Archives of Biochemistry and Biophysics 589 (2016) 158–167

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Review article

Application of quantitative metabolomics in systems genetics in rodent models of complex phenotypes

Dominique Gauguier [a, b, *]

[a] Sorbonne Universities, University Pierre & Marie Curie, University Paris Descartes, Sorbonne Paris Cité, INSERM UMR_S 1138, Cordeliers Research Centre, 75006 Paris, France
[b] Institute of Cardiometabolism and Nutrition (ICAN), University Pierre & Marie Curie, 91 boulevard de l'Hôpital, 75013 Paris, France

ARTICLE INFO

Article history:
Received 2 August 2015
Received in revised form
10 September 2015
Accepted 17 September 2015
Available online 25 September 2015

Keywords:
Functional genomics
Genetic linkage
Genetic association
Genome
Mouse
Quantitative trait locus
QTL
Metabotypes
Rat

ABSTRACT

Genome-wide association studies (GWAS) have provided remarkable advances in our understanding of the etiology of complex diseases in humans and have underlined the need to improve patients' phenotype characterization with intermediate molecular phenotypes. High resolution metabolomics is becoming an increasingly popular and robust strategy for metabolic phenotyping large cohorts of patients and controls in genetic studies, in order to map the genetic control of metabotypes in various biological matrices (organ extracts and biofluids) through Quantitative Trait Locus (mQTL) analysis. This article reviews results from ongoing research in mQTL mapping in rodent models of human complex traits, with a specific focus on the cardiometabolic syndrome, and prospects of applications of untargeted metabolomics to improve knowledge of multilevel genome expression control in health and disease and to detect potential novel biomarkers for complex phenotypes in experimental systems in mice and rats.

Contents

1. Introduction

Functional genomic tools (mainly transcriptomics, proteomics, metabolomics) are exploratory systems of genomic regulations that

* INSEMR UMRS_1138, Cordeliers Research Centre, 15 rue de l'Ecole de Médecine, 75006 Paris, France.
E-mail address: dominique.gauguier@crc.jussieu.fr.

http://dx.doi.org/10.1016/j.abb.2015.09.016
0003-9861/

document different but connected and strongly interdependent dimensions in the continuum of genome-wide gene expression [1]. Transcriptome analysis based on microarrays and RNA sequencing is a practical approach for gene expression studies because signals directly correspond to coding or non coding transcripts localized in the genome and often well-annotated in public genome databases (www.ensembl.org). It uses universal and standardized systems for both data production and analysis, allowing data generated in different laboratories to be directly compared. The proteome covers a complex dimension of genome expression, which includes regulation of protein abundance, conformation and activity directly relevant to physiological phenotypes, but the technologies used still lack throughput. The metabolome addresses the regulation of low molecular weight compounds (lipids, carbohydrates, amino acids, …) which are often end products of genome expression in various biological matrices (organ extracts, cell preparation, urine, plasma, cerebrospinal fluid) [2,3]. It is therefore highly relevant to pathophysiological phenotypes, with broad ranging applications in deciphering the complex pathogenesis and etiology of increasingly frequent and prevalent multifactorial disease conditions combining genetic risk and environmental influences. These include primarily pathophysiological elements of the cardiometabolic syndrome (CMS) (glucose intolerance, obesity, hypertension, dyslipidemia) and its vascular complications [4], but also autoimmune diseases and neurobehavioral and neurodegenerative conditions.

There is growing interest in metabolome profiling technologies for a wide range of applications, including primarily in toxicology and drug discovery [3] but also, more recently, in genetics [5]. Deciphering the genetic determinants of complex disorders is rapidly evolving from pure genetic association studies based on disease status to the analysis of the genetic control of molecular phenotypes used as disease biomarkers [6]. In contrast to transcriptomics, which requires the use of organ biopsies or primary cells, that are only partly relevant to organ function, metabolomic studies can be carried out with biofluids (plasma and urine) collected using minimally invasive sampling methods thus allowing repeated measures and longitudinal studies. Metabolomics is therefore expected to become an increasingly powerful molecular phenotyping technique for genetic studies, by inferring causality between genetic polymorphisms and variations in concentration of metabolites in a biological sample. These may represent disease associated molecular markers when the underlying genetic effects co-localise in the genome with those of disease relevant phenotypes. Genetic mapping of metabolomic traits in mammalian species was originally pioneered in a rat model of type 2 diabetes [7] and later applied in mice [8] and human GWAS [9], but causality between metabolomic features and disease remains to be established.

In this article, I review the application of metabolomics in the detection of potential biomarkers for complex phenotypes in experimental systems in mice and rats, through their genetic mapping in experimental crosses and panels. I also address the important prospects of untargeted metabolomic profiling strategies in quantitative genetics to extend genome mapping analyses to all metabolite features that can be detected and quantified in spectral data, and the integration of metabolomic datasets with other dimensions of genome expression.

1.1. Animal models of human complex phenotypes

Even though genome-wide association studies (GWAS) have provided remarkable advances in our understanding of the etiology of complex diseases in humans [10], the biological mechanisms underlying the function of disease risk loci remain elusive in the vast majority of studies. Physiological and genomic studies in animal models can provide important clues on the biological function of these loci. Mouse and rat models are the preferred mammalian systems for such investigations because large progenies can be obtained in short period of time and can be maintained in standardized or carefully monitored environmental (diet, external challenge) and maintenance (temperature, humidity, light/dark cycle) conditions, thus reducing interindividual variability. They can be used for invasive physiological studies and provide a source of biological material from organs that cannot be accessed in humans for genomic studies and for production of primary cells for *in vitro* studies. A wide range of disease models that mirror human disease conditions are available in both rats and mice, which allow extensive physiological and genomic studies to identify the mechanisms involved in disease onset and progression.

A key advantage of rat and mouse models lies in the availability of inbred strains that are genetically homogeneous and have been fully characterized for genetic polymorphisms by genome sequencing [11–13]. In addition to strains that carry mutations or naturally occurring genetic polymorphisms that lead to phenotypic alterations, a wide-range of models are established for disease phenotypes induced experimentally by gene editing (knock out, transgenesis), random mutagenesis (N-ethyl-N-nitrosourea-ENU), environmental changes (obesity and diabetes induced by high fat high sucrose diet; hypertension induced high salt intake) or chemical treatment (diabetes induced by streptozotocin, atopy induced by gold salts). Mouse and rat strains provide complementary models for physiological and molecular investigations, depending on the scientific questions addressed.

The laboratory mouse is often preferred to study immunological phenotypes and to carry out gene editing, whereas the laboratory rat is the leading model species in pharmacology and toxicology [14]. Rat models provide the most relevant models for the accurate analysis of whole organism, organ and cellular phenotypes relevant to multifactorial disorders. They are therefore particularly powerful systems for the study of pathophysiological elements of the CMS that requires complex physiological procedures to analyse longitudinally blood pressure, glucose