A Network Approach to Wound Healing

Tomasz Arodz,^{1,2} Danail Bonchev,^{2,3} and Robert F. Diegelmann^{2,4,*}

Departments of ¹Computer Science and ⁴Biochemistry & Molecular Biology, ²VCU Reanimation, Engineering, and Science Center, and ³Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, Virginia.

Objective: The wound healing process is well-understood on the cellular and tissue level; however, its complex molecular mechanisms are not yet uncovered in their entirety. Viewing wounds as perturbed molecular networks provides the tools for analyzing and optimizing the healing process. It helps to answer specific questions that lead to better understanding of the complexity of the process. What are the molecular pathways involved in wound healing? How do these pathways interact with each other during the different stages of wound healing? Is it possible to grasp the entire mechanism of regulatory interactions in the healing of a wound?

Approach: Networks are structures composed of nodes connected by links. A network describing the state of a cell taking part in the healing process may contain nodes representing genes, proteins, microRNAs, metabolites, and drug molecules. The links connecting nodes represent interactions such as binding, regulation, co-expression, chemical reaction, and others. Both nodes and links can be weighted by numbers related to molecular concentration and the intensity of intermolecular interactions. Proceeding from data and from molecular profiling experiments, different types of networks are built to characterize the stages of the healing process. Network nodes having a higher degree of connectivity and centrality usually play more important roles for the functioning of the system they describe.

Results: We describe here the algorithms and software packages for building, manipulating and analyzing networks proceeding from information available from a literature or database search or directly extracted from experimental gene expression, metabolic, and proteomic data. Network analysis identifies genes/proteins most differentiated during the healing process, and their organization in functional pathways or modules, and their distribution into gene ontology categories of biological processes, molecular functions, and cellular localization. We provide an example of how network analysis can be used to reach better understanding of regulation of key wound healing mediators and microRNAs that regulate them.

Innovation: Univariate statistical tests widely used in clinical studies are not enough to improve understanding and optimize the processes of wound healing. Network methods of analysis of patients "omics" data, such as transcriptoms, proteomes, and others can provide a better insight into the healing processes and help in development of better treatment practices. We review several articles that are examples of this emergent approach to the study of wound healing.

Conclusion: Network analysis has the potential to considerably contribute to the better understanding of the molecular mechanisms of wound healing and to the discovery of means to control and optimize that process.

Robert F. Diegelmann, PhD

Submitted for publication November 6, 2012.
Accepted in revised form December 22, 2012.

*Correspondence: Department of Biochemistry & Molecular Biology, Virginia Commonwealth University Medical Center, 1101 East Marshall St., Richmond, VA 23235-3314 (e-mail: rdieglm@vcu.edu).

Abbreviations and Acronyms

ARACNE = algorithm for the reconstruction of accurate cellular networks

BioGRID = biological general repository for interaction datasets

CAM = chorioallantoic membrane

CD4 = cluster of differentiation 4

CEBPA = CCAAT enhancer-binding protein alpha

CLR algorithm = context likelihood of relatedness algorithm

DAVID = database for annotation, visualization, and integrated discovery

EMBL-EBI = European Molecular Biology Laboratory—European Bioinformatics Institute

(continued)

Abbreviations and Acronyms (*continued*)

ESR1=estrogen specific receptor 1
 FOSL1=fos-like protein 1
 GENIE3= gene network inference with ensemble of trees
 GO= gene ontology
 GSEA= gene set enrichment analysis
 HNRNPK= heterogeneous nuclear ribonucleoprotein K
 HOXA9= homeobox protein Hox-A9
 IL4= interleukin 4
 IL10= interleukin 10
 IRAK= interleukin receptor-associated kinase
 KEGG= Kyoto encyclopedia of genes and genomes
 KHSRP= KH splicing regulatory protein
 MAPK8= mitogen-activated protein kinase 8
 MCP-1= monocyte chemoattractant protein 1
 MITF= microphthalmia-associated transcription factor
 MRMR= maximum relevance/minimum redundancy
 MRNET= MRMR network
 NF1A= nuclear factor 1-alpha
 PINA= protein interaction network analysis
 STAT3= signal transducer and activator of transcription 3
 SRF= serum response factor
 TGFBR1= transforming growth factor receptor 1
 TNF= tumor necrosis factor
 TP53= tumor protein 53
 TRAF6= TNF receptor-associated factor 6
 VEGFA= vascular endothelial growth factor A

INTRODUCTION

THE HIGHLY ORCHESTRATED sequence of events that occur during normal, acute wound healing have been known for a very long time.¹ Likewise, the pathologic features that are characteristic of delayed wound healing as observed in chronic pressure ulcers, diabetic wounds, and venous stasis ulcers are also well known. On the opposite end of the spectrum the aberrant over healing seen in fibrotic responses such as keloid, hypertrophic scar, strictures, and adhesions are well described on a pathologic level.² The various cells, cytokines, and the spectrum of enzymes that participate in both normal and pathologic wound repair are also well known. What is not well known are the specific signaling switches that control all of these various normal and abnormal responses.

