# Comprehensive Lipid Analysis: A Powerful Metanomic Tool for Predictive and Diagnostic Medicine

Steven M. Watkins PhD

FAME Analytics, Inc., West Sacramento, CA, USA

**Key words:** lipid analysis, genomics, informatics, predictive medicine, diagnosis

*IMAJ 2000;2:722–724*

*See also the article on page 657*

The recent explosion of data acquisition and analysis technology, termed informatics, promises to revolutionize predictive and diagnostic medicine. The information readily available to clinicians and researchers today dwarfs that of even a few years ago, and will expand at an even more accelerated rate in the next few years. Managing this information and applying it to useful purpose are formidable challenges. Currently, genomics is the most developed and recognized form of biological informatics. However, genomics is not a panacea for predictive medicine because phenotype is not necessarily predicted by genotype. In the chain of biomolecules from genes to phenotype, metabolites are the quantifiable molecules with the closest link to phenotype. Thus, metabolite informatics, or metanomics, represents a more logical approach than genomics for identifying trends or metabolic profiles of specific diseases. Discouragingly, the complexity of comprehensive metabolite analysis has hindered the progression of metanomics. However, the technology to profile specific subsets of metabolites, in particular lipids, is currently available and should be utilized for this purpose. This report outlines the reasons that lipid analysis represents an attractive and feasible approach to developing a metanomic database capable of producing predictive and diagnostic profiles of disease.

## Informatics

Informatics represents a subtle but significant shift in perspective among biologists. Whereas historically, scientists were accustomed to simplifying their systems to make metabolic interpretations, informatics allows scientists to embrace biological complexity and to make metabolic or phenotypic inference on the basis of as much information as possible. Genomics has brought us the concept of high throughput science, and as a result has demonstrated the power of non-targeted and unbiased data acquisition. Although non-targeted data acquisition is uncommon in metabolite analysis, it does not violate the hypothesis-oriented procedure for scientific study. Rather, high throughput and non-targeted data acquisition simply allow scientists to test their specific hypotheses on a larger, non-biased dataset. This investigative process functions differently from a traditional reductionist approach, where experiments are designed to address single questions. Instead, informatics focuses on obtaining accurate data that can be integrated with other datasets so that future hypotheses can be tested on a database *in silico* rather than at the laboratory bench. This method of investigation is obviously suited to genomics, where sequences from disparate sources are integrated easily into one database. Implementing a comprehensive metanomic database involves planning for considerably more complexity.

## Phenotype is the goal

Although the efforts of the scientific community are still focused on the genome project, it is wise to look forward to the broader usefulness of informatics in medicine. Although genomics is obviously an unprecedented advance for science and medicine, it is, and will continue to be phenotype that interests the medical community most. Genes clearly define the capacity of a cell to process metabolites. Just as clearly, genes and phenotype are not predictably related [1–3]. The reasons for the tenuous association between genotype and phenotype range from an understandable naivete with regard to gene annotation, to the plain fact that genotypes simply code for the metabolic architecture. The gap between genes and phenotype is spanned by many biochemical steps, each with individual specificities and a sensitivity to the environment. Phenotype, although most useful for researchers and clinicians alike, is by far the most difficult to quantify. Consequently, new strategies need to be employed to bring the obvious advantages of informatics into the doctor's office.

## Lipid metanomics

Lipids are a subset of metabolites for which true metanomic study is currently feasible. The technology to quantify the fatty acid, glycerolipid and sterol content of biological tissues or fluids already exists. Thus, only subtle changes in current methodology

are necessary to produce data of sufficient quality to generate metanomic databases of lipids and disease. The major fatty acids in human metabolism and the enzymes that modify them are depicted in Figure 1. Fatty acids are an interesting subject matter for metanomics because they are the only major macronutrients to survive digestion intact, and yet humans possess the biochemical machinery to process dietary fatty acids further into new forms of fatty acid. As a result, the fatty acid composition of tissues and fluids reflects the influence of both diet and metabolism. By quantifying the fatty acids present in human plasma, for example, a researcher could determine the dietary preferences of that individual. Alternatively, and perhaps more interestingly, a researcher could assay endogenous lipid metabolism by comprehensive lipid analysis, because every lipid substrate and product is measured simultaneously from a single sample. Since current technology allows for the comprehensive analysis of lipid composition in a sample, metabolic interpretations can be extended to the activities of the enzymes that modify lipids. Quantitative analysis of fatty acid concentration provides data not only on the fatty acids but on the relative activities of the desaturases and elongases that modify them as well. Moreover, a quantitative analysis of fatty acids from individual glycerolipid classes yields data on the mass of each glycerolipid class, thereby enabling the investigation of pathways involved in glycerolipid metabolism. The ability to not only profile diseases, but also to identify the complex metabolic dysregulations involved in that disease will be a major advance for medicine.

