

BluePen Biomarkers LLC: integrated biomarker solutions

BluePen Biomarkers provides a unique comprehensive multi-omics biomarker discovery and validation platform. We can quantify, integrate and analyze genomics, proteomics, metabolomics and lipidomics biomarkers, alongside clinical data, demographics and other phenotypic data. A unique bio-inspired signal processing analytic approach is used that has the proven ability to identify biomarkers in a wide variety of diseases. The resulting biomarkers can be used for diagnosis, prognosis, mechanistic studies and predicting treatment response, in contexts from core research through clinical trials. BluePen Biomarkers provides an additional groundbreaking research goal: identifying surrogate biomarkers from different modalities. This not only provides new biological insights, but enables least invasive, least-cost tests that meet or exceed the predictive quality of current tests.

First draft submitted: 4 May 2016; Accepted for publication: 6 May 2016; Published online: 9 June 2016

Keywords: bioinformatics • biomarkers • lipidomics • metabolomics • proteomics • systems biology

BluePen Biomarkers was established in April 2016 in Philadelphia PA through a collaboration between the University of Pennsylvania's Center for Innovation, Blueprint Bio and Emerald Logic from Aliso Viejo (www.bluepenbio.com). The company was founded by experts from the field of biomarker discovery and validation at the University of Pennsylvania, which complements the bioinformatics expertise provided by Emerald Logic from Aliso Viejo (CA, USA). A major goal of the company is to address the problem of poorly predictive biomarkers that was identified by George Poste in his provocative 2011 article in *Nature* [1]. A coordinated systems biology approach has been developed to replace the patchwork of different technologies that are currently used to identify biomarkers [1]. Our approach involves the use of integrated multi-omics technologies as described below:

- BluePen Biomarker's team of management and scientific advisors comprises leading biomedical scientists who provide a coordinated approach to fully integrated biomarker solutions;
- Genomics, proteomics, metabolomics and lipidomics research is conducted in BluePen Biomarkers' laboratories using state-of-the-art instrumentation;
- The products of multi-omics research are integrated using sophisticated bioinformatics based on combinatorial signal processing and analytic approaches developed by Emerald Logic to provide a fully integrated biomarker solution (Figure 1);
- The fundamental conviction of the company's founders is that any intervention that has a biologic outcome will have a detectable predictive biomarker signature. The challenge is to detect that signature;

Ian A Blair^{*1,2,3}, Clementina Mesaros^{2,3}, Patrick Lilley^{4,5} & Matthew Nunez^{1,5}

¹BluePen Biomarkers LLC, 3401 Grays Ferry Avenue, Building 212, Suite 132, Philadelphia, PA 19146-2701, USA

²Penn SRP Center, University of Pennsylvania, Philadelphia, PA 19104-6160, USA

³Department of Systems Pharmacology & Translational Therapeutics, University of Pennsylvania, Philadelphia, PA 19104-6160, USA

⁴Emerald Logic, Aliso Viejo, CA 92656, USA

⁵Blueprint Bio, Aliso Viejo, CA 92656, USA

*Author for correspondence:

Tel.: +1 215 573 9885

Fax: +1 215 573 9889

ianblair@mail.med.upenn.edu

- The comprehensive multi-omics systems biology platform available at BluePen Biomarkers coupled with Emerald Logic's innovative signal processing approach provides a powerful solution to the biomarker needs of academia as well as pharmaceutical and biotech industries (Figure 1);
- The substantial expertise available within the company provides a unique resource for assessing potential clinical biomarker needs;
- Biomarker discovery and validation can be tailored to the particular project under consideration;
- BluePen Biomarkers now provides a commercial multi-omics platform that can quickly yield highly predictive biomarkers.