CLINICAL PROBLEM ADDRESSED

To achieve better understanding of processes in wound healing and ultimately arrive at better wound management and treatment approaches, it is not enough to proceed from analysis of experimental data based on univariate statistical tests. Novel, more elaborate, network-oriented methods are needed. We review two articles^{3,4} that apply network approach to analyze wound healing, involving gene expression and proteomic profiling, and a third article⁵ that focuses on relations between microRNAs and genes involved in the processes of inflammation and angiogenesis during wound healing.

MATERIALS AND METHODS

Relevant basic science context

The last decade in biological and biomedical studies has been characterized by a rethinking of the basic approaches and a shift in the main

focus of studies from individual objects (genes, proteins, and RNAs) to the system as a whole (cell, tissue, and organ). This change in the paradigm is termed a transition from reductionism to systems approach.⁶ *Systems Biology*⁷⁻⁹ underscores the new, emergent properties in the whole, properties/functions that are not available in the individual parts of the system. More generally, systems biology studies all the flows of matter, energy, and information within the living species, and the organization of these flows in space and time. *Networks* are the most natural language for describing functional systems of any kind. They are structures built of *nodes*, representing the individual elements of the system, and *links*, characterizing the interactions between these elements. The structure of a network and its major parts (modules and pathways) is intimately related to function. Complex networks of any kind share common properties and basic patterns.¹⁰ Highly connected nodes (termed *hubs*) play major roles in complex systems both for good and for bad. They are crucial for the functioning of the system, but upon being directly attacked—say, a pathogen releases a toxin that inhibits the activity of a key protein in the cell, or a mutation silences a key gene—they also could contribute to the malfunctioning of the whole system and even to systems disintegration. Complex networks are highly connected, which greatly reduces their diameter, that is, the number of connections that need to be traversed to reach from any node in the network to any other node. As a result, any interaction from the environment can be spread out quickly through the network. This process proceeds to a higher degree via more centrally located nodes, so *node centrality* is another major local characteristic, along with *node connectivity*. These net-

Figure 1. Schematic view of network concepts: nodes, edges, hubs, centrality, and connectivity.

work concepts are schematically depicted in Fig. 1. A network approach to diseases views the disease as a *perturbed network*.¹¹ Redefining human diseases in terms of perturbed networks is one of the major tasks of network biology. A closely related task is the change in the strategy of drug design from “one defective gene/protein, one drug molecule” to a combination of drugs accounting for the connectivity of the malfunctioning gene/protein, thus reducing the unavoidable side effects.

A network approach is most often used in conjunction with “-omics” data from high-throughput profiling experiments. One central characteristic of profiling experiments is the large number of variables that are being measured. It reaches from hundreds of endpoints in proteomic studies, to tens of thousands of probes in transcriptomic studies of gene expression. This leads to problems with interpretation of the results that involve large number of intercorrelated variables—for example, expression levels for all genes measured with a microarray. Network analysis approaches, essentially, take the structure of known or data-extracted relationships between variables, and use that structure to illuminate the results of profiling studies.

Generally, the applications of networks to disease studies are concerned not only with identifying individual new disease genes, but also in studying the properties of the “disease networks,” identifying disease-related subnetworks (modules and pathways), and discovery of markers for early identification of diseases.¹² Biological networks are defined by the structure and types of connections between individual nodes, for example, genes or proteins. Many methods do not reach this high level of detail, instead operating at the level of

groups of genes or proteins, without going into specifics of the structure of their interconnections. Such methods can be put broadly into two categories—data-driven and knowledge-based.

In knowledge-based approaches, expression data is analyzed in the context of existing, predefined groupings of genes or proteins in the form of gene ontology (GO) or signaling pathways. Knowledge-based techniques aim at extrapolating information from statistical analysis of individual genes to a higher-level view by using the concept of gene sets and gene set enrichment. Gene set enrichment analysis (GSEA)^{13,14} requires that individual genes significantly correlated with a given trait are first identified using some statistical test. Given a group of identified statistically significant genes for a given trait, GSEA proceeds to quantify in which predefined function- or pathway-related gene sets the significant ones are over-represented. Enrichment analysis identifies the top up- and downregulated genes/proteins in patients as compared with “normal” samples, and their distribution into specific pathways, and into GO categories, as a basis for mechanistic hypotheses. The GO categories include detailed lists of biological processes, molecular functions, and cellular localizations that are most strongly affected in comparison with healthy state. This type of analysis can be performed either with the software packages for network analysis mentioned below or with other ones like DAVID (Database for Annotation, Visualization, and Integrated Discovery), which is specialized in GO analysis.¹⁵

Data-driven techniques take a different approach—they aim to detect sets of genes with similar behavior based only on the experimental results. This allows for discovery of clusters that do not correspond to predefined reference groups, but makes the results more susceptible to noise, especially in low-sample-size scenarios. The user has a choice of many clustering methods, with hierarchical clustering and k-means being the most popular. Some methods aim to extend clustering to include annotation of genes as a factor that influences the resulting clusters.^{16,17} Bi-clustering can also be used to simultaneously cluster genes and samples before presenting them in the form of a heat map. The genes that cluster together can be then analyzed for similarity of their function based on gene annotations.

Fully network-oriented methods reach beyond clustering and enrichment analysis to the level of connectivity between individual proteins or genes. To build a relevant network, one needs relevant data that would allow distinguishing in full detail

the normal and disease state. Such are the data for gene expression, protein expression, and metabolite concentrations, along with information for node–node interactions. The latter can be very diverse ranging from gene expression regulation by transcription factors and microRNAs, to biochemical reactions of metabolites, to protein–protein specific and nonspecific interactions (regulatory, post-translational, signaling, van der Waals attraction), and to molecular transport.

