Protein Biomarker Druggability Profiling

Subramani Mani, MBBS PhD[a,*], **Daniel Cannon, MS**[a], **Robin Ohls, MD**[a], **Tudor Oprea, MD PhD**[a], **Stephen Mathias, PhD**[a], **Karri Ballard, PhD**[b], **Oleg Ursu, PhD**[a], and **Cristian Bologa, PhD**[a]

[a]University of New Mexico Health Sciences Center, Albuquerque, NM, 87131

[b]Myriad RBM, Austin, TX 78759

Abstract

Developing automated and interactive methods for building a model by incorporating mechanistic and potentially causal annotations of ranked biomarkers of a disease or clinical condition followed by a mapping into a contextual framework in disease-linked biochemical pathways can be used for potential drug-target evaluation and for proposing new drug targets. We demonstrate the potential of this approach using ranked protein biomarkers obtained in neonatal sepsis by enrolling 127 infants (39 infants with late onset neonatal sepsis and 88 control infants) and by performing a focused proteomic profile of the sera and by applying the interactive druggability profiling algorithm (DPA) developed by us.

Keywords

Protein biomarkers; pathway analysis; druggability profiling; mechanistic annotation; neonatal sepsis; machine learning

Introduction and background

Identification of proteomic biomarkers from blood for early detection of systemic infections or malignancies will facilitate prompt treatment leading to improved outcomes. Various biomarkers are useful for diagnosis, assessment of treatment response to drugs and for prognostic evaluation. However, there is limited work on mechanistic modeling of the biomarkers and identification of biomarkers that are likely to be druggable, that is, they could be subject to manipulation by small molecules. Simply identifying the protein biomarkers differentially expressed in diseased versus healthy populations may be appropriate for predicting outcomes, but this approach does not offer the ability to differentiate between biomarkers that drive disease manifestations (causal influences) and those that result from the disease process (effects).

During the last decade, there has been increasing interest in looking into the peripheral blood for protein biomarkers for early detection, diagnosis, risk stratification, treatment planning

*Corresponding author: Associate Professor, Division of Translational Informatics, Department of Medicine, MSC 08-4560, 1 University of New Mexico, Albuquerque, NM 87131-0001, manis@salud.unm.edu, Phone (work): 505-272-8933, Phone (cell): 615-618-7355.

or therapeutic response prediction for various infections, inflammatory conditions, degenerative diseases and cancers [1–16]. Compared to traditional tissue biopsies, drawing peripheral blood for biomarker assay is relatively non-invasive and the process can be easily repeated if needed for studying temporal patterns. Though there were some concerns early on about the process of biomarker discovery, validation and clinical utility [17], recent advances in proteomic biomarker assay development [18–24] have considerably improved the utility and value for examining protein biomarkers in blood. Since the proteome is downstream relative to the genome and transcriptome, assaying proteins holds considerable promise for discovering disease-linked biomarkers. It is clear that detection of multiparameter protein biomarkers from blood for prompt and early detection of systemic infections or cancers could lead to much more effective treatment of the condition with possible improvement in outcomes. Though researchers have recognized the potential of multiple biomarker measurements for rationalizing the discovery of suitable drug targets [7], there is limited work on mechanistic modeling of the biomarkers and in identifying the set of druggable biomarkers [25–27] (that is, compounds/drugs that can modulate the protein(s) and/or receptor(s)). Moreover, from a druggability perspective, large tracts of the exome have been left unexplored. Based on review of literature, Hopkins et al. report that 399 non-redundant molecular targets have been shown to bind efficaciously with small molecules [28] out of more than 10,000 likely such targets in the human genome using projections of ligand binding domains [29]. The challenge is to identify the relevant subset of potential druggable targets that are represented in disease-linked proteins.

A set of predictive, diagnostic or prognostic biomarkers relevant to a specific clinical condition or disease can be identified using a focused literature search or obtained from research studies designed specifically for biomarker discovery. Although biomarkers can play an effective role in early detection, diagnostic reasoning and assessment of prognosis, a mechanistic understanding is required for evaluating biomarkers from a druggability perspective. The resulting mechanistic knowledge will eventually move the promising ones to clinical trials and for therapeutic interventions.