Discussion

Whole blood is normally provided to BluePen Biomarkers for genomics analysis; whereas serum is normally supplied for proteomics, metabolomics and lipidomics analyses. However, buffy coat or peripheral blood monocytes (PBMCs) are sometimes provided for proteomics analyses and platelets are sometimes provided for metabolism studies or specific mitochondrial protein quantification [2–4]. BluePen Biomarkers major services are summarized below:

- Genomics analyses generally involve RNA sequencing (RNA-Seq), which is sometimes known as whole transcriptome shotgun sequencing. RNA-Seq, which is conducted by next-generation sequencing (NGS) using an Illumina HiSeq 3000 instrument, has largely replaced microarray analysis for the quantification of mRNA transcripts in a biofluid sample at a given moment in time. This can provide biomarker signatures of particular disease states. RNA-Seq can also be used to assess alternative gene-spliced transcripts, post-transcriptional modifications, gene fusion, mutations/single-nucleotide polymorphisms, different populations of RNA, total RNA, micro-RNA and transfer-RNA;
- Proteomics analyses normally involve targeted quantification of the major proteins that are identified as being upregulated by the RNA-Seq data. Analyses are conducted using nanospray ionization coupled with two-dimensional ultraperformance liquid chromatography/high-resolution mass spectrometry (UPLC-HRMS/MS) coupled with parallel reaction monitoring (PRM). These studies are conducted on a Thermo Q Exactive HF Hybrid Quadrupole-Orbitrap mass spectrometer operating in a targeted mode. Validation of differentially regulated proteins is conducted using rigorous sta-

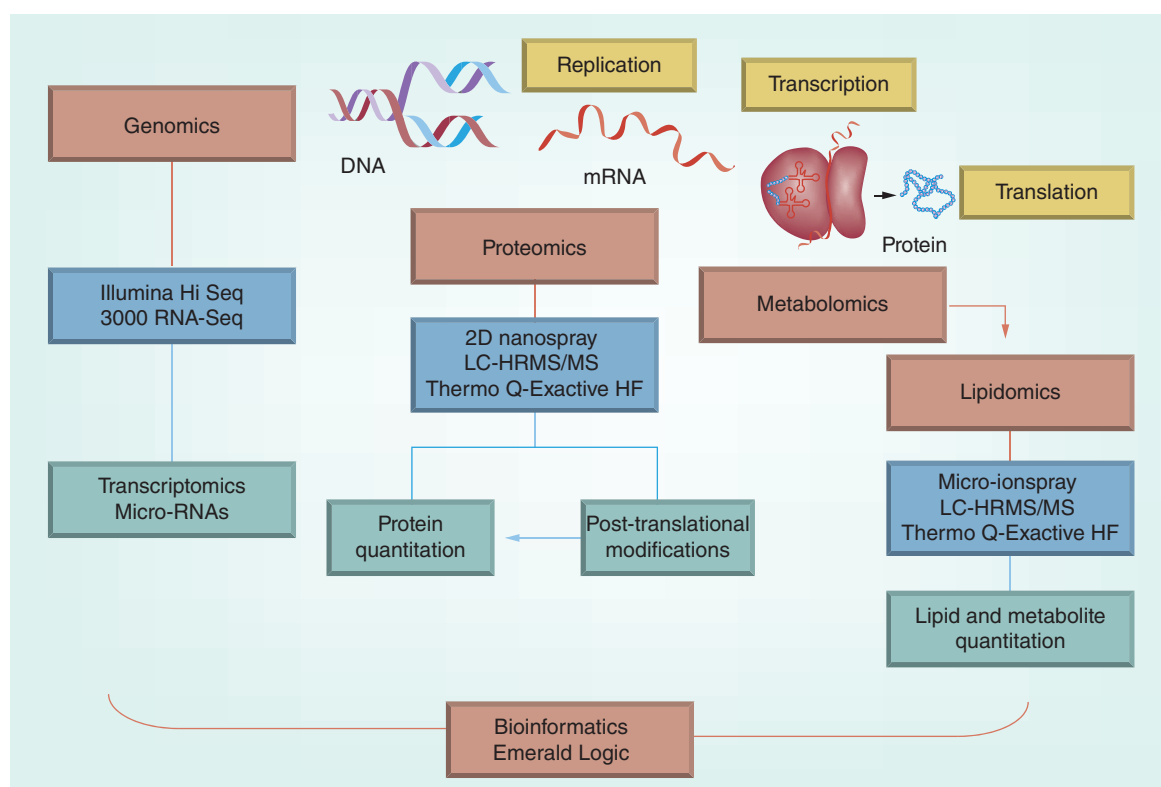


Figure 1. BluePen Biomarkers systems biology approach to biomarker discovery and validation.