Which are the specific ways to apply a network approach to identifying and curing diseases and injuries? Again, we can identify the same two broad categories of methods—data-driven and knowledge-based. The main distinction is in the role the network plays in the analysis. Data-driven methods focus on network inference, with the network being the result of analyzing the data. The knowledge-based approaches start with a large structure of connections reflecting current knowledge of biology, gathered and curated from databases and literature. That background knowledge serves as a basis for deriving a focused network that is related to the results of the experiment under study.

Network inference relies on data from profiling experiments on a studied group of subjects. The simplest inference method is relevance network,¹⁸ which ranks all the possible gene pairs based on their mutual information and predicts the presence of a regulatory interaction if the coefficient is higher than a given threshold. The ARACNE (algorithm for the reconstruction of accurate cellular networks) method^{19,20} was proposed as an improvement to relevance networks. It removes possibly indirect edges from triplets of genes with the use of data processing inequality. The CLR (context likelihood of relatedness) algorithm,²¹ on the other hand, applies an adaptive background correction step. Another example of information-theoretic algorithm is a maximum relevance/minimum redundancy (MRMR) network (termed MRNET),²² which formulates the network inference problem as a series of supervised gene selection procedures, adopting an MRMR principle.²³ On the other hand, a development of machine learning theory, especially classification, regression, and feature selection theory, successfully applied in various other domains, brought promising new possibilities to the field of inferring biological networks. A good example is the recently published GENIE3 (gene network inference with ensemble of trees) algorithm,²⁴ which makes use of regression trees to score important transcription factors. The methods listed above may not agree on the exact structure of the inferred network, and studies suggest that a meta-network, con-

structed by aggregating results from several methods, leads to higher accuracy.

Knowledge-based methods also make use of experimental data, but they interpret them in the context of information taken from the online accessible databases. Here, one may list the PINA (protein interaction network analysis) database for protein–protein interactions;²⁵ IntAct—the European Molecular Biology Laboratory—European Bioinformatics Institute (EMBL-EBI) Protein Interaction DB;²⁶ The BioGRID (biological general repository for interaction datasets) database for genetic and physical interactions;²⁷ KEGG (Kyoto encyclopedia of genes and genomes) for metabolic information²⁸ to mention a few. Some of the best professional software packages for network analysis, such as Ingenuity²⁹ and Pathway Studio,³⁰ also provide their databases for analysis with their proprietary software. Another popular software package for network analysis is Cytoscape,³¹ which offers a number of additional functions (plugins) and is publicly available.

The essence of knowledge-based methods is not to discover new links, but to find out which subset of known connections form a subnetwork that differentiates between phenotypes, disease states, or time-points. One of the first approaches to integration of experimental gene expression data with prior knowledge in a form of protein–protein interaction network was focused on identifying highly active subnetworks.³² The activity of the subnetwork was measured as an average of expression of individual genes, and highly active groups of genes were identified using simulated annealing. Another method uses spectral mapping of expression profiles on eigenfunctions of gene network to detect subnetworks, which are later used as features in a supervised machine learning framework based on support vector machines.³³ Other methods directly overlay expression data on protein–protein interactions and perform a greedy search to find subnetworks with discriminatory power.^{34–36} The concept of random walk on a protein–protein interaction network was recently employed to partition the network of genes into groups with suspected functional similarity. Once the subnetworks are specified, the averaged expression in a subnet is evaluated for its predictive ability using an information-theoretic criterion.³⁷

As the final step of the analysis, the networks produced from experimental or literature data can be further manipulated to include the closest neighborhood of the most important nodes or to find common regulators and common targets of the genes/proteins/microRNAs these nodes represent.

Table 1. Software tools for network analysis

Software Tool	Main Functionality
DAVID ¹⁵	Gene set enrichment analysis (GSEA) based on Gene Ontology
ARACNE ^{19,20}	Inference of transcription regulatory network from gene expression
CLR ²¹	Inference of transcription regulatory network from gene expression
MRNET ²²	Inference of transcription regulatory network from gene expression
GENIE3 ²⁴	Inference of transcription regulatory network from gene expression
Ingenuity ²⁹	Inference of from gene expression, proteomic data and proprietary database. GSEA
Pathway Studio ³⁰	Inference of protein–protein interaction networks (direct interaction, shortest paths, common regulators and common targets, miRNA/protein targets), from gene expression, proteomic data and proprietary database. GSEA
Cytoscape ³¹	Visualization of biological networks; plugins extend the functionality to various types of network calculations and analysis
jActiveModules ³² (Cytoscape plugin)	Discovery of active modules in a regulatory or signaling network

DAVID, database for annotation, visualization, and integrated discovery; ARACNE, algorithm for the reconstruction of accurate cellular networks; CLR, context likelihood of relatedness algorithm; MRNET, maximum relevance/minimum redundancy network; GENIE3, gene network inference with ensemble of trees.