We illustrate the proof of concept of our approach using the domain of late onset neonatal sepsis.

Methods

Biomarker discovery and ranking

We performed a focused proteomic assay of 90 potential biomarkers suspected to play a role in infection and/or inflammation using serum samples collected from 39 cases of late onset neonatal sepsis (culture positive) and 88 controls (culture negative) that we enrolled over a five year period from 2007 to 2012 (n=127). The Institutional Review Board at our institution approved the protocol, and written, informed parental consent was obtained. The potential biomarkers were selected based on literature and domain knowledge of experts Dr. Ballard and Dr. Ohls. The quantitative proteomic assay was performed by Myriad RBM using a customized implementation of the Luminex xMAP technology, a microsphere-based multiplexed immunoassay platform. Our first modeling objective was to develop a classification method capable of detecting late onset neonatal sepsis (LOS) on the day the

clinician first suspected sepsis. Toward this end, we first defined t_0 as the day on which the first positive blood culture was drawn, as this likely corresponds to be the day sepsis was first suspected. For each control, we selected a day t_0 taking into consideration the infant's age and sample availability. We then included only data from samples drawn on or before t_0 for computational predictive modeling. Ranking of predictive biomarkers was performed using a machine learning (ML) approach—the Random Forest variable importance method [30 31] to score and rank each variable. The methodological details of predictive modeling for biomarker discovery and ranking are described in detail in [32] which is under review. The methodological framework is summarized in Figure 1.

The methodological approach for biomarker discovery and ranking consists of predominantly two integrated modules—the predictive modeling and evaluation module (PMEM) and the biomarker ranking module (BRM). The PMEM proposes and validates a set of biomarkers for the condition of interest (neonatal sepsis). The BRM module incorporates three biomarker ranking approaches based on (1) variables representation in models, (2) feature set selection algorithms and (3) random forest variable importance. For this illustrative study we used the random forest variable importance method for ranking the neonatal sepsis biomarkers. Though we tried various thresholds in the model building stage to optimize sensitivity or specificity, the area under the ROC curve (AUC) remained the same and based on the best performance of RF we selected RF variable importance for variable ranking.

Note that though the 1st step of the DPA algorithm combines biomarker discovery and ranking, the operationalization of the step involves two procedures and they are separated out into two modules in Figure 1. For biomarker ranking to be effective and meaningful we need predictive models with high performance leading to the identification of useful predictive biomarkers. In the predictive modeling phase a representative set of machine learning algorithms are used to generate different types of predictive models and their performance evaluated. High performing models are then leveraged for the identification of predictive biomarkers which can be ranked by different methods using the biomarker ranking module.

Mechanistic annotation of ranked biomarkers

The top biomarkers selected using machine learning methods produce an incomplete mechanistic picture of their role in the disease process and its manifestations, which can be completed by including mechanistic information from pathway databases. We developed a workflow to extract, merge and analyze pathways from the "KEGG: Kyoto Encyclopedia of Genes and Genome" database [33 34] that are relevant to biomarkers of interest. First, all pathways implicated by any of the selected top biomarkers were identified; next, all these pathways were extracted from the KEGG pathways database and expanded.

Using (1) pathway biomarker representation and (2) the state-of-the-art drug database consisting of detailed information on approved and discontinued drugs worldwide (DrugCentral) [35 36] developed by our group, we identified biomarkers that can serve as potential therapeutic targets for further evaluation and drug development. We used the gene

set over-representation analysis tool from ConsensusPathDB [37] to obtain a ranked list of KEGG pathways in which the top biomarkers for neonatal sepsis were over-represented.

We implemented an automated tool to extract drug associations from the KEGG REST API[1]. For each pathway, we first retrieved the list of drugs linked directly to the pathway. Then, we retrieved the list of genes in the pathway and the list of drugs linked to each of those genes. We next retrieved a list of activities for each drug linked to the pathway or a gene in the pathway. The pathway diagrams were annotated manually using this extracted information.

We additionally obtained a list of drugs from DrugCentral that are potentially active on the top-ranked pathways we had identified. That is, for each pathway, we extracted a list of drugs that have known bioactivities with a gene in the pathway recorded in DrugCentral.

