENES	GENE NAME	PRODUCT DESCRIPTION	ANNOTATED GO FUNCTION TERM	ANNOTATED GO PROCESS TERM
	PF11_0202		AP-4 complex subunit mu, putative	Protein binding	Intracellular protein transport, vesicle-mediated transport
	PF14_0758	*PTP3*	EMP1-trafficking protein (PTP3)		
	PF10_0013		Plasmodium exported protein (hyp12), unknown function (PfJ13)		
	PF11_0225	*GCN20*	Protein GCN20 (GCN20)	ATP binding, ATPase activity, coupled to transmembrane movement of substances	Regulation of translation, transport
	PF13_0074	*SURF13.1*	Surface-associated interspersed protein 13.1 (SURFIN 13.1), pseudo-gene (SURF13.1)		
MAL7P1.92		*CRMP2*	Cysteine repeat modular protein 2 (CRMP2)	Protein binding	Intracellular receptor mediated signal-ling pathway, intracellular transport
	PFD1160w	*SURF4.2*	Surface-associated interspersed protein 4.2 (SURFIN 4.2) (SURF4.2)		
	PFB0680w	*RON6*	Rhoptry neck protein 6 (RON6)		Entry into host
	PF11_0168	*RON4*	Rhoptry neck protein 4 (RON4)		
	PF08_0036	*SEC23*	Protein transport protein SEC23 (SEC23)	Zinc ion binding	ER to Golgi vesicle-mediated trans-port, intracellular protein transport
	PF14_0529		AP-1 complex subunit gamma, putative	Binding, protein transporter activity	Intracellular protein transport, vesicle-mediated transport
PF08_0002		*SURF8.2*	Surface-associated interspersed protein 8.2 (SURFIN 8.2) (SURF8.2)	Binding, protein transporter activity	
	MAL7P1.164		AP-4 complex subunit beta, putative	Binding, protein transporter activity	Intracellular protein transport, vesicle-mediated transport
	PF10_0352	*MSP11*	Merozoite surface protein (MSP11)		
	PFC1090w	*EPF4*	Exported protein family 4 (EPF4)		
	PF08_0110	*RAB18*	Ras-related protein Rab-18 (RAB18)	GTP binding, GTPase activity	GTP catabolic process, intracellular protein transport, nucleocytoplasmic transport, regulation of vesicle-mediated transport, small GTPase mediated signal transduction

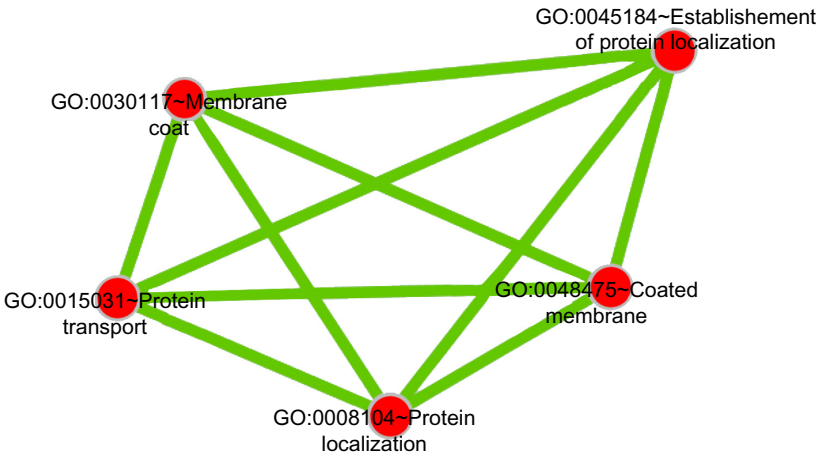

Figure 4. Enrichment analysis map of all predicted TFs and their target genes for the MCs' pathway at a significant level of $P = 0.001$ and $q = 0.05$. The enrichment map gives a graphical representation of the GO terms that were found to be significantly overrepresented in the context of the GO hierarchy.

membrane coat, protein localization, coated membrane, and establishment of protein localization.