In such a way, predictions can be made for other genes involved in the disease of interest and its curing, and novel drug candidates can be identified. All this will bring medicine much closer to individualized and more efficient treatments. Examples of different cases of network analysis are given in the rest of the article, while the summary of software tools that can be used to facilitate analysis are listed in Table 1. A comparative analysis of some of these tools is given.³⁸

Experimental model or material: advantages and limitations

Network analysis relies on experimental measurements of quantities that build the network. Most often, these are captured using profiling experiments. Soulet *et al.*³ relied on transcriptomic profiling using Affymetrix GeneChip microarray in their exploration of healing of wounds in the chick embryo model. Transcriptomic profiling has the advantage of capturing genome-wide profile of gene expression in a single experiment, but the results often require confirmation using quantitative real-time polymerase chain reaction (qPCR) for selected small subset of identified genes. Other studies of wound samples also rely on the combination of transcriptomic profiling using microarrays with qPCR validation step.^{39–41}

Measurements at the protein level show a more direct view of molecular processes involved in wound healing. Two complementary main approaches are being used for proteomic profiling. In the first one, wound samples are analyzed to uncover the protein profiles without any prior specification of which proteins are measured. Technical approaches for that goal are limited to proteins present with high abundance, and thus may fail to detect low-abundance species that are nonetheless crucial in the healing processes, such as cytokines.

One example of that approach is polyacrylamide gel electrophoresis followed up by mass spectrometry (MS) identification.⁴² Another example is MS coupled with non-gel techniques, such as isobaric tags for relative and absolute quantitation, used by Edsberg *et al.*⁴ A complementary approach that uses protein microarrays to construct proteomic profiles of wounds allows for detection of proteins with low concentrations, but on the other hand it requires that the set of proteins is decided prior to the quantification process. In a recent article, Edsberg *et al.*⁴ used several formats of antibody arrays, including glass and membrane printed microarrays, planar microarray, and microsphere-based suspension microarray, to measure presence and abundance of different proteins. The study highlights that differences in array format have substantial effect on detected proteins—in the case of samples analyzed using both glass and membrane support surfaces only 15% of proteins were detected by both methods. In fields outside the wound healing domain, another type of protein microarray, the reverse-phase protein microarray, is becoming popular for studying signaling pathways.

RESULTS AND DISCUSSION

The methods used in analyzing results from profiling studies oriented on wound healing are currently limited mostly to enrichment analysis using GO terms, and to clustering.^{39–41} However, studies that take a more detailed look at networks involved in response to wounds are emerging. Here, we review two such recent reports, one focused on proteomics and another on transcriptomic profiling.

A multi-modal proteomic profiling study by Edsberg *et al.*⁴ focused on pressure ulcers, tracked

over a time span of up to 6 weeks in 32 subjects, with wound fluid samples collected from the interior and separately from the periphery of the wounds. The network analysis was performed using Ingenuity,¹⁸ and focused on two aspects: uncovering differences in interior wound fluid between chronic and healed wounds, and, within chronic wounds, the differences between fluid samples collected from interior and periphery of the wound. Matrix metalloproteinases are more abundant in wounds that eventually healed, while tissue inhibitors of metalloproteinases and proinflammatory cytokines are more abundant in chronic wounds. Keratins and members of the S100 family are significantly more abundant in periphery of chronic wounds, while some matrix metalloproteinases and proinflammatory cytokines are overabundant in wound interior.

The design of the study limits the detected proteins to extracellular and, to a limited extent, membrane proteins. Using network analysis allows for linking the detected species to other proteins in cell membrane and to intracellular proteins that interact with them. One protein pointed to by network analysis is transcription factor AP-1, which is a complex that includes c-Jun protein. Matrix metalloproteinases are regulated by AP-1.⁴³ Appreciation of its role in wound healing is emerging. Activation of AP-1 through JNK signaling was shown crucial for dorsal closure in *Drosophila melanogaster*, which can serve as a model for wound repair.⁴⁴ In humans, it was shown to play a role in epidermal leading edge organization.⁴⁵

Another tightly linked intracellular protein pointed to by network analysis is nuclear factor (NF)- κ B transcription factor. NF- κ B is activated, among others, by tumor necrosis factor (TNF)-beta and interleukin (IL)-1. It has a known role in regulation of immune response, its activation leading to expression of inflammatory cytokines.⁴⁶ It also plays a role in activation of matrix metalloprotein (MMP)-9 and other matrix metalloproteinases.^{47–49} Activation of NF- κ B was directly observed upon wounding.⁵⁰

One other molecular species that was pointed by network analysis is fibrinogen. The relationship of fibrinogen to differences between chronic and healing wounds has been observed in other studies, with beta and gamma chains significantly more abundant in ulcers.⁴²

The second study we describe here, by Soulet *et al.*,³ is a transcriptomic profiling of wounded chorioallantoic membrane (CAM) of three chicken embryos and of unwounded, control CAMs in an-

other three embryos. The CAM model allows for observation of changes in vasculature and stromal fibroblasts. On top of enrichment analysis using DAVID,¹⁵ network analysis was performed using Ingenuity.²⁹ Here, the experimental method is genome-wide. The benefit of network approach comes not from identifying genes that were not measured, but from showing the interconnections between measured genes, and bringing focus to networks that may be regulated in a concerted manner even though not all individual elements are differentially expressed, or are not highly ranked in terms of fold change.