We also performed a target tissue localization (TTL) analysis of the top sepsis biomarkers using the Human Proteome Map (HPM) [38] to understand the tissue specificity of the biomarkers. The official gene names of the top biomarkers listed in Table 1 were used to query the HPM database and generate Figure 3. Note that DPA is an interactive algorithm and needs manual intervention during all step transitions.

We now introduce the DrugCentral before presenting the results. DrugCentral database aggregates information on approved and discontinued drugs worldwide, except for biological entities (antibodies, vaccines, etc.). More than 4400 active ingredients manually curated are stored currently in a relational database, with 3,943 small organic molecule entries; INN and USAN assigned names mapped to Active Pharmaceutical Ingredients (APhIs). Each APhI is linked to biological activity records collected from public (ChEMBL, IUPHAR[39], PDSP[40]), and commercial (WOMBAT-PK) databases. New FDA-approved APhIs are stored with biological activities data published in scientific literature. Currently, 22,760 biological activity records for human protein targets (1,886 unique), and 4,696 activity records for non-human targets (1,184 unique) are stored in DrugCentral. Drug targets related to known mechanisms of action (MoA) are currently under evaluation and to date 138 out of 838 human target-APhI pairs have been expert curated. Approved drug labels are collected and stored in text fields organized by LOINC (Logical Observation Identifiers Names and Codes, http://loinc.org/) section headings.

The steps of the Druggability Profiling Algorithm (DPA) are shown in Box 1.

Results

Fifteen biomarkers were identified that predicted neonatal sepsis infection with high accuracy. These biomarkers are shown in Table 1.

We identified five pathways extracted from the KEGG database that incorporate many of the top ranked sepsis biomarkers. See Table 2 for a listing of these pathways and the biomarkers represented in each.

---

[1]"REST-style KEGG API." 2012. 4 Dec. 2014 <http://www.kegg.jp/kegg/docs/keggapi.html>

Most of the top ranked biomarkers are cytokines and of the top selected pathways in sepsis the NOD-like receptor signaling pathway includes four biomarkers selected using the methods described in the Methods section. A subset of the protein targets present in this pathway are modulated by approved and investigational drugs as depicted in Figure 2. Only drugs with a known mechanism of action are shown in Figure 2.

Table 3 provides a summary of the top KEGG pathways of neonatal Sepsis showing the number of biomarkers represented in each of the pathways, the number of drugs in KEGG acting on these biomarkers and the number of drugs in DrugCentral acting on them.

The tissue specific expression data of the various sepsis biomarkers are shown in Figure 3. Figure 3 provides the relative expression of some of the top ranked sepsis biomarkers in 30 clinically defined healthy tissues (17 adult tissues, 6 primary hematopoietic cells and 7 fetal tissues). It also indicates the lack of expression for five of them in healthy tissues.

Discussion and Conclusion

Biomarker profiling for potential drug target identification using KEGG pathway database and the DrugCentral database has the potential to identify drug-target interactions related to the biomarkers of a specific disease or clinical condition. Even though KEGG incorporates drug-target pairs, KEGG doesn't provide numerical bioactivity values relating drugs and targets; in comparison DrugCentral has quantitative values associating drug-target pairs. Moreover, DrugCentral contains additional ~900 drug entries not present in KEGG [36]. The six drug-target interactions extracted from KEGG and the seven drug-target interactions with mechanism of action identified using DrugCentral shown in Figure 2 and the multiple drug-target interaction counts presented in Table 3 need further evaluation with respect to altering the pathology and clinical course of neonatal sepsis.

The results presented in Table 2 indicate that multiple signaling pathways are dysregulated in sepsis, suggesting that a successful treatment of this severe condition would require a polypharmacological approach using combinations of multiple drugs in order to restore the equilibrium in these multiple pathways. As seen in Table 3, some drugs act on targets included in these pathways, and a combination of them could be considered by expert clinicians as promising candidates for further evaluation in future experiments on animal models or clinical trials.

