ENZENE

Passion. Innovation. Life.

Multi-omics Approach to Biomarker Discovery

Abhishek Mathur, PhD, MBA
Senior Vice President,
Head of Research and Development

Mandar Bhonde, PhD
Senior General Manager,
Head of Discovery Operations and Bioanalysis

Introduction

A biomarker, also referred to as a molecular biosensor, is typically a protein, metabolite or any cellular component that informs researchers and clinicians about the nature, status and severity of a disease. Biomarkers can be classified (Figure 1) into (i) risk stratification, (ii) prognostic, (iii) diagnostic, (iv) predictive or (v) disease/therapy monitoring (1).

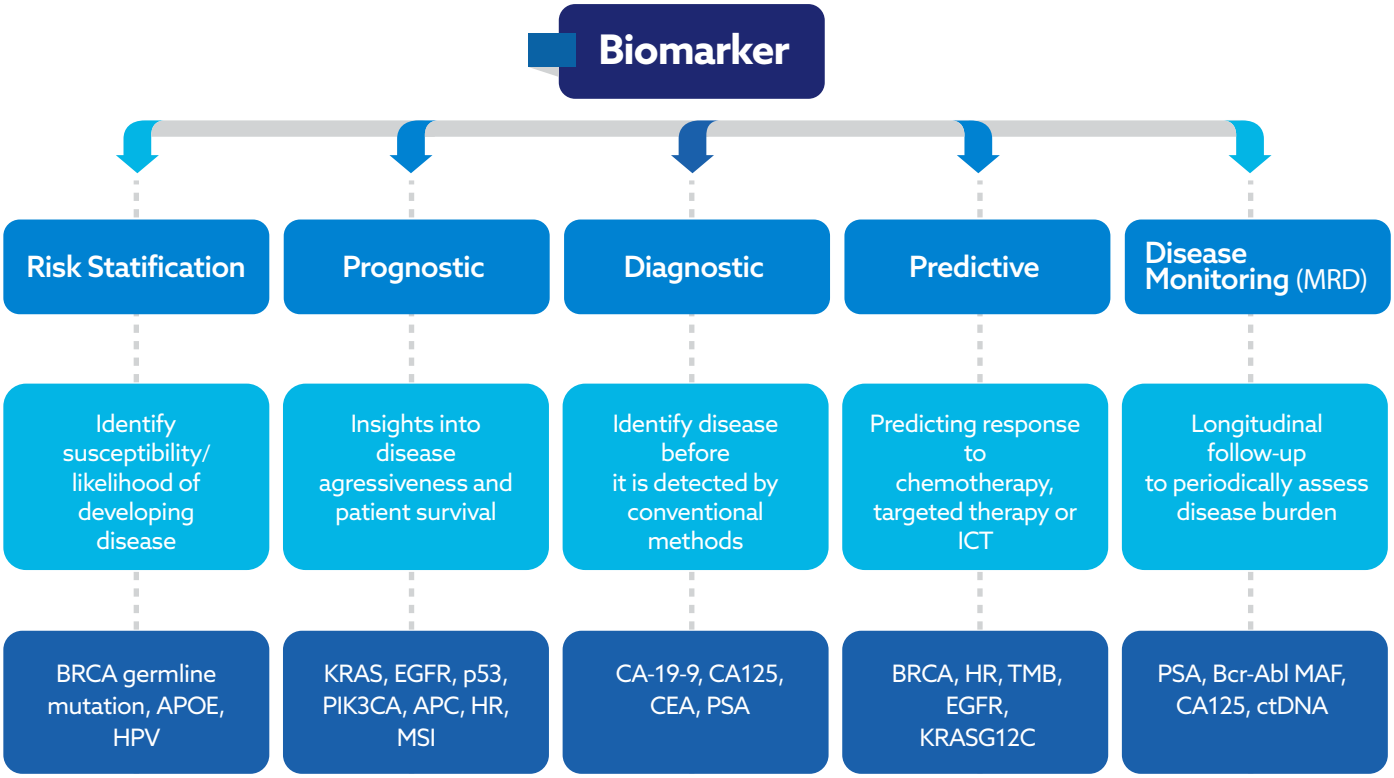

Figure 1: Biomarker classification

Biomarkers can be broadly categorized based on five main aspects related to (i) risk of developing a disease, (ii) prognostication, (iii) screening & diagnosis, (iv) predicting response/resistance to therapy, and (v) minimal residual disease (MRD) monitoring. Each of these categories is further exemplified with certain validated biomarkers such as presence of a BRCA germline mutation and the susceptibility to breast cancer. HR: Homologous Recombination; MSI: Microsatellite Instability; TMB: Tumor Mutation Burden; PSA: Prostate Specific Antigen; MAF: Mutant Allele Frequency; ctDNA: circulating tumor DNA

Why do we need biomarkers?

Advances in therapeutic strategies have significantly transformed from being one-size-fits-all to being patient/disease specific, thus heralding in the advent of personalized or precision medicine. This has been demonstrated effectively by the use of genomics that generate actionable information in the form of genetic alterations leading eventually to precision oncology.