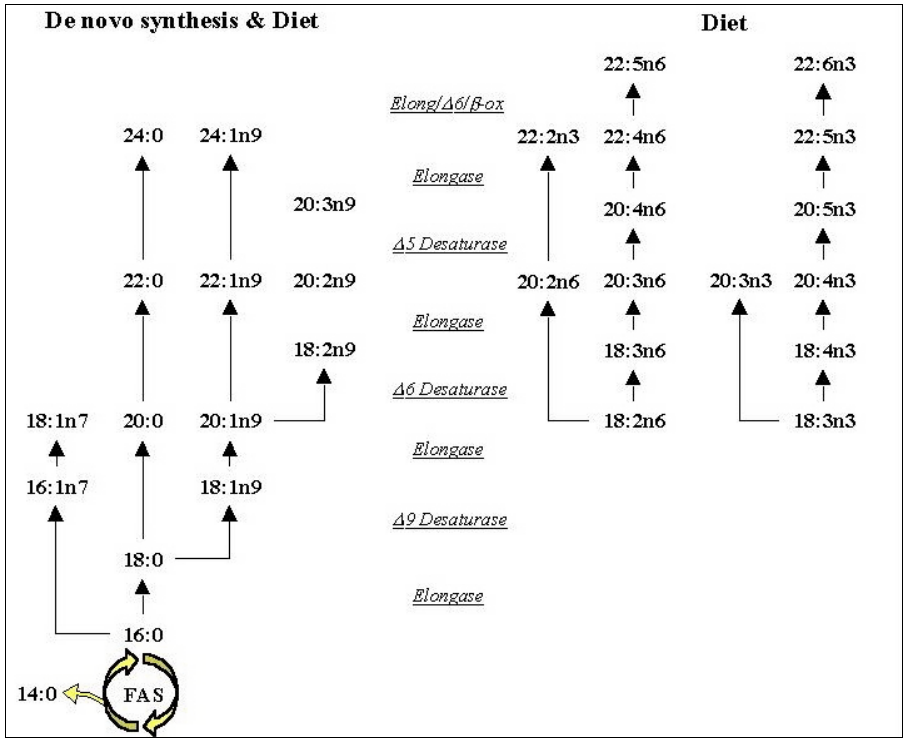


**Figure 1.** Fatty acid metabolism in humans.

## Application of lipid metanomics to predictive medicine

Relative to biomolecules, science has very few ways to quantify phenotype. Alternatively, medicine has, at its very core, a system for identifying, categorizing and recording phenotypic information about individuals. Because science has become exceedingly good at quantifying large numbers of molecules at an astonishing rate of throughput, it is logical that analytical science and medicine should couple their expertise to develop the metabolite-phenotype relationship. By developing a database that allows clinicians to input the patient information and high throughput science to contribute to the analytical data, powerful new predictive and analytical tools are sure to emerge.

The data from a comprehensive lipid analysis produce information useful for this purpose. The applications of a quantitative lipid database are myriad. In one variation, the data from comprehensive lipid analysis are used to generate biomarkers of phenotype. These biomarkers are not, as traditionally defined, single measurements, but rather complex lipid metabolite profiles that include a large number of metabolites and even relations between metabolites. These profiles, when compared among experimental groups, generate a series of significant differences that can be used to construct reliable database filters. A database filter is essentially a list of the most consistent and unique metabolite concentrations or interactions that exist between experimental groups. These differences and interactions are determined by standard statistical methods. The purpose of creating database filters for specific phenotypes is twofold. First, using discriminant analysis or an analogous statistical technique, a database filter can identify entries in a database that match the phenotype of interest. This is an essential element to metanomics and informatics in general, since it allows the scientist to query a database of individuals who were not specifically tested for the phenotype of interest. The second purpose for creating a list of reliable and unique differences between experimental groups is to identify the points in the lipid metabolism pathways most closely linked with the phenotype. As an example of this approach, a researcher might perform an experiment to determine the complete lipid profile of patients with type II diabetes. These data would be recorded with all of the

phenotype and clinical information relevant to the patient in a database. At a later time, another researcher could generate metabolic profiles for individuals consuming dietary olive or fish oils, respectively, and enter this information into the same database. Both researchers would now have the ability to identify groupings of patients that match either diabetic or dietary profiles. Once the data are collected, it is a simple matter of asking the appropriate question *in silico* to determine whether there are relations between dietary oil consumption and diabetes. Additionally, the identified differences act as clues for the metabolic basis of the effect. There are innumerable advantages to an *in silico* approach such as outlined above, including increased statistical power, the avoidance of cumbersome financial and practical limitations to experimentation, and the ability to reassess data as new information emerges. Subject matching, dataset selection and the grouping of experimental sets are all done through *in silico* querying. It is easy to envision that unanticipated relationships between diet, metabolism and phenotype will quickly emerge.

## Methodology

Lipid analysis is an old technology that is well described in the literature; however, several aspects of lipid analysis need updating for lipid metanomics to advance. Currently, most fatty acid analyses are performed by gas chromatography, a technique that provides exquisite separation and quantification of analytes. However, most research groups continue to report their results as percentage of total fatty acids. Data in this format are not comparable between experiments, nor are they comparable between individual lipid classes within an experiment. For example, a scientist interested in the metabolism of oleic acid could not determine the distribution of oleic acid among lipid classes in plasma from mole percentage or weight percentage data. For lipid metanomics as a field to advance, the data produced in each experiment should be expressed as a concentration, for example, micrograms per milliliter, so that a database of lipids in health and disease can be assembled from multiple experiments.

Taking the long view, lipid researchers will need to define the functional and associative domains of lipids to truly understand the association between lipid composition and the function of lipids in biology. Despite the already comprehensive nature of lipid analyses, the selection and preparation of the sample to be analyzed is still a critical step. Each investigator must be aware that lipids of similar chemical composition do not always serve similar physiological functions. Examples of functional or associative domains of lipids are pervasive throughout biology. For example, low density lipoproteins and lymphocytes each contain phosphatidylcholine, although their respective phosphatidylcholine is not likely to serve a similar physiological function. Without separating lipoproteins from lymphocytes prior to analysis, as the example indicates, a database relating the composition of phosphatidylcholine to phenotype would be intrinsically noisy. It is not critical that all associative or functional domains be isolated prior to the analysis. A tremendous amount of useful information can be obtained even at a minimum level of sample definition. However, it is critical for the creation of a useful metanomic database that each investigator records the exact nature of the sample. Lipid metanomics should strive to avoid the confusion surrounding the annotation of the genome.

## Summary

The power and accuracy of predictive diagnostics stand to improve dramatically as a result of lipid metanomics. The high definition of data obtained with this approach allows multiple rather than single metabolites to be used in markers for a group. Since as many as 40 fatty acids are quantified from each lipid class, and up to 15 lipid classes can be quantified easily, more than 600 individual lipid metabolites can be measured routinely for each sample. Because these analyses are comprehensive, only the most appropriate and unique metabolites are selected for their predictive value. Thus, comprehensive lipid analysis promises to greatly improve predictive diagnostics for phenotypes that directly or peripherally involve lipids.

A broader and possibly more exciting aspect of this technology is the generation of metabolic profiles that are not simply markers for disease, but metabolic maps that can be used to identify specific genes or activities that cause or influence the disease state. Metanomics is, in essence, functional genomics from metabolite analysis. By defining the metabolic basis for phenotype, researchers and clinicians will have an extraordinary opportunity to understand and treat disease. Much in the same way that gene chips allow researchers to observe the complex expression response to a stimulus, metanomics will enable researchers to observe the complex metabolic interplay responsible for defining phenotype. By extending this approach beyond the observation of individual dysregulations, medicine will begin to profile not single diseases, but health. As health is the proper balance of all vital metabolic pathways, comprehensive or metanomic analysis lends itself very well to identifying the metabolite distributions necessary for optimum health. Comprehensive and quantitative analysis of lipids would provide this degree of diagnostic power to researchers and clinicians interested in mining metabolic profiles for biological meaning.

## References


1. Strohman RC. Linear genetics, non-linear epigenetics: complementary approaches to understanding complex diseases. *Integr Physiol Behav Sci* 1995;30:273–82.
2. Strohman RC. The coming Kuhnian revolution in biology [see comments]. *Nat Biotechnol* 1997;15:194–200.
3. Schilling CH, Edwards JS, Palsson BO. Toward metabolic phenomics: analysis of genomic data using flux balances. *Biotechnol Prog* 1999;15:288–95.



**Correspondence:** Dr. S.M. Watkins, FAME Analytics, Inc., 2545 Boatman Avenue, West Sacramento, CA 95691, USA. Tel: (1-916) 371-7974; Fax: (1-916) 375-6508;
email: steve.watkins@fameanalytics.com