The tissue-specific expression data shown in Figure 3 can be used to prioritize drug targets in terms of tissue relevance for many clinical conditions. In addition to antibiotics which are critical in the management of neonatal sepsis druggability profiling of proteomic biomarkers provides additional avenues for modifying the course of neonatal sepsis based on host response. The intense proinflammatory host response with elevated chemokines and cytokines is an important factor in neonatal sepsis pathology and these immunological reactions can adversely affect distant organs including the brain [41]. Based on the clinical presentation and organ involvement TTL information from Step 5 of the DPA (Figure 3) can be used to select the most appropriate drug from Step 4 of DPA (Figure 2). This study opens up the possibility of biomarker druggability profiling using informatics methods and

databases to propose new drug targets and for a systematic identification of drug-target interactions.

For purposes of this study we adopted the biomarker ranking method based on RF variable importance. However, we also developed and implemented two other methods for biomarker ranking based on variable representation in tree models and feature sets.

We make two significant contributions—(1) we provide an interactive algorithm that takes as input a set of biomarkers and outputs a ranked list of pathways that incorporate the biomarkers while also providing a list of drugs with a mechanism of action that act on potential targets of the pathway and (2) we illustrate the usefulness of the algorithm in the domain of neonatal sepsis using top-ranked protein biomarkers.

Acknowledgments

References

1. Addona TA, Shi X, Keshishian H, et al. A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. Nature biotechnology. 2011; 29(7):635–43.

2. Aebersold R, Anderson L, Caprioli R, Druker B, Hartwell L, Smith R. Perspective: a program to improve protein biomarker discovery for cancer. Journal of proteome research. 2005; 4(4):1104–09. [PubMed: 16083259]

3. Bogdanov M, Matson WR, Wang L, et al. Metabolomic profiling to develop blood biomarkers for Parkinson's disease. Brain. 2008; 131(2):389–96. [PubMed: 18222993]

4. Doecke JD, Laws SM, Faux NG, et al. Blood-based protein biomarkers for diagnosis of Alzheimer disease. Archives of neurology. 2012; 69(10):1318–25. [PubMed: 22801742]

5. Domenici E, Willé DR, Tozzi F, et al. Plasma protein biomarkers for depression and schizophrenia by multi analyte profiling of case-control collections. PLoS One. 2010; 5(2):e9166. [PubMed: 20161799]

6. Hewitt SM, Dear J, Star RA. Discovery of protein biomarkers for renal diseases. Journal of the American Society of Nephrology. 2004; 15(7):1677–89. [PubMed: 15213255]

7. Hood L, Heath JR, Phelps ME, Lin B. Systems biology and new technologies enable predictive and preventative medicine. Science. 2004; 306(5696):640–43. [PubMed: 15499008]

8. Jacobs JM, Adkins JN, Qian W-J, et al. Utilizing human blood plasma for proteomic biomarker discovery. Journal of proteome research. 2005; 4(4):1073–85. [PubMed: 16083256]

9. Kövesdi E, Lückl J, Bukovics P, et al. Update on protein biomarkers in traumatic brain injury with emphasis on clinical use in adults and pediatrics. Acta neurochirurgica. 2010; 152(1):1–17. [PubMed: 19652904]

10. Liao H, Wu J, Kuhn E, et al. Use of mass spectrometry to identify protein biomarkers of disease severity in the synovial fluid and serum of patients with rheumatoid arthritis. Arthritis & Rheumatism. 2004; 50(12):3792–803. [PubMed: 15593230]

11. Liotta LA, Petricoin EF. Serum peptidome for cancer detection: spinning biologic trash into diagnostic gold. Journal of Clinical Investigation. 2006; 116(1):26–30. [PubMed: 16395400]

12. Paczesny S, Krijanovski OI, Braun TM, et al. A biomarker panel for acute graft-versus-host disease. Blood. 2009; 113(2):273–78. [PubMed: 18832652]

13. Petricoin EF, Belluco C, Araujo RP, Liotta LA. The blood peptidome: a higher dimension of information content for cancer biomarker discovery. Nature Reviews Cancer. 2006; 6(12):961–67. [PubMed: 17093504]

14. Rosas IO, Richards TJ, Konishi K, et al. MMP1 and MMP7 as potential peripheral blood biomarkers in idiopathic pulmonary fibrosis. PLoS medicine. 2008; 5(4):e93. [PubMed: 18447576]

15. Tsimikas S, Willerson JT, Ridker PM. C-reactive protein and other emerging blood biomarkers to optimize risk stratification of vulnerable patients. Journal of the American College of Cardiology. 2006; 47(8s1):C19–C31. [PubMed: 16631506]

16. Welsh JB, Sapinoso LM, Kern SG, et al. Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. Proceedings of the National Academy of Sciences. 2003; 100(6):3410–15.

17. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nature biotechnology. 2006; 24(8):971–83.

18. Adkins JN, Varnum SM, Auberry KJ, et al. Toward a Human Blood Serum Proteome analysis by multidimensional separation coupled with mass spectrometry. Molecular & Cellular Proteomics. 2002; 1(12):947–55. [PubMed: 12543931]

19. Fan R, Vermesh O, Srivastava A, et al. Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. Nature biotechnology. 2008; 26(12):1373–78.

20. Piliarik M, Bocková M, Homola J. Surface plasmon resonance biosensor for parallelized detection of protein biomarkers in diluted blood plasma. Biosensors and Bioelectronics. 2010; 26(4):1656–61. [PubMed: 20864329]

21. Seibert V, Ebert MP, Buschmann T. Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. Briefings in functional genomics & proteomics. 2005; 4(1):16–26. [PubMed: 15975261]

22. Stern E, Vacic A, Rajan NK, et al. Label-free biomarker detection from whole blood. Nature nanotechnology. 2010; 5(2):138–42.

23. Whiteaker JR, Zhao L, Anderson L, Paulovich AG. An automated and multiplexed method for high throughput peptide immunoaffinity enrichment and multiple reaction monitoring mass spectrometry-based quantification of protein biomarkers. Molecular & Cellular Proteomics. 2010; 9(1):184–96. [PubMed: 19843560]

24. Zhang H, Liu AY, Loriaux P, et al. Mass spectrometric detection of tissue proteins in plasma. Molecular & Cellular Proteomics. 2007; 6(1):64–71. [PubMed: 17030953]

25. Danhof M, Alvan G, Dahl SG, Kuhlmann J, Paintaud G. Mechanism-based pharmacokinetic–pharmacodynamic modeling—a new classification of biomarkers. Pharmaceutical research. 2005; 22(9):1432–37. [PubMed: 16132354]

26. Jack CR Jr, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. The Lancet Neurology. 2013; 12(2):207–16. [PubMed: 23332364]

27. Mayr M, Zhang J, Greene AS, Gutterman D, Perloff J, Ping P. Proteomics-based Development of Biomarkers in Cardiovascular Disease Mechanistic, Clinical, and Therapeutic Insights. Molecular & Cellular Proteomics. 2006; 5(10):1853–64. [PubMed: 16733263]

28. Hopkins AL, Groom CR. The druggable genome. Nature reviews Drug discovery. 2002; 1(9):727–30. [PubMed: 12209152]

29. Bailey D, Zanders E, Dean P. The end of the beginning for genomic medicine. Nature biotechnology. 2001; 19(3):207–08.

30. Breiman L. Random forests. Machine Learning. 2001; 45(1):5–32.

31. Genuer R, Poggi J-M, Tuleau-Malot C. Variable selection using random forests. Pattern Recognition Letters. 2010; 31(14):2225–36.

32. Mani S, Cannon DC, Hartenberger C, et al. Focused Proteomic Profiling for Late-Onset Neonatal Sepsis. 2016 (under review).

33. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic acids research. 2014; 42(D1):D199–D205. [PubMed: 24214961]

34. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic acids research. 2012; 40(D1):D109–D14. [PubMed: 22080510]

35. Oprea TI, Nielsen SK, Ursu O, et al. Associating Drugs, Targets and Clinical Outcomes into an Integrated Network Affords a New Platform for Computer-Aided Drug Repurposing. Molecular informatics. 2011; 30(2–3):100–11. [PubMed: 22287994]

36. Ursu O, Holmes J, Knockel J, et al. DrugCentral: online drug compendium. Nucleic Acids Research. 2016 gkw993.

37. Kamburov A, Cavill R, Ebbels TM, Herwig R, Keun HC. Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. Bioinformatics. 2011; 27(20):2917–18. [PubMed: 21893519]