Clinically validated blood-based biomarkers offer a convenient and minimally invasive gateway to assess disease progression and response/resistance to therapeutic interventions. Prior knowledge about family history, combined with patient genomics, can allow us to identify risks and probabilities in onset and severity of disease (e.g., BRCA germline mutation, APOE e4 protein or

HPV infection leading to breast cancer, Alzheimer's Disease or cervical cancer). Secondly, cancer evolution involves multiple genetic changes including point mutations, gain/loss-of-function (GOF/LOF) mutations, gene copy number amplifications, chromosomal translocations, homologous recombination defects and microsatellite instabilities. Thus, cancer patients with mutations in genes such as KRAS, EGFR and p53 have a poor prognosis, meaning, the chances of survival are low and that conventional chemotherapy will limit benefit. Equally crucial is the diagnosis of cancer or any disease condition which must be based on authenticated biomarker(s). Ideally, diagnostic biomarkers inform about the disease condition even without undergoing extensive imaging examinations.

Biomarker Identification

One of the most challenging aspects is the identification of a biomarker that will be a true reflection of the dynamic changes that take place during disease progression and/or therapy. Several techniques and technological platforms have been utilized for this purpose. These include immunohistochemistry (IHC), fluorescence in situ hybridization (FISH), microarrays, next generation sequencing (NGS) and proteomics (Figure 2).