ble isotope dilution UPLC-HRMS. Stable isotope-labeled protein standards are generated using stable isotope labeling by amino acids in cell culture (SILAC) based methodology [5];

- Metabolomics and lipidomics analyses are conducted on serum samples that have been extracted using Bligh and Dyer methodology [6,7]. For biomarker discovery, the aqueous phase is subjected to metabolomics analysis and the organic phase is subjected to lipidomics analysis. Both analyses are conducted using heated electrospray ionization (HESI) coupled with UPLC-HRMS/MS on a Thermo Q Exactive HF Hybrid Quadrupole-Orbitrap mass spectrometer. Metabolites and lipids that are differentially expressed between cases and controls are thoroughly characterized by comparison of UPLC-HRMS/MS characteristics with authentic samples either obtained commercially or synthesized in house;
- Stable isotope-labeled metabolite and lipid standards are generated using stable isotope labeling by essential nutrients in cell culture (SILEC) methodology [8–10]. There are potentially >100,000 lipids present in serum, although Lipid Maps has only identified some 48,360 lipids (www.lipidmaps.org). This means that BluePen Biomarkers can potentially discover lipid biomarkers where no standards are available to confirm structures. In this case, total synthesis is conducted using the extensive synthetic expertise that is available within the company so that absolute structural confirmation can be conducted before embarking on extensive clinical studies;
- Bioinformatics analyses are conducted by Emerald Logic using a proprietary analytic platform that they have developed. This will identify signals in mRNA, protein, metabolite and lipid expression between cases and controls. An optimal panel of biomarkers will then emerge with maximal clinical specificity and sensitivity for a particular indication.

Developing clinically useful biomarkers in biofluids is a complex process that often results in failure [1]. Therefore, it is important that the biomarker discovery studies are conducted in appropriate biofluids and that they are followed by rigorous bioanalytical validation [11]. This involves establishing limits of quantification and estimation of the errors that are introduced by sample handling analogous to methods that are used in the validation of drug assays [12]. The bioanalytical validation is an important first step that is conducted by BluePen Biomarkers before clinical validation is con-

ducted using precious patient samples. As a result, time-consuming and expensive clinical validation studies do not fail as a result of bioanalytical issues and so this rigorous approach ends up being highly cost-effective. Biomarker validation studies require quality control samples to be prepared in the relevant biofluid that is being used. For endogenous biomarkers this can be challenging because it is likely to already be present in the biofluid. In addition, sophisticated methods are required to determine whether an endogenous signal is actually the analyte of interest but not interfering substances with similar physicochemical properties. To best distinguish the target analyte from chemical background peaks, assays have to be conducted with the maximal specificity and sensitivity that is possible using stable isotope analogs as internal standards. Gas chromatography (GC)- and LC-MS are the two most widely used instrument platforms to employ stable isotope dilution methodology. LC-MS is more applicable to the analysis of a wider range of biomarkers than GC-MS and is also inherently easier to use for rigorous validation. Consequently, BluePen Biomarkers uses bioanalytical platforms for proteomics, metabolomics and lipidomics analyses that are all based on LC-MS methodology.

Quantitative biomarker studies generally require the most sensitive means of detection possible [7,11]. The new generation of high-resolution mass spectrometers based on the Thermo Q-Exactive platform can provide data using either the HR selected ion monitoring (SIM) mode or the HRPRM mode. BluePen Biomarkers uses both of these techniques on their Thermo Q-Exactive HF instruments. The specificity of this approach to biomarker analysis arises from the ability to monitor masses with an accuracy of 0.001 Da and the sensitivity derives from the unique capabilities of the Orbitrap that is installed in the HF instrument. A stable (heavy) isotope labeled internal standard is used to establish the presence of an endogenous analyte using both the LC retention time and accurate mass selection of the Q-Exactive HF mass spectrometer. This level of specificity cannot be attained with any other bioanalytical technique employed for biomarker analysis.