38. Kim M-S, Pinto SM, Getnet D, et al. A draft map of the human proteome. Nature. 2014; 509(7502):575–81. [PubMed: 24870542]

39. Sharman JL, Benson HE, Pawson AJ, et al. IUPHAR-DB: updated database content and new features. Nucleic acids research. 2013; 41(D1):D1083–D88. [PubMed: 23087376]

40. Roth BL, Lopez E, Beischel S, Westkaemper RB, Evans JM. Screening the receptorome to discover the molecular targets for plant-derived psychoactive compounds: a novel approach for CNS drug discovery. Pharmacology & therapeutics. 2004; 102(2):99–110. [PubMed: 15163592]

41. Ng PC, Ma TPY, Lam HS. The use of laboratory biomarkers for surveillance, diagnosis and prediction of clinical outcomes in neonatal sepsis and necrotising enterocolitis. Archives of Disease in Childhood-Fetal and Neonatal Edition. 2015 fetalneonatal-2014-307656.

<div style="border: 1px solid; background-color: #d6e4f0; padding: 1em;">

Box 1: Druggability Profiling Algorithm (DPA)

1. Protein biomarker discovery and ranking (Figure 1 and Table 1)

2. Identify pathways containing ranked biomarkers from KEGG

3. Analyze and rank based on number of biomarkers represented (Table 2)

4. Identify drugs from KEGG and DrugCentral acting on biomarkers and other proteins/genes in the pathways (Figure 2 and Table 3)