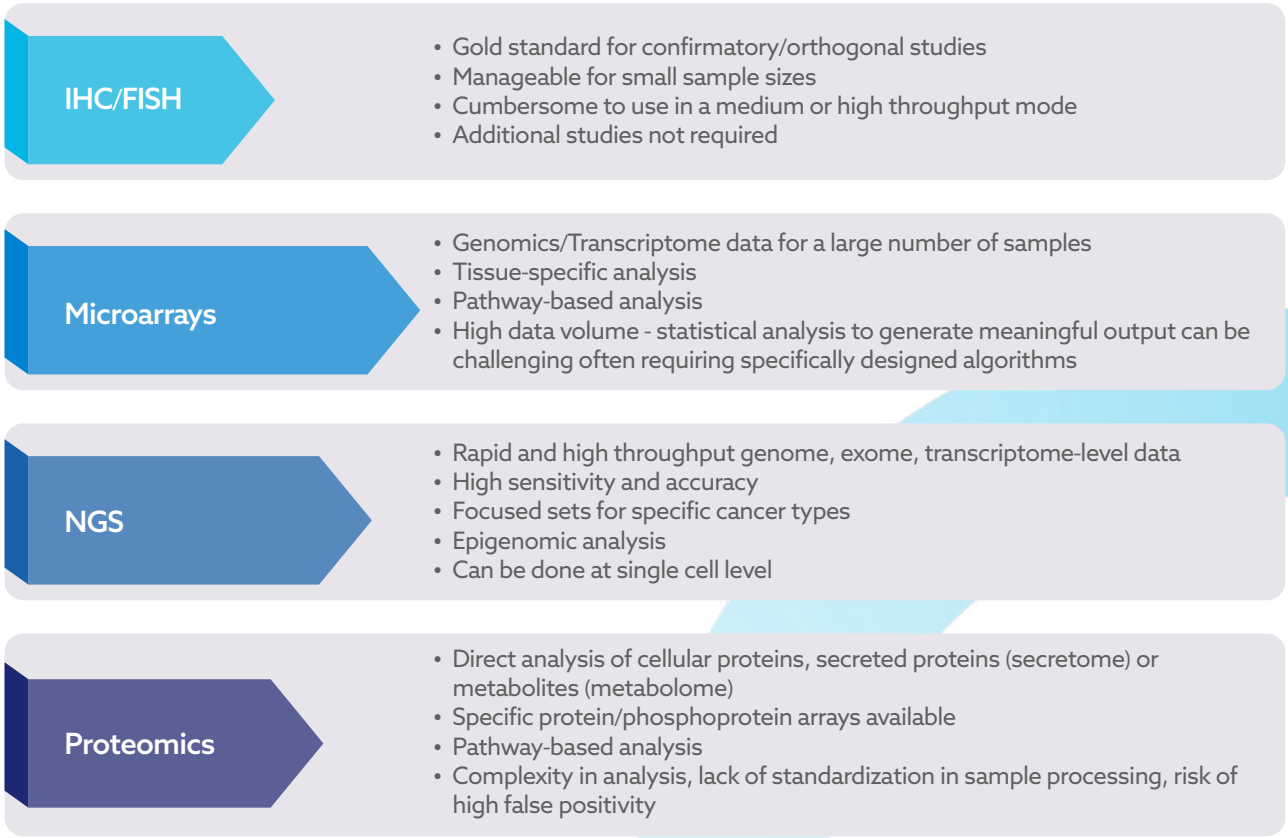

Figure 2: Biomarker identification technology platforms
Identification of biomarkers can be accomplished in several ways including IHC, FISH, microarrays, NGS and proteomic analysis. Each of these techniques/technological platforms has certain advantages and drawbacks. Transcriptomic analysis via NGS and analysis of specific protein arrays followed by confirmatory orthogonal assays are presently the most widely used modalities.
IHC: Immunohistochemistry; FISH: Fluorescence in situ hybridization; NGS: Next generation sequencing

Databases for Biomarker Discovery – Comparative Assessments

Several databases are available for data mining in biomarker applications. These databases have a comprehensive collection of molecular and/or biochemical biomarkers from a rich collection of samples across several therapeutic areas and represent various disease types and subtypes. Data accessibility for for-profit organizations remains a concern for some of these databases as they are available only for the research community. Figure 3 provides the highlights of each of these databases known in the public domain (2, 3).

TCGA

- Molecular characterization of 20,000 primary cancer and matched normal samples across 33 cancer types.
- >2.5 petabytes of genomic, epigenomic, transcriptomic and proteomic data.
- Data available for research community; accessibility to for-profit organizations remains a concern.

ICGC

- Characterization of genomic abnormalities in ~25,000 samples across 50 major cancer types.
- The Data Portal currently contains data from 24 cancer projects including ICGC, The Cancer Genome Atlas (TCGA), Johns Hopkins University, and the Tumour Sequencing Project. It consists of 3,478 genomes and 13 cancer types and subtypes.
- Data availability to for-profit organizations remains a concern.

MSIgDB

- Most widely used comprehensive databases of gene sets for performing gene set enrichment analysis.
- >10,000 gene sets that better represent a wider range of biological processes and diseases.

MarkerDB

- Comprehensive database biomarkers - clinically approved and pre-clinical / investigational markers described in the literature.
- Four major types of molecular biomarkers (chemical, protein, DNA and karyotypic) and four biomarker categories (diagnostic, predictive, prognostic and exposure) associated with more than 27 broad disease types and >600 different conditions or diseases.

GOBIOM

- Comprehensive database provides proteomics, genomic, biochemical, imaging, metabolite clinical scoring and cellular biomarker information.
- Utility in diagnosis, prognosis, disease monitoring, treatment response, surrogate efficacy and toxicity.
- Covers 18 different therapeutic areas; ~3,000 therapeutic indications.

Figure 3: Databases for biomarker discovery

The enlisted databases serve as a key starting point for deriving clinical correlates of potential biomarkers. The collection of biomarkers is based on thousands of clinical samples from diverse therapeutic areas.