An authentic stable isotope-labeled analog of a biomarker is identical to the endogenous molecule, except for its mass [11]. The response ratio between the unlabeled biomarker and its labeled internal standard is then interpolated into a standard curve in order to calculate the absolute amount of a biomarker analyte in the unknown sample. Variations of this technique are also used (most extensively in proteomics) where metabolically labeled proteins can be used at unknown concentrations for relative quantification purposes [13–15]. If necessary, absolute protein quantification can also be performed using appropriate protein standards [16,17]. In all cases, stable

isotope internal standards offer the means to verify the presence of the analyte and normalize experimental variables such as sample storage and matrix suppression. Differential ionization cannot occur between an identified biomarker and its stable isotope analog internal standard in the source of the mass spectrometer. This minimizes suppression effects even when chromatographic retention times vary with biofluid samples from different individuals. It is therefore impossible to standardize the amount of suppression occurring within any particular sample. Stable isotope analog internal standards also compensate for the selective binding to active sites on glassware or other surfaces that can occur during extraction and chromatography, preventing significant losses of the biomarker being analyzed. Finally, in some cases, compound enrichment is required to improve the analytical performance, as we showed for amyloid β -peptides [18] and DNA-adducts [19,20], which both required immunoaffinity purification. In these cases, stable isotope analog internal standards were essential.

The search for small molecule biomarkers that are predictive of cardiovascular disease, neurodegenerative diseases and cancers has been particularly intense and frustrating [1]. However, BluePen Biomarkers strongly believes that by linking biomarker discovery and validation with the sophisticated bioinformatics provided by Emerald Logic that it will be possible to rationally approach the development of biomarkers for specific diseases.

Highlights

- BluePen Biomarkers was established in April 2016 to provide a coordinated approach to fully integrated biomarker solutions;
- The products of multi-omics research are integrated using sophisticated bioinformatics based on combinatorial signal processing and analytic approaches developed by Emerald Logic to provide a fully integrated biomarker solution;
- BluePen Biomarkers believes that by linking biomarker discovery and validation with the sophis-

ticated bioinformatics provided by Emerald Logic that it will be possible to rationally approach the development of biomarkers for specific diseases;

- RNA-Seq is used to assess the potential utility of transcript profiling for biomarker analysis and upregulation of relevant proteins is confirmed by targeted proteomics;
- BluePen Biomarkers uses methods based on UPLC-HRMS for proteomics, metabolomics and lipidomics methods on their Thermo Q-Exactive HF instruments;
- This provides the only commercial multi-omics platform currently available that can quickly yield highly predictive biomarkers;
- The fundamental conviction of the company's founders is that any intervention that has a biologic outcome will have a detectable predictive biomarker signature.

Disclaimer

The opinions expressed in this article are those of the authors and do not necessarily reflect the views of Future Science, Ltd.

Financial & competing interests disclosure

IA Blair is a Co-founder and Board member of BluePen Biomarkers. P Lilley is a Board Member of BluePen Biomarkers and CEO of Emerald Logic. M Nunez is a Co-founder, Chairman and CEO of BluePen Biomarkers, as well as Founder, Chairman and Board member of Blueprint Bio. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Creative Commons Attribution 4.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

References

- 1 Poste G. Bring on the biomarkers. *Nature* 469(7329), 156–157 (2011).
- 2 Basu SS, Deutsch EC, Schmaier AA, Lynch DR, Blair IA. Human platelets as a platform to monitor metabolic biomarkers using stable isotopes and LC-MS. *Bioanalysis* 5(24), 3009–3021 (2013).
- 3 Worth AJ, Basu SS, Deutsch EC *et al.* Stable isotopes and LC-MS for monitoring metabolic disturbances in Friedreich's ataxia platelets. *Bioanalysis* 7(15), 1843–1855 (2015).
- 4 Worth AJ, Marchione DM, Parry RC *et al.* LC-MS analysis of human platelets as a platform for studying mitochondrial metabolism. *J. Vis. Exp.* 103(6), 612–617 (2016).
- 5 Yan Y, Weaver VM, Blair IA. Analysis of protein expression during oxidative stress in breast epithelial cells using a stable isotope labeled proteome internal standard. *J. Proteome. Res.* 4(6), 2007–2014 (2005).