Conclusion

We have successfully predicted the GRN of a very important pathway in the deadly malaria parasite *P. falciparum*. A number of predictions have been made from our analyses that should be further validated experimentally. Our results have also confirmed findings from previous studies and yielded new information, which shed more light and help improve our understanding of the genetic regulation activities of *P. falciparum* in our quest to fully overcome the malaria endemic. The knowledge of the predicted TFs, their target genes, and their predicted binding sites will indeed increase our understanding of the mechanism that underlay the tight transcriptional control of *P. falciparum* and also help in the rational design of promoters to ease the study of essential genes of the parasite.

Acknowledgments

The authors would like to acknowledge the contribution of the reviewers in improving the quality of this article.

Author Contributions

All authors contributed to this work. Original idea and conception: II, JO, BB, EA. Implementation of research methods: II. Writing and revision of the manuscript: II, JO. All authors reviewed and approved the final manuscript.

Supplementary Material

Supplementary File S01. Gene_list.txt.
Supplementary File S02. Computational pipeline.pdf.
Supplementary File S03. MC-GRN_sourcecode.R.
Supplementary File S04. Filtered_network.xlsx.
Supplementary File S05. Motifs.txt.

REFERENCES

1. Maurer G. Die malaria perniciosa. *Zentralbl Bakteriol Parasitenkunde*. 1902;23:695–719.
2. Trager W, Rudzinska MA, Bradbury PC. The fine structure of *Plasmodium falciparum* and its host erythrocyte in natural malarial infections in man. *Bull World Health Organ*. 1966;35(3):883–5.
3. Langreth SG, Jensen JB, Reese RT, Trager W. Fine structure of human malaria in vitro. *J Protozool*. 1978;25(4):443–52.
4. Lanzer M, Wickert H, Krohne G, Vincensini L, Braun Breton C. Maurer's clefts: a novel multi-functional organelle in the cytoplasm of *Plasmodium falciparum*-infected erythrocytes. *Int J Parasitol*. 2006;36(1):23–36.
5. Ewejobi IM, Bulashevska S, Brors B, Adebiyi EF. In-silico prediction of the genetic regulatory interactions in Maurer's cleft pathway of *Plasmodium falciparum*. *Proceeding of the EMBnet-RIBio*. Vol 16. EMBnet.journal, Cancun, Mexico; 2009.
6. Frischknecht F, Lanzer M. The *Plasmodium falciparum* Maurer's clefts in 3D. *Mol Microbiol*. 2008;67(4):687–91.
7. Mundwiler-Pachlatko E, Beck H-P. Maurer's clefts, the enigma of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A*. 2013;110(50):19987–94.
8. Sam-Yellowe TY. The role of the Maurer's clefts in protein transport in *Plasmodium falciparum*. *Trends Parasitol*. 2009;25(6):277–84.
9. Miller EA, Beilharz TH, Malkus PN, et al. Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. *Cell*. 2003;114(4):497–509.
10. Kirchgatter K, Del Portillo HA. Clinical and molecular aspects of severe malaria. *An Acad Bras Cienc*. 2005;77(3):455–75.
11. Haldar K, Mohandas N, Bhattacharjee S, et al. Trafficking and the tubulovesicular membrane network. In: Sherman IW, ed. *Molecular Approaches to Malaria*. 1st ed. Washington, DC: ASM Press; 2005:253–71.
12. Young JA, Johnson JR, Benner C, et al. In silico discovery of transcription regulatory elements in *Plasmodium falciparum*. *BMC Bioinformatics*. 2008;9:70.
13. Iengar P, Joshi NV. Identification of putative regulatory motifs in the up-stream regions of co-expressed functional groups of genes in *Plasmodium falciparum*. *BMC Genomics*. 2009;10:18.
14. Coulson RMR, Hall N, Ouzounis CA. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Res*. 2004;14:1548–54.
15. Bischoff E, Vaquero C. In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocyctic development of *Plasmodium falciparum*. *BMC Genomics*. 2010;11:34.
16. Schlitt T, Brazma A. Current approaches to gene regulatory network modeling. *BMC Bioinformatics*. 2007;8(suppl 6):S9.