Data Analytical and Visualization Tools

Once putative biomarkers are identified and confirmed through orthogonal assays, appropriate programs are required for understanding the biomarker role and positioning in disease biology (including general and specific cellular signaling pathways). Several analytical and visualization tools are available which enhance our understanding of the breadth of potential biomarkers involved in disease and whether they are spanning more than one signalling pathway. Both data sets are critical for designing of clinical trials and for assessment of samples generated from such trials. Figure 4 provides a compilation of tools available for data analytics.

Ingenuity Pathway Analysis

- Licensed by Qiagen
- False positives
- Cost and effort involved in confirmatory experiments

iPathway Guide

- Advaita Bioinformatics
- Eliminate false positives and correctly identify significantly impacted true positive pathways
- Proprietary Impact Analysis quickly identifies the significantly impacted pathways based on two forms of evidence: Over Representation Analysis, and Perturbation Analysis.

KEGG

- Kyoto Encyclopedia of Genes and Genomes: a knowledge base for systematic analysis of gene functions, linking genomic information with higher order functional information
- Extremely complex meshwork
- Expected connections often missing when specific data is queried/retrieved

STRING

- STRING (Search Tool for the Retrieval of Interacting Genes/Proteins), a biological database and web resource for mining known and predicted protein–protein interactions
- The graphical representation of the protein interactions provides a high-level view of functional linkage, facilitating the analysis of modularity in biological processes
- Protein linkages across >10,000 organisms, >50 million proteins and >20 billion interactions

Clarivate Analytics

- Enriched data, insights & analytics, workflow solutions and expert services in the areas of Academia & Government, Intellectual Property and Life Sciences & Healthcare
- Competitive intelligence
- Costly licensing

Figure 4: Avenues for data analysis and visualization

Several tools are available for analyzing data obtained from biomarker studies. These tools enable researchers to identify upstream and downstream connections of the target biomarker as well as their role across cell signaling networks.

Biomarker Discovery Services at Enzene – Intelligent, Fit-for-Purpose Multi-omics Approach

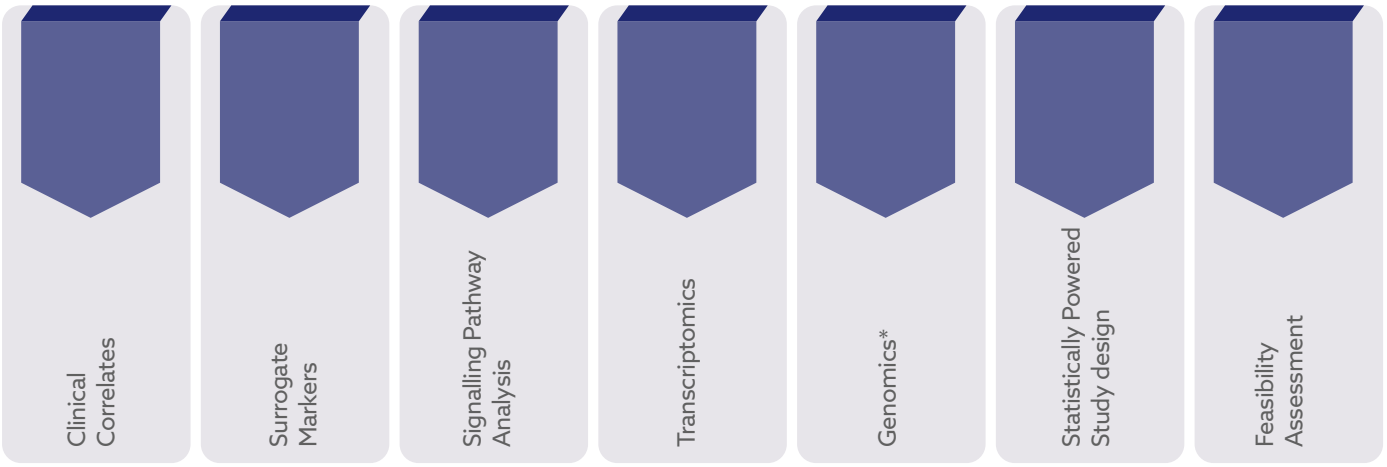

Figure 5: Biomarker identification – design & strategy
Feasibility analysis is a critical aspect involving cell line and/or animal models with a keen eye on the human disease phenotype. Statistically designed study parameters enable us to generate reliable and robust data with high translatability from *in vitro* to *in vivo* to patient. Several paradigms can be utilized to discover novel potential biomarkers and further orthogonal assessments as confirmatory studies.
*indicates outsourced activity.