5. Perform target tissue localization (TTL) analysis of the top biomarkers using the Human Proteome Map

</div>
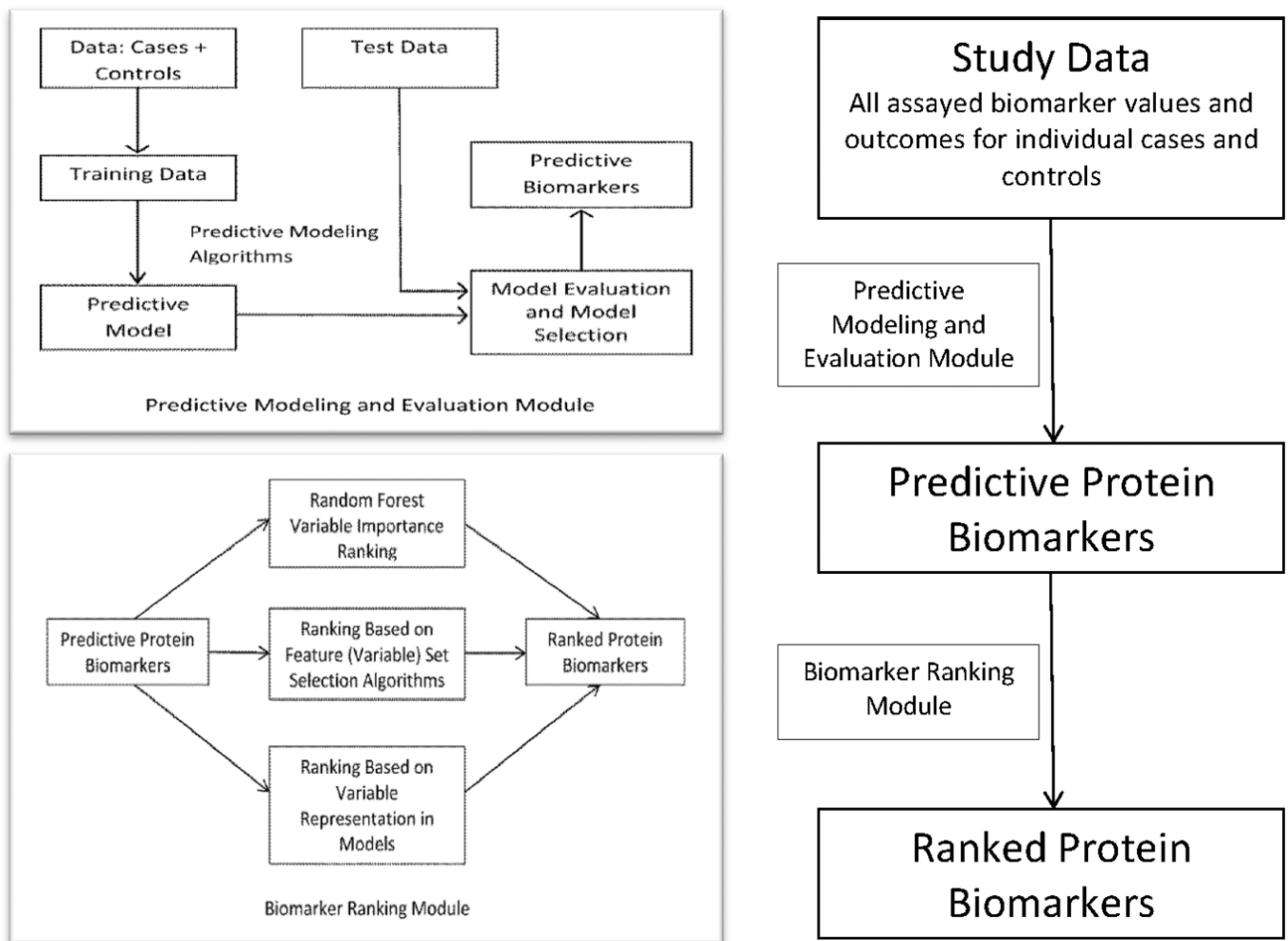
An integrated machine learning (ML) framework for biomarker discovery and ranking. To the left are shown two modules—(1) Predictive Modeling and Evaluation Module (2) Biomarker Ranking Module. These two modules are incorporated in the integrated ML framework provided on the right for obtaining a set of ranked protein biomarkers from study data composed of cases and control samples.
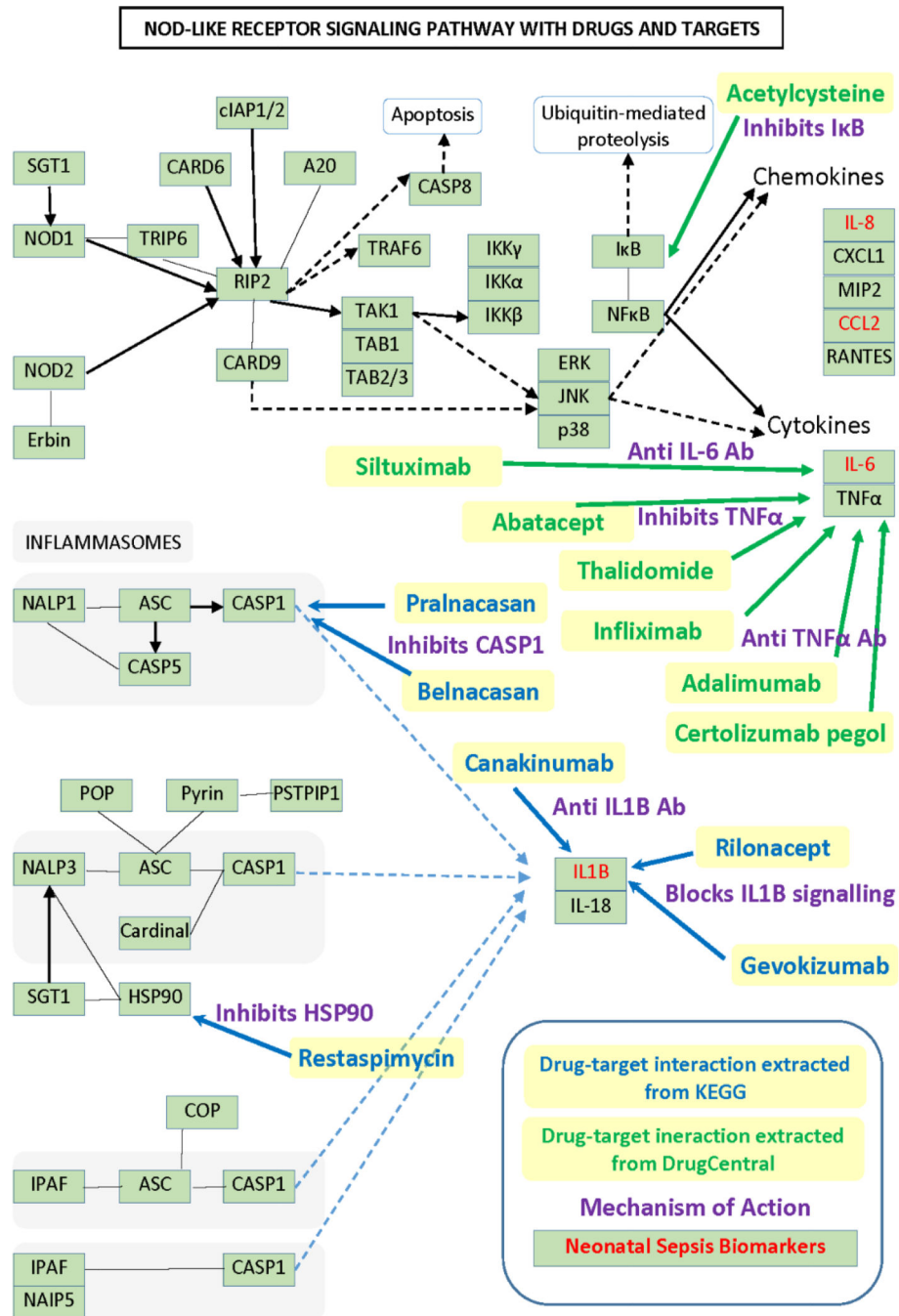
Figure 1.
ML Framework