17. Davidson E, Levin M. Gene regulatory networks. *Proc Natl Acad Sci U S A*. 2005;102(14):4935.
18. Albert R. Boolean modeling of genetic regulatory networks. Ben-Naim E, Frauenfelder H, Toroczkai Z, eds. *Complex Networks*. Berlin: Springer; 2004:459–81.
19. Lovrics A, Gao Y, Juhász B, et al. Boolean modelling reveals new regulatory connections between transcription factors orchestrating the development of the ventral spinal cord. *PloS One*. 2013;9(11):e111430.
20. Durzinsky M, Wagler A, Marvan W. Reconstruction of extended petri nets from time series data and its application to signal transduction and to gene regulatory networks. *BMC Syst Biol*. 2011;5:113.
21. Marwan W, Rohr C, Heiner M. Petri nets in snoopy: a unifying framework for the graphical display, computational modelling, and simulation of bacterial regulatory networks. van Helden J, Toussaint A, Thieffry D, eds. *Bacterial Molecular Networks*. New York: Springer; 2012:409–37.
22. Xu R, Wunsch IID, Frank R. Inference of genetic regulatory networks with recurrent neural network models using particle swarm optimization. *IEEE/ACM Trans Comput Biol Bioinform*. 2007;4(4):681–92.
23. Cheng L, Hou ZG, Lin Y, Tan M, Zhang WC, Wu FX. Recurrent neural network for non-smooth convex optimization problems with application to the identification of genetic regulatory networks. *IEEE Trans Neural Netw*. 2011;22(5):714–26.
24. Li Z, Li P, Krishnan A, Liu J. Large-scale dynamic gene regulatory network inference combining differential equation models with local dynamic Bayesian network analysis. *Bioinformatics*. 2011;27(19):2686–91.
25. Cao J, Qi X, Zhao H. Modeling gene regulation networks using ordinary differential equations. Wang J, Tan AC, Tian T, eds. *Next Generation Microarray Bioinformatics*. Humana Press, Totowa, NJ; 2012:185–97.
26. Butte AS, Kohane IS. Relevance networks: a first step toward finding genetic regulatory networks within microarray data. In: Parmigiani G, Garett ES, Irizarry RA, Zeger SL, eds. *The Analysis of Gene Expression Data*. New York: Springer; 2003:428–46.
27. Ma S, Gong Q, Bohnert HJ. An *Arabidopsis* gene network based on the graphical Gaussian model. *Genome Res*. 2007;17(11):1614–25.
28. Ingkasuwan P, Netrphan S, Prasitwattanaseree S, et al. Inferring transcriptional gene regulation network of starch metabolism in *Arabidopsis thaliana* leaves using graphical Gaussian model. *BMC Syst Biol*. 2012;6:100.
29. Krouk G, Lingeman J, Colon AM, Coruzzi G, Shasha D. Gene regulatory networks in plants: learning causality from time and perturbation. *Genome Biol*. 2013;14(6):123.
30. Brazhnik P, de la Fuente A, Mendes P. Gene networks: how to put the function in genomics. *Trends Biotechnol*. 2002;20(11):467–72.
31. Harel A, Kenett RS, Ruggeri F. Modeling web usability diagnostics on the basis of usage statistics. In: Jank W, Shmueli G, eds. *Statistical Methods in e-Commerce Research*. Hoboken, NJ: John Wiley & Sons, Inc; 2008:131–72.
32. Gallo CA, Carballido JA, Ponzoni I. Inference of gene regulatory networks based on association rules. In: Elloumi M, Zomaya AY, eds. *Biological Knowledge Discovery Handbook: Preprocessing, Mining, and Postprocessing of Biological Data*. Hoboken, NJ: John Wiley & Sons, Inc; 2013:803–40.
33. Armaanzas R, Inza I, Larraaga P. Detecting reliable gene interactions by a hierarchy of Bayesian network classifiers. *Comput Methods Programs Biomed*. 2008;91(2):110–21.
34. Djebbari A, Quackenbush J. Seeded Bayesian networks: constructing genetic networks from microarray data. *BMC Syst Biol*. 2008;2:57.
35. Dojer N, Gambin A, Mizera A, Wilczyński B, Tiuryn J. Applying dynamic Bayesian networks to perturbed gene expression data. *BMC Bioinformatics*. 2006;7:249.
36. Kim S, Imoto S, Miyano S. Dynamic Bayesian network and nonparametric regression for nonlinear modeling of gene networks from time series gene expression data. *Biosystems*. 2004;75(1–3):57–65.