Network analysis of chicken embryo wound transcriptomes show that IL-1 and MMP-9 play central roles in the most highly scoring network, even though these two genes individually have low fold change between wound and control. NF- κ B is also highly linked to in the network. IL-1 also plays a prominent role in the second-highest scoring network, while IL-8 is a hub in the fourth-highest scoring network, which also includes NF- κ B as highly connected nodes.

Networks may include nodes that are not genes or proteins, but other entities such as hormones or cofactors, which are not measurable with proteomic or transcriptomic profiling. Network analysis can pinpoint importance of such entities in wound healing. In the analyzed study, retinoic acid and beta-estradiol form central hubs of highly-scored networks in the CAM wound model. Retinoic acid is known to improve wound healing,⁵¹ while estrogen was shown to be involved in age-related slowing of wound healing.⁴⁰

Network methods can extend beyond analysis of proteins and protein-coding genes, for example to include microRNAs. MicroRNAs, as recently reviewed by Roy and Sen,⁵ play a role in wound inflammation and angiogenesis. After discovery in 1993 by Ambros, Lee, and Feinbaum,⁵² microRNAs quickly became a field of intensive studies, which revealed their role as another level of regulation of processes in the living cell, and have the potential to bring new healing strategies for diseases, including wound inflammation and angiogenesis. In the reviewed target article,⁵ Roy and Sen followed one of the basic network approaches—the key role of the highly connected network nodes (hubs). More specifically, they emphasize the importance of some miRNAs in regulating several hubs, which in turn control the wound inflammation and angiogenesis processes. The review also singles out the role in wound inflammation of the regulatory loops in which by miRNAs and inflammatory

mediators elicited following injury mutually regulate their expression and activity.⁵

Some more ideas about applying network approaches to wound healing are presented in this and the following paragraph making use of data from the target article of Roy and Sen.⁵ MicroRNAs usually jointly act with transcription factors and other regulatory proteins to form a larger and more effective regulatory system. Table 1 of the target article presents a list of six key mediators of injury inflammation—TNF, transforming growth factor receptor 1, IL10, TNF receptor-associated factor 6, interleukin receptor-associated kinase, monocyte chemotactic protein 1, and miRNAs that regulate these proteins (miR-125b, miR-128a, miR-466l, miR-146a, and miR-124a). One may start with the list of the six proteins and five miRNAs as seed nodes and design such an integrated regulatory network. We shall apply first the “direct interactions” algorithm of the Pathway Studio software.³⁰ The ResNet 9.0 database of this software reproduces five of the six “seed” miRNA regulatory interactions, as shown by the red lines in Fig. 2a. Eight more regulatory interactions are added on this figure between pairs of proteins, along with the activation of miR-146A by IL-10, activation of miR-125B1 by TNF, and suppression of miR-128-1 by miR124-1. Using the “common regulators” algorithm of Pathway Studio 9.0, the same “seed” set of miRNA–protein interactions from Table 1 is expanded in Fig. 2b to include also regulatory interactions of three more miRNAs, 14 transcription

factors, six more receptor proteins, and five other proteins involved in the regulation of the “seed” proteins.

Similar expansion of the limited information about the miRNA expression and control of several important immune system proteins (PU.1, C/EBP α and NFIA) is shown in Fig. 3a, b. The first two proteins are denoted there with their systematic names—SPI1 and CCAAT enhancer-binding protein alpha (CEBPA), respectively.

The inflammatory response of the three proteins is shown to activate the transcription of miR-424, miR-21, and miR-223, while being under the regulatory control of miR-155. The overexpression of miR-223 is on the other hand shown to be controlled by auto-regulatory feedback loop, which enhances the granulocytic differentiation. As shown in Fig. 3b, the “common regulators” algorithm reveals the complexity of the immune system regulation during the inflammation process. Thus, miR-155 activity is modulated by a number of transcription factors (signal transducer and activator of transcription 3 [STAT3], homeobox protein Hox-A9, microphthalmia-associated transcription factor), one receptor protein (cluster of differentiation 4), several other proteins (vascular endothelial growth factor A, mitogen-activated protein kinase 8, v-akt murine thymoma viral oncogene homolog 1, KH splicing regulatory protein [KHSRP]), and two more miRNAs (miR-223 and miR-424). Similarly, miR-21 expression is influenced by regulatory interactions, which be-

Figure 2. Combined miRNAs and transcription factor regulation of inflammation in wound healing. **(a)** The “seed” miRNA regulatory network built with data from Table 1 of Roy and Sen.⁵ **(b)** Integrated regulatory network built by using Pathway Studio 9.0 software, which added three more miRNAs (highlighted in yellow), 13 transcription factors, 7 receptor proteins, and 5 other proteins and their relations. The proteins highlighted in blue and the miRNA in red represent the seed set. Relation colors: red, miRNA regulation; green, promoter binding; black, direct regulation; dotted line, indirect regulation. To see this illustration in color, the reader is referred to the web article at www.liebertpub.com/wound

Figure 3. The miRNA regulation of CCAAT enhancer-binding protein alpha (CEBPA), SPI1 and NFIA proteins activated in immune cells upon wound healing (discussed in the target article) is used in (a) as seed network. In (b), it is supplemented by regulation from transcription factors and other proteins by applying the common regulator algorithm of Pathway Studio 9.0 software. Relation colors: red, miRNA regulation; green, promoter binding; black, direct regulation; dotted line, indirect regulation. To see this illustration in color, the reader is referred to the web article at www.liebertpub.com/wound

sides SPI1 and CEBPA includes seven more transcription factors (Sp1 transcription factor, *Jun* proto-oncogene, STAT3 serum response factor, estrogen specific receptor 1, tumor protein 53, and fos-like protein 1) and three more proteins (IL4, KHSRP and heterogeneous nuclear ribonucleoprotein K). This wealth of regulators might ultimately lead to the design of new drugs and effective procedures for accelerated wound healing.

INNOVATION

The network approach is a paradigm that has been successfully applied to the study of regulation and signaling. The articles reviewed here are among the first efforts to bring networks into the study of wound healing. The innovation from applying network methods comes from its ability to enhance understanding of high-throughput proteomic or transcriptomic profiling. They can point to more links between observed genes or proteins, and provide relations to other entities not measured, due to limitations of the chosen experimental methods. Ultimately, network approaches can help for a deeper, integrated understanding of the wound healing process and make it more optimal.

CONCLUSIONS

Caution, critical remarks, and recommendations

Application of the network approach faces certain limits and caveats. Currently available reference networks are incomplete and not fully

accurate—the confidence in links varies widely, from interactions or regulation links that are thoroughly verified experimentally, to results of high-throughput experiments or computational inference that may not be accurate. Thus, some links in a known network may represent false positives—links that do not exist in reality. Another problem is false negatives—interactions that are missing from reference networks. Here, the scope of the problem is harder to quantify, but for example recent estimates of the number of protein–protein interactions are on the order of 100,000 for humans, yet not even 40,000 are currently known.⁵³

Another problem that affects network methods is the small sample size of typical experiments. This is a problem for analysis of profiling experiments in general, since the number of observed variables vastly exceeds the number of samples, but it becomes even more pronounced in methods that aim at inferring networks from data. For example, in differential expression study involving m genes classical approach has to make a decision on which of the m genes are differentially expressed. In network inference, the task is more daunting—there are m^2 possible direct connections, and thus m^2 decisions to be made based on data.

Caution is also advised when dealing with simple concepts such as centrality or hub status of nodes in a network. Studies confirm that high-degree genes are more likely to be associated with disease but this rule does not hold universally in all contexts. For example, genes associated with heritable genetic disease are enriched in hubs. However,

this correlation results from a subset of essential genes, for which any aberration leads to lethal phenotype. It disappears for nonessential heritable genetic disease genes, which do not exhibit increased connectivity. On the other hand, genes whose somatic mutations lead to cancer are very often hubs.⁵⁴

Future developments of interest

As this review shows, the application of network methods to basic research in wound healing is only in its beginning. Future profiling studies will benefit from incorporating systems biology approaches to data analysis. Once the human proteome is completed with all of the missing protein–protein interactions, the benefits of the network approach to wound healing and all diseases will be tremendous. Another expected direction is the emergence of integrative studies, which is already happening in the cancer domain. There, efforts such as the Cancer Genome Atlas⁵⁵ or NCI-60 CellMiner⁵⁶ integrate data at genetic, epigenetic, transcriptomic, and proteomic levels, and network methods are being adapted to deal with such multi-modal datasets.^{57,58} In the wound healing domain, we should expect studies that analyze together gene expression at mRNA level, proteins and their post-translational modifications, microRNA expression, and even epigenetic markers, as epigenetic reprogramming was recently observed in wound healing.⁵⁹ These would bring a more complete picture of the molecular processes involved in wound healing. The complexity of such multi-modal datasets makes the network approach an indispensable tool.

AUTHOR DISCLOSURE AND GHOSTWRITING

The authors have no commercial associations and have total responsibility for the preparation and writing of this article.

KEY FINDINGS

Basic science advances

Network analysis is an innovative strategy that has the potential to be used to better characterize the molecular switches that control normal and abnormal wound healing mechanisms.

Clinical science advances

At present, the network approach to wound healing is still in the developmental stage. Once the process is better understood, it is hoped to become a valuable tool for clinicians, enabling the treatment of many pathologic wound healing responses, such as fibrosis and chronic, nonhealing wounds.

Relevance to clinical care

A better understanding of the molecular signaling switches that control the wound repair process could help the development of new clinical approaches to optimize the healing process.

ABOUT THE AUTHORS

Tomasz Arodz, PhD, is an assistant professor in the Department of Computer Science, Virginia Commonwealth University. He holds PhD and MS degrees in Computer Science from AGH University of Science and Technology in Krakow, Poland, and an MS in Biotechnology from Jagiellonian University in Krakow. His current research interests include network methods and machine learning in bioinformatics. **Danail G. Bonchev, PhD, DSc**, is Professor of Mathematics and Director of Research in Bioinformatics at the Center for the Study of Biological Complexity, Virginia Commonwealth University. He holds a PhD in quantum chemistry from the Bulgarian Academy of Sciences in Sofia, and a *Doktor Nauk* of Chemical Sciences in mathematical chemistry from Moscow State University in Russia. **Robert F. Diegelmann, PhD**, is Professor of Biochemistry and Molecular Biology at Virginia Commonwealth University Medical Center and is a founding member of the Wound Healing Society and served as the president of the society in 2012–2013. His current research interests include chronic non-healing wounds and combat casualty wounds.