- 6 Snyder NW, Khezam M, Mesaros CA, Worth A, Blair IA. Untargeted metabolomics from biological sources using ultraperformance liquid chromatography-high resolution mass spectrometry (UPLC-HRMS). *J. Vis. Exp.* 75, e50433 (2013).
- 7 Snyder NW, Mesaros C, Blair IA. Translational metabolomics in cancer research. *Biomark. Med.* 9(9), 821–834 (2015).
- 8 Basu SS, Mesaros C, Gelhaus SL, Blair IA. Stable isotope labeling by essential nutrients in cell culture for preparation of labeled coenzyme A and its thioesters. *Anal. Chem.* 83(4), 1363–1369 (2011).
- 9 Basu SS, Blair IA. SILEC: a protocol for generating and using isotopically labeled coenzyme A mass spectrometry standards. *Nat. Protoc.* 7(1), 1–12 (2012).
- 10 Snyder NW, Tomblin G, Worth AJ *et al.* Production of stable isotope-labeled acyl-coenzyme A thioesters by yeast stable isotope labeling by essential nutrients in cell culture. *Anal. Biochem.* 474, 59–65 (2015).
- 11 Ciccimaro E, Blair IA. Stable-isotope dilution LC-MS for quantitative biomarker analysis. *Bioanalysis* 2(2), 311–341 (2010).
- 12 Bansal S, DeStefano A. Key elements of bioanalytical method validation for small molecules. *AAPS J.* 9(1), E109–E114 (2007).
- 13 Wehr AY, Hwang WT, Blair IA, Yu KH. Relative quantification of serum proteins from pancreatic ductal adenocarcinoma patients by stable isotope dilution liquid chromatography-mass spectrometry. *J. Proteome Res.* 11(3), 1749–1758 (2012).
- 14 Wehr AY, Furth EE, Sangar V, Blair IA, Yu KH. Analysis of the human pancreatic stellate cell secreted proteome. *Pancreas* 40(4), 557–566 (2011).
- 15 Shah SJ, Yu KH, Sangar V, Parry SI, Blair IA. Identification and quantification of preterm birth biomarkers in human cervicovaginal fluid by liquid chromatography/tandem mass spectrometry. *J. Proteome. Res.* 8(5), 2407–2417 (2009).
- 16 Ciccimaro E, Hanks SK, Yu KH, Blair IA. Absolute quantification of phosphorylation on the kinase activation loop of cellular focal adhesion kinase by stable isotope dilution liquid chromatography/mass spectrometry. *Anal. Chem.* 81(9), 3304–3313 (2009).
- 17 Wang Q, Zhang S, Guo L *et al.* Serum apolipoprotein A-1 quantification by LC-MS with a SILAC internal standard reveals reduced levels in smokers. *Bioanalysis* 7(22), 2895–2911 (2015).
- 18 Oe T, Ackermann BL, Inoue K *et al.* Quantitative analysis of amyloid beta peptides in cerebrospinal fluid of Alzheimer's disease patients by immunoaffinity purification and stable isotope dilution liquid chromatography/negative electrospray ionization tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* 20(24), 3723–3735 (2006).
- 19 Mangal D, Vudathala D, Park JH, Lee SH, Penning TM, Blair IA. Analysis of 7,8-dihydro-8-oxo-2'-deoxyguanosine in cellular DNA during oxidative stress. *Chem. Res. Toxicol.* 22(5), 788–797 (2009).
- 20 Mesaros C, Arora JS, Wholer A, Vachani A, Blair IA. 8-Oxo-2'-deoxyguanosine as a biomarker of tobacco-smoking-induced oxidative stress. *Free Radic. Biol. Med.* 53(3), 610–617 (2012).