37. Wang M, Chen Z, Cloutier S. A hybrid Bayesian network learning method for constructing gene networks. *Comput Biol Chem*. 2007;31(5–6):361–72.
38. Werhli AV, Husmeier D. Reconstructing gene regulatory networks with Bayesian networks by combining expression data with multiple sources of prior knowledge. *Stat Appl Genet Mol Biol*. 2007;6(1):15.
39. Lee WP, Yang KC. A clustering-based approach for inferring recurrent neural networks as gene regulatory networks. *Neurocomputing*. 2008;71(4–6):600–10.
40. Neapolitan RE. *Learning Bayesian Networks*. Upper Saddle River, NJ: Prentice Hall; 2003.
41. Ristevski B. A survey of models for inference of gene regulatory networks. *Nonlinear Anal Model Control*. 2013;18(4):444–65.
42. Zhang Y, Deng Z, Jiang H, Jia P. Inferring gene regulatory networks from multiple data sources via a dynamic Bayesian networks with structural EM. *Lect Notes Comput Sci*. 2007;4544:204–14.
43. Zou M, Conzen SD. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics*. 2005;21(1):71–9.
44. Whittaker J. *Graphical Models in Applied Multivariate Statistics*. New York: Wiley; 1990.
45. Toh H, Horimoto K. Inference of a genetic network by a combined approach of cluster analysis and graphical Gaussian modeling. *Bioinformatics*. 2002;18(2):287–97.
46. Kishino H, Waddell PJ. Correspondence analysis of genes and tissue types and finding genetic links from microarray data. *Genome Inform Ser Workshop Genome Inform*. 2000;11:83–95.
47. Magwene PM, Kim J. Estimating genomic coexpression networks using first-order conditional independence. *Genome Biol*. 2004;5(12):R100.
48. Schäfer J, Strimmer K. Learning large-scale graphical Gaussian models from genomic data. In: Mendes J, ed. *Proceedings of "Science of Complex Networks: From Biology to the Internet and WWW" (CNET 2004)*. Aveiro, Portugal: The American Institute of Physics; 2005:263–76.
49. Wille A, Buhlmann P. Low-order conditional independence graphs for inferring genetic networks. *Stat Appl Genet Mol Biol*. 2006;5:1.
50. Schäfer J, Strimmer K. An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics*. 2005;21(6):754–64.
51. Schäfer J, Strimmer K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat Appl Genet Mol Biol*. 2005;4:Article32.
52. Schäfer J, Opgen-Rhein R, Strimmer K. Reverse engineering genetic networks using the "GeneNet" package. *R News*. 2006;6(5):50–3.
53. Opgen-Rhein R, Strimmer K. From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data. *BMC Syst Biol*. 2007;1:37.
54. Foth BJ, Zhang N, Chaal BK, Sze SK, Preiser PR, Bozdech Z. Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics*. 2011;10(8):M110.006411.
55. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*. 2004;5(10):R80.
56. RStudio. *RStudio: Integrated Development Environment for R (Version 0.98.978) [Computer software]*. Boston, MA: 2014. Available at: http://www.rstudio.org/.
57. Davis S, Meltzer PS. GEOquery: a bridge between the gene expression omnibus (GEO) and bioconductor. *Bioinformatics*. 2007;14:1846–7.
58. Schäfer J, Opgen-Rhein R, Strimmer K. *GeneNet: Modeling and Inferring Gene Networks*. 2015. R package version 1.2.12.
59. Gansner ER, North SC. An open graph visualization system and its applications to software engineering. *Software Pract Exp*. 2000;30(11):1203–33.
60. Fang H, Gough J. The 'dnet' approach promotes emerging research on cancer patient survival. *Genome Med*. 2014;6:64.
61. Swinton J. *Vennerable: Venn and Euler Area-Proportional Diagrams*. 2013. R package version 3.0/r82.
62. Csardi G, Nepusz T. The igraph software package for complex network research. *Int Complex Syst*. 2006:1695(5).
63. Ameur A. *BCRANK: Predicting Binding Site Consensus from Ranked DNA Sequences*. 2010. R package version 1.26.0.
64. Aoki S, Li J, Itagaki S, et al. Serine repeat antigen (SERA5) is predominantly expressed among the SERA multigene family of *Plasmodium falciparum*, and the acquired antibody titers correlate with serum inhibition of the parasite growth. *J Biol Chem*. 2002;277(49):47533–40.
65. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
66. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One*. 2010;5(11):e13984.