REFERENCES

1. Diegelmann RF and Evans MC: Wound healing: an overview of acute, fibrotic and delayed healing. *Front Biosci* 2004; **9**: 283.
2. Ehrlich HP, Desmoulière A, Diegelmann RF, Cohen IK, Compton CC, Garner WL, Kapanci Y, and Gabbiani G: Morphological and immunochemical differences between keloid and hypertrophic scar. *Am J Pathol* 1994; **145**: 105.
3. Soulet F, Kilarski WW, Antczak P, Herbert J, Bicknell R, Falciani F, and Bikfalvi A: Gene signatures in wound tissue as evidenced by molecular profiling in the chick embryo model. *BMC Genomics* 2010; **11**: 495.
4. Edsberg LE, Wyffels JT, Brogan MS, and Fries KM: Analysis of the proteomic profile of chronic pressure ulcers. *Wound Repair Regen* 2012; **20**: 378.
5. Roy S and Sen CK: miRNA in wound inflammation and angiogenesis. *Microcirculation* 2012; **19**: 224.
6. Barabási A-L: The network takeover. *Nat Physics* 2012; **8**: 14.
7. Palsson B: *Systems Biology: Properties of Reconstructed Networks*. Cambridge: Cambridge University Press, 2006.

8. Alon U: *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC, Boca Raton, FL: Taylor and Francis Group, 2006.
9. Kitano H: Computational systems biology. *Nature* 2002; **420**: 206.
10. Estrada E: *The Structure of Complex Networks*. New York: Oxford University Press, 2012.
11. Hood L: Presentation on the First Summit on Systems Biology, Virginia Commonwealth University, May 2007.
12. Ideker T and Sharan R: Protein networks in disease. *Genome Res* 2008; **18**: 644.
13. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* 2005; **102**: 15545.
14. Ackermann M and Strimmer K: A general modular framework for gene set enrichment analysis. *BMC Bioinform* 2009; **10**: 47.
15. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat Genet* 2000; **25**: 25 (<http://david.abcc.ncifcrf.gov>).
16. Huang D and Pan W: Incorporating biological knowledge into distance-based clustering analysis of microarray gene expression data. *Bioinformatics* 2006; **22**: 1259.
17. Tseng GC: Penalized and weighted K-means for clustering with scattered objects and prior information in high-throughput biological data. *Bioinformatics* 2007; **23**: 2247.
18. Butte AJ, Kohane IS: Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. *Pacific Symp Biocomputing* 2000; **5**: 415.
19. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera RD, and Califano A: ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinform* 2006; **7** (Suppl 1): S7.
20. Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, and Califano A: Reverse engineering cellular networks. *Nat Protocols* 2006; **1**: 662.
21. Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, Kasif S, Collins JJ, and Gardner TS: Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol* 2007; **5**: e8.
22. Meyer PE, Kontos K, Lafitte F, and Bontempi G: Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J Bioinform Syst Biol* 2007: Article 79879.
23. Ding C and Peng H: Minimum redundancy feature selection from microarray gene expression data. *J Bioinform Comput Biol* 2005; **03**: 185.
24. Huynh-Thu VA, Irrthum A, Wehenkel L, and Geurts P: Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* 2010; **5**: e12776.
25. Wu J, Vallenius T, Ovaska K, Westermarck J, Makela TP, and Hautaniemi S: Integrated network analysis platform for protein-protein interactions. *Nat Methods* 2000; **6**: 75.
26. Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U, Jandrasits C, Jimenez RC, Khadake J, Mahadevan U, Masson P, Pedruzzi I, Pfeifferberger E, Porras P, Raghunath A, Roechert B, Orchard S, and Hermjakob H: The IntAct molecular interaction database in 2012. *Nucleic Acids Res* 2011; **40**: D841.
27. Stark C, Breitkreutz BJ, Chatr-Aryamontri A, Boucher L, Oughtred R, Livstone MS, Nixon J, Van Auken K, Wang X, Shi X, Reguly T, Rust JM, Winter A, Dolinski K, and Tyers M: The BioGRID interaction database: 2011 update. *Nucleic Acids Res* 2011; **39**: D698.
28. Kanehisa M, Goto S, Sato Y, Furumichi M, and Tanabe M: KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Res* 2012; **40**: D109.
29. IPA: Ingenuity Systems. www.ingenuity.com
30. Nikitin A, Egorov S, Daraselia N, and Mazo I: Pathway Studio—the analysis and navigation of molecular networks. *Bioinformatics* 2003; **19**: 2155 (www.ariadnegenomics.com).
31. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T: Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; **13**: 2498.
32. Ideker T, Ozier O, Schwikowski B, and Siegel AF: Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 2002; **18** Suppl 1: S233.
33. Rapaport F, Zinoviyev A, Dutreix M, Barillot E, and Vert J-P: Classification of microarray data using gene networks. *BMC Bioinform* 2007; **8**: 35.
34. Chuang HY, Lee E, Liu YT, Lee D, and Ideker T: Network-based classification of breast cancer metastasis. *Mol Syst Biol* 2007; **3**: 140.
35. Guo Z, Li Y, Gong X, Yao C, Ma W, Wang D, Li Y, Zhu J, Zhang M, Yang D, and Wang J: Edge-based scoring and searching method for identifying condition-responsive protein-protein interaction sub-network. *Bioinformatics* 2007; **23**: 2121.
36. Liu M, Liberzon A, Kong SW, Lai WR, Park PJ, Kohane IS, and Kasif S: Network-based analysis of affected biological processes in type 2 diabetes models. *PLoS Genet* 2007; **3**: e96.
37. Nibbe RK, Koyutürk M, and Chance MR: An integrative -omics approach to identify functional sub-networks in human colorectal cancer. *PLoS Comput Biol* 2010; **6**: e1000639.
38. Thomas S and Bonchev D: A survey of current software for network analysis in molecular biology. *Hum Genomics* 2010; **4**: 353.
39. Chen L, Arbieva ZH, Guo S, Marucha PT, Mustoe TA, and DiPietro LA: Positional differences in the wound transcriptome of skin and oral mucosa. *BMC Genomics* 2010; **11**: 471.
40. Hardman MJ and Ashcroft GS: Estrogen, not intrinsic aging, is the major regulator of delayed human wound healing in the elderly. *Genome Biol* 2008; **9**: R80.
41. Roy S, Khanna S, Rink C, Biswas S, and Sen CK: Characterization of the acute temporal changes in excisional murine cutaneous wound inflammation by screening of the wound-edge transcriptome. *Physiol Genomics* 2008; **34**: 162.
42. Eming SA, Koch M, Krieger A, Brachvogel B, Kreft S, Bruckner-Tuderman L, and Fox JW: Differential proteomic analysis distinguishes tissue repair biomarker signatures in wound exudates obtained from normal healing and chronic wounds. *J Proteome Res* 2010; **9**: 4758.
43. Angel P, Szabowski A, and Schorpp-Kistner M: Function and regulation of AP-1 subunits in skin physiology and pathology. *Oncogene* 2001; **20**: 2413.
44. Martin P and Parkhurst SM: Parallels between tissue repair and embryo morphogenesis. *Development* 2004; **131**: 3021.
45. Li G, Gustafson-Brown C, Hanks SK, Nason K, Arbeit JM, Pogliano K, and Johnson RS: c-Jun is essential for organization of the epidermal leading edge. *Dev Cell* 2003; **4**: 865.
46. Mercurio F and Manning AM: Multiple signals converging on NF-kappa B. *Curr Opin Cell Biol* 1999; **11**: 226.
47. Nguyen J, Gogusev J, Knapnougol P, and Bauvois B: Protein tyrosine kinase and p38 MAP kinase pathways are involved in stimulation of matrix metalloproteinase-9 by TNF-alpha in human monocytes. *Immun Lett* 2006; **106**: 34.
48. Gordon GM, Ledee DR, Feuer WJ, and Fini ME: Cytokines and signaling pathways regulating matrix metalloproteinase-9 (MMP-9) expression in corneal epithelial cells. *J Cell Physiol* 2009; **221**: 402.
49. Chase AJ, Bond M, Crook MF, and Newby AC: Role of nuclear factor-kappaB activation in metalloproteinase-1, -3, and -9 secretion by human macrophages *in vitro* and rabbit foam cells produced *in vivo*. *Arterioscler Thromb Vasc Biol* 2002; **22**: 765.
50. Haas AF, Wong JW, Iwahashi CK, Halliwell B, Cross CE, and Davis PA: Redox regulation of wound healing? NF-kappaB activation in cultured human keratinocytes upon wounding and the effect of low energy HeNe irradiation. *Free Radic Biol Med* 1998; **25**: 998.
51. Ubels JL, Edelhauser HF, and Austin KH: Healing of experimental corneal wounds treated with topically applied retinoids. *Am J Ophthalmol* 1983; **95**: 353.
52. Lee RC, Feinbaum RL, and Ambros V: The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 1993; **75**: 843.
53. Bonetta L: Protein-protein interactions: interactome under construction. *Nature* 2010; **468**: 851.

54. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, and Barabási AL: The human disease network. *Proc Natl Acad Sci USA* 2007; **104**: 8685.
55. The Cancer Genome Atlas: <http://tcga-data.nci.nih.gov>
56. Shankavaram UT, Varma S, Kane D, Sunshine M, Chary KK, Reinhold WC, Pommier Y, and Weinstein JN: CellMiner: a relational database and query tool for the NCI-60 cancer cell lines. *BMC Genomics* 2009; **10**: 277.
57. Yeang CH: An integrated analysis of molecular aberrations in NCI-60 cell lines. *BMC Bioinform* 2010; **11**: 495.
58. Ovaska K, Laakso M, Haapa-Paananen S, Louhimo R, Chen P, Aittomäki V, Valo E, Núñez-Fontarnau J, Rantanen V, Karinen S, Nousiainen K, Lahesmaa-Korpinen AM, Miettinen M, Saarinen L, Kohonen P, Wu J, Westermarck J, and Hautaniemi S: Large-scale data integration framework provides a comprehensive view on glioblastoma multiforme. *Genome Med* 2010; **2**: 65.
59. Shaw T and Martin P: Epigenetic reprogramming during wound healing: loss of polycomb-mediated silencing may enable upregulation of repair genes. *EMBO Rep* 2009; **10**: 881.