Communicating Research to the General Public

At the March 5, 2010 UW-Madison Chemistry Department Colloquium, Prof. Bassam Z. Shakhashiri, the director of the Wisconsin Initiative for Science Literacy (WISL), encouraged all UW-Madison chemistry Ph.D. candidates to include a chapter in their Ph.D. thesis communicating their research to non-specialists. The goal is to explain the candidate's scholarly research and its significance to a wider audience that includes family members, friends, civic groups, newspaper reporters, program officers at appropriate funding agencies, state legislators, and members of the U.S. Congress.

Over 20 Ph.D. degree recipients have successfully completed their theses and included such a chapter.

WISL encourages the inclusion of such chapters in all Ph.D. theses everywhere through the cooperation of Ph.D. candidates and their mentors. WISL is now offering additional awards of $250 for UW-Madison chemistry Ph.D. candidates.

Wisconsin Initiative for Science Literacy

The dual mission of the Wisconsin Initiative for Science Literacy is to promote literacy in science, mathematics and technology among the general public and to attract future generations to careers in research, teaching and public service.

UW-Madison Department of Chemistry
1101 University Avenue
Madison, WI 53706-1396
Contact: Prof. Bassam Z. Shakhashiri
bassam@chem.wisc.edu
www.scifun.org

May 2014

Advancing Mass Spectrometry-based Metabolomics and Proteomics: from Method Development to Disease Applications

By

Ling Hao

A dissertation in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Pharmaceutical Sciences)

at the

University of Wisconsin-Madison

2017

Date of final oral examination: 05/19/2017

This dissertation is approved by the following members of the Final Oral Committee:

Lingjun Li, Professor, Pharmacy and Chemistry

Ron Burnette, Professor, Pharmacy

Paul C. Marker, Professor, Pharmacy

Michael R. Sussman, Professor, Biochemistry

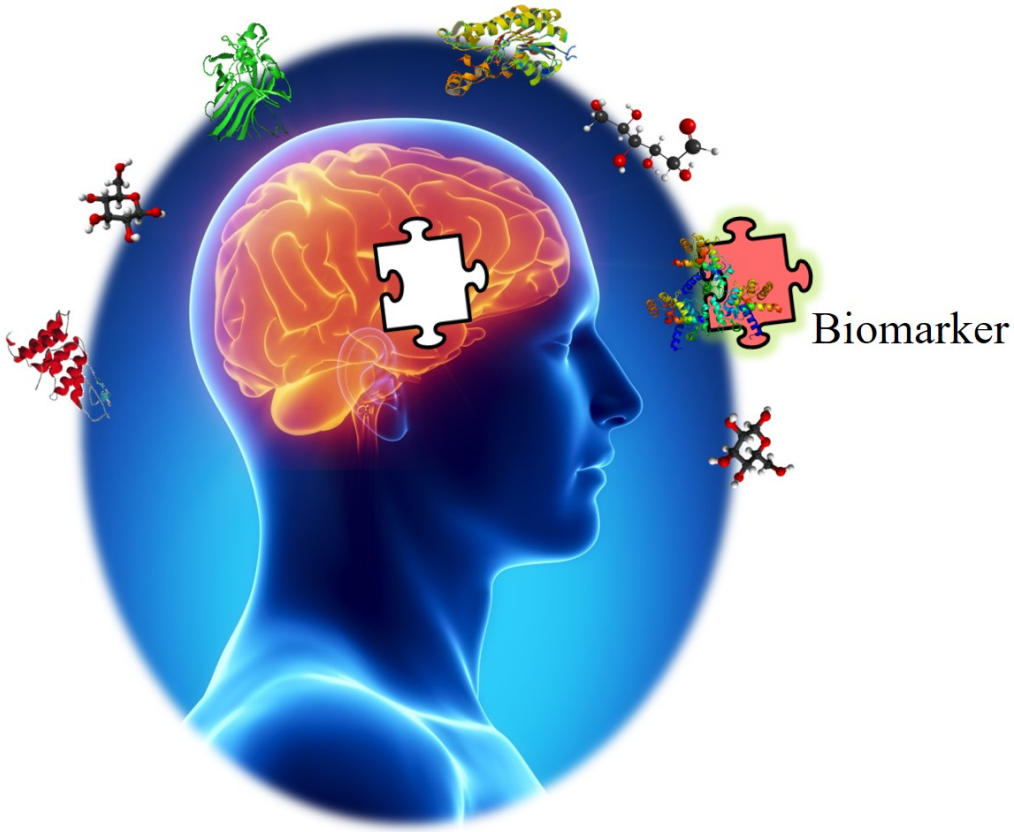
Chad M. Vezina, Associate Professor, Veterinary Medicine

William Ricke, Associate Professor, Urology

Chapter 12

Chapter for Wisconsin Initiative for Science Literacy Award:

Discovering the Molecular Signatures of Human Disease



Brain image adapted from images.google.com

Introduction

Imagine you are going to see your doctor and get your urine and blood tested in the hospital. Your test results reflect some changes in your body fluid compared to the healthy range, which can help the doctor to diagnose an underlying disease. Yet far before the tests can capture some changes, the disease has already affected the body. Some proteins, amino acids, carbohydrates, or lipids levels in your body have changed because of the disease. These large or small molecules present in biological systems are defined as biomolecules. Similar as using blood glucose level for the diagnosis of diabetes, some changed biomolecules in the body can be used to indicate specific health or disease state. We also call them biomarkers. My PhD research is focused on developing novel and improved methodologies for the analysis of biomolecules and applying these techniques to human disease biomarker discovery. By discovering the biomarker changes in human body, we can better understand the disease mechanism and support the future development of new diagnostic tests.

The Process of Human Disease Biomarker Discovery

For human disease biomarker discovery, we first need to design the study and recruit patients from the hospital including a diseased group and a healthy control group. We then collect biological samples from these consenting patients, for instance tissue, blood, urine, and other body fluids depending on the disease of interest. As shown in the pipeline in Figure 1, we can screen the biomolecules in the biological samples using an untargeted discovery approach and compare the biomolecules between the diseased patient group and the control patient group. The significantly changed biomolecules are identified as candidate disease biomarkers. After this,

we need to recruit more patients and perform targeted method to validate candidate biomarkers and measure the accurate amount of those biomarkers in each sample. Before the application to clinical practice, biomarkers also need to be validated in the general population (>1000 patients) to evaluate their diagnostic performance.

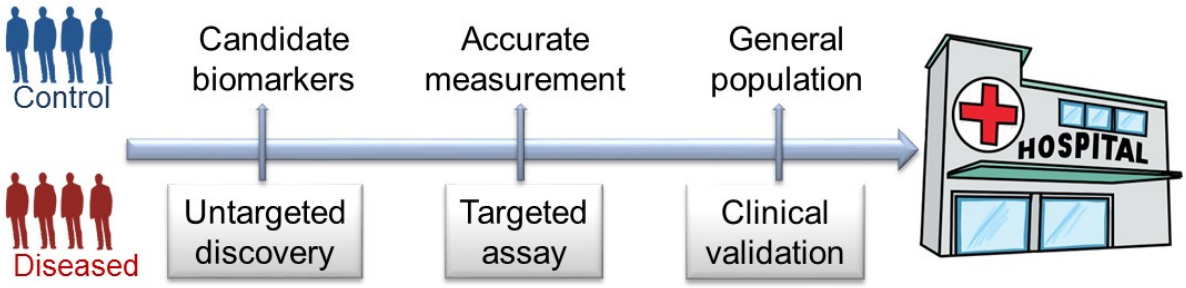


Figure 1. The pipeline of human disease biomarker discovery.

A Crucial Technology for Biomarker Discovery: Mass Spectrometry (MS)

The basis of disease biomarker discovery is the analysis of biomolecules in biological samples. There are hundreds of thousands of biomolecules within a biological sample. We need a robust chemical/biological technology for the effective and efficient analysis of biomolecules. Mass spectrometry (MS) is currently the foremost technology to study proteins, peptides, and small molecules as well as their dynamic alterations involved in disease processes. MS measures the mass-to-charge ratios of molecules and is comprised of a sample inlet, an ionization source, a mass analyzer and an ion detector. Sample molecules are introduced into the instrument through the sample inlet. They can become positively charged or negatively charged ions in the ionization source and fly into the instrument under a magnetic field. The molecules can be separated based on mass-to-charge ratios in the mass analyzer. Mass analyzer can also isolate

charged ions and fragment them to detect both parent ions and fragmentation ions. The detector can then convert the ion energy into electrical signals and transmits to a computer.

For the study of complex biological samples, we often couple an additional instrument platform with mass spectrometry to separate different biomolecules before the MS detection. The most widely used separation platform is called liquid chromatography (LC). A typical workflow of LC-MS analysis is provided in Figure 2, where we first collect biological samples and extract molecules of interest from the sample. We then purify the sample extract and inject the sample into an LC-MS instrument, followed by data analysis using bioinformatic software packages. LC-MS platform provides accurate, sensitive, and reproducible measurements of thousands of biomolecules simultaneously, and has therefore become a central tool for biomarker discovery.

However, the biomolecules in complex biological systems have high structural diversity and wide dynamic range of concentrations. Datasets generated by an LC-MS instrument is multi-dimensional and requires special software packages for data analysis and statistical tools for data interpretation. My PhD research focuses on addressing some of these challenges by developing novel and improved LC-MS-based methods and data analysis strategies, as well as applying these techniques to the studies of human diseases.

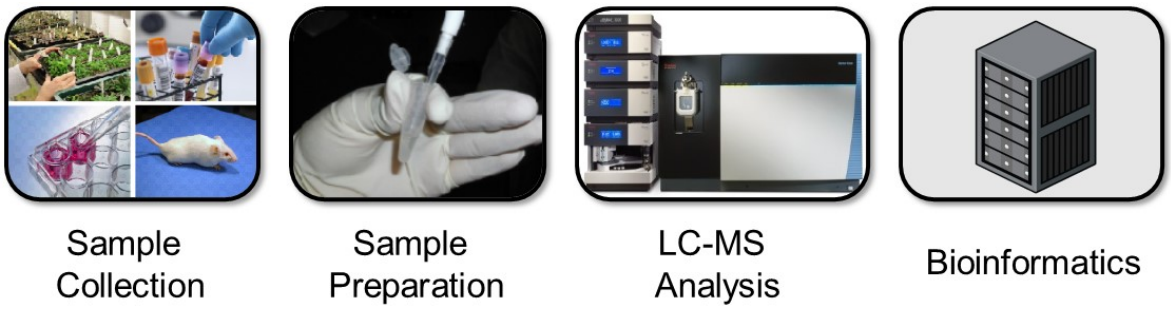


Figure 2. A typical workflow of LC-MS-based study of biomolecules.

My Research Project: Biomarker Discovery of Preclinical Alzheimer's Disease

Alzheimer's disease (AD) is a neurodegenerative disorder affecting millions of elderly people worldwide. As it progresses, the brain tissue starts to shrink which causes problems with memory, thinking and behavior. Unfortunately, the disease process begins years before the diagnosis of dementia. And even worse, currently there is no cure for the disease. The development of AD has evolved to be a continuum over the past decade, with a new disease framework including the preclinical (pre-symptomatic) stage, mild cognitive impairment, and dementia. The pathophysiological process of AD starts years before the emergence of clinical syndrome, yet unequivocal diagnosis and treatment in the early phase of AD is still lacking. It is highly possible that patients could be optimally treated in the preclinical stage of AD before the occurrence of clinical symptoms.

Cerebrospinal fluid (CSF) circulates within the brain ventricular system, maintains metabolic homeostasis of the brain, and is therefore the most direct and valuable biofluid sample to evaluate the dysfunction of the brain. Discovering the metabolic changes in CSF samples derived from preclinical AD patients vs. control patients can provide critical insights into disease progression and support the early diagnosis and treatment of AD.

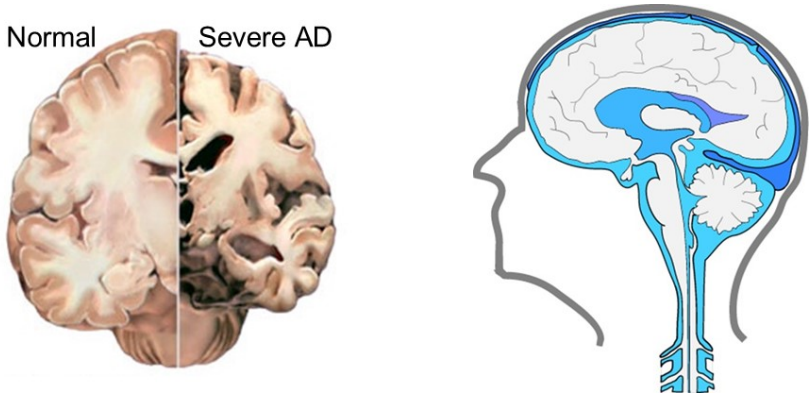


Figure 3. Left: Comparison of human brain tissues between normal state and severe Alzheimer's disease (Image obtained from Alzheimer's Association: http://www.alz.org). Right: The distribution of cerebrospinal fluid in human central nervous system.

Our study aimed to discover small molecule biomarkers of preclinical AD using human CSF samples. As a chemist, I first developed an LC-MS-based analytical method to detect and quantify small molecule metabolites in the human cerebrospinal fluid. Then, I compared the molecular profiles of preclinical Alzheimer's disease patients (N=16) to control patients (N=14). I detected over 4000 compounds from the LC-MS dataset and constructed a statistical model to separate preclinical AD and control groups (Figure 4). Using bioinformatic tools, we identified a panel of most significantly changed molecules in preclinical AD vs. control patients as candidate biomarkers. A representative list of candidate biomarkers is provided in Table 1. We also mapped these dysregulated compounds into functional metabolic pathways to understand how these molecules work together in the human body.

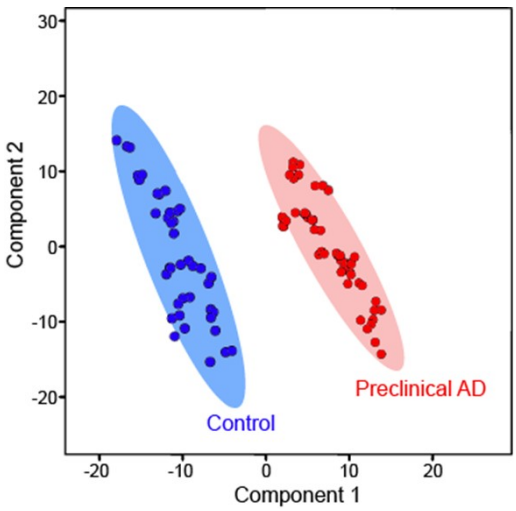


Figure 4. Visualization of group separation (Preclinical AD vs. Control) using the molecular profile of all detected compounds. Each dot represents the CSF profile from one human patient in a two-dimensional space.

Table 1. Representative candidate biomarkers of preclinical AD in human CSF samples.

Name	Detected molecular weight	Changing trend
Free fatty acid	256.2406	Decreased in Preclinical AD
Sulfosalicylic acid	217.9875	Decreased
Valyl-valine	216.1468	Decreased
Acylcarnitine	309.1929	Decreased
3-Oxododecanoic acid	214.157	Decreased
Methionine	149.0504	Decreased
Uridine	244.0698	Decreased
3-Dehydroquinic acid	190.0472	Decreased
Glutamine	146.0681	Increased in Preclinical AD
Tyrosine	181.0731	Increased
Pimelic acid	160.0729	Increased
2-Oxovaleric acid	116.0471	Increased
Hydroxyadipic acid	162.0522	Increased
Acetamidopropanal	115.0631	Increased
Formylanthranilic acid	165.0419	Increased
Glucose/Fructose	180.0628	Increased
Lithocholic acid	376.3003	Increased

The candidate biomarkers we discovered for preclinical AD can offer new insights and serve as important molecular targets for the study of disease mechanisms. We can also build a classification model with the biomarkers to differentiate disease and control patients, which can be used to predict the status of an unknown patient and develop diagnostic tests in the future. Our laboratory continues to analyze CSF samples of human patients from different stages of AD, including preclinical, mild cognitive impairment, and dementia. We can elucidate the dysregulations and functional roles of CSF biomolecules during disease progression and validate candidate biomarkers through targeted strategies. But just like the process of drug discovery, disease biomarker discovery is really challenging and requires strict validation processes and collaborative efforts between chemist, biologist, statistician, physician, and even patient. I believe my research can make a difference and change the way we think about disease diagnosis. By looking at the molecular signatures of disease, we can solve the puzzle of disease and help doctors to make a more precise diagnosis and provide individualized treatment to the patient.