An intelligently designed combination of the various platforms would yield a credible biomarker or a set of biomarkers. Such biomarkers (or molecular indicators) have the potential to be used as depicted in Figure 1 depending on the nature and extent of their involvement in disease pathology. Although precision oncology has revolutionized cancer medicine, further deep diving is required to do the following:

- Understand and uncover the molecular aspects of inherent and/or acquired drug resistance
- Decipher molecular changes in subtypes of certain cancers
- Delineate genomic alterations indicative of disease relapse/recurrence
- Devise strategies for effective and long-term longitudinal disease monitoring

In addition to oncology, there are other disease areas in dire need of credible biomarkers. These include CNS disorders such as Alzheimer's Disease (AD), Parkinson's Disease (PD), chronic pain (neuropathic pain), inflammation, etc. Intelligently designed laboratory approaches coupled with state-of-the-art technologies that have a clinical implication in mind will allow rapid advances in these areas with unmet biomarker need. Enzene Biosciences Ltd. adopts a customized, client-centric approach and offers versatility by having multiple technological platforms for meeting project and client needs. Enzene possesses expertise in performing biomarker and related studies using a flow cytometer and cell sorter as well as a Luminex xMAP® multiplexing platform. This rich infrastructural arsenal enables Enzene to cater to and execute a plethora of project ideas, thus meeting diverse client requirements.

About the Authors

Dr. Abhishek Mathur is a seasoned healthcare executive with over 17 years of expertise in Biologics R&D, Product Development, and Operations. With a proven track record of interdisciplinary leadership across various functions spanning early research to product commercialization, his extensive experience reinforces our commitment to innovation and patient-centricity. Prior to joining Enzene, he held prominent roles at pioneering biopharmaceutical firms like Amgen and Regeneron in the US, where he successfully guided numerous products from early development to commercialization.

Dr. Mathur holds a Chemical Engineering degree from IIT Bombay, a PhD in Biological Sciences from Northwestern University (USA) and an MBA from Duke University (USA).

At Enzene, Abhishek heads Research and Development operations, guiding the team in innovative strategies and ensuring the smooth execution of projects to achieve the company's goals in biotechnology advancements.

Dr. Mandar Bhonde has a Ph.D. in Cancer Biology and is an experienced drug discovery and development specialist in the area of novel small molecule therapeutics. With 15+ years in the industry, Mandar has a proven research and innovation track record in the form of several publications and patents. In his previous tenure at Lupin Ltd., Mandar successfully led the oncology portfolio driving it to commercialization in the form of two multi-million dollar deals involving big pharma companies. Currently, at Enzene, Mandar heads Discovery Operations and Bioanalysis, bringing in scientific and conceptual leadership and ensuring smooth project execution.

References

1. Fang-Shu Ou et al., J Thorac Oncol. 2021 Apr; 16(4): 537–545. Biomarker Discovery and Validation: Statistical Considerations.
2. Otília Menyhárt 1 and Balázs Győrffy. Comput Struct Biotechnol J. 2021 Jan 22:19:949-960. doi: 10.1016/j.csbj.2021.01.009. eCollection 2021. Multi-omics approaches in cancer research with applications in tumor subtyping, prognosis, and diagnosis.
3. David S Wishart et al., Nucleic Acids Res. 2021 Jan 8;49(D1):D1259-D1267. doi: 10.1093/nar/gkaa1067. MarkerDB: an online database of molecular biomarkers.

ENZENE

Passion. Innovation. Life.

Ready to bring your next molecule to life?

Contact us today and DISCOVER how ENZENE can help you: BD@enzene.com

www.enzene.com

India • USA