Figure 2.
The NOD-like receptor signaling pathway from KEGG shown with drugs, targets and known mechanism of action (→ denotes activation, ---> and ---> denote indirect effect and — denotes binding/association).

Figure 3.
Expression across samples of selected sepsis proteomic biomarkers. Only six fetal tissues are represented. White squares denote absence of any expression, faded red denotes some level of expression and bright red denotes high levels of expression in the specific tissue for the particular protein biomarker providing a relative measure of tissue target localization.

Table 1

Top Fifteen Predictive Biomarkers of Late Onset Neonatal Sepsis Based on RF Variable Importance. Gene Names in [] Map to Table 2 and Figures 2 and 3.

Monocyte Chemotactic Protein-1 (MCP-1) [CCL2]
Tumor necrosis factor receptor-2 (TNFR2) [TNFRSF1B]
Interleukin-6 (IL-6) [IL6]
Interleukin-1 receptor antagonist (IL-1ra) [IL1RN]
Prostatic Acid Phosphatase (PAP) [ACPP]
Macrophage Inflammatory Protein-1 beta (MIP-1 beta) [CCL4]
Granulocyte Colony-Stimulating Factor (G-CSF) [CSF3]
Calcitonin [CALCA]
C-Reactive Protein (CRP) [CRP]
Interleukin-8 (IL-8) [CXCL8]
Interleukin-10 (IL-10) [IL10]
Interleukin-1 beta (IL-1 beta) [IL1B]
Myeloperoxidase (MPO) [MPO]
Intercellular Adhesion Molecule 1 (ICAM-1) [ICAM1]
Tumor Necrosis Factor beta (TNF-beta) [LTA]

Table 2

Top ranked KEGG pathways containing selected biomarkers from Table 1

Disease	Pathway	Biomarkers (Gene Names)
Sepsis	Cytokine-cytokine receptor interaction - human	CCL2, TNFRSF1B, CCL4, IL10, CSF3, IL6, LTA, IL1B
	TNF signaling pathway - human	CCL2, IL6, LTA, TNFRSF1B, IL1B
	Chagas disease (American trypanosomiasis) - human	IL10, CCL2, IL1B, IL6
	NOD-like receptor signaling pathway - human	CCL2, IL6, IL1B, IL8
	Cytosolic DNA-sensing pathway - human	IL6, IL1B, CCL4

Table 3

Top ranked KEGG pathways with biomarker and drug counts from KEGG and DrugCentral

Disease	Pathway	Biomarker Count	KEGG Drug Count	DrugCentral Drug Count
Sepsis	Cytokine-cytokine receptor interaction - human	9	17	12
	TNF signaling pathway - human	5	10	7
	Chagas disease (American trypanosomiasis) -human	5	8	12
	NOD-like receptor signaling pathway - human	4	7	12
	Cytosolic DNA-sensing pathway - human	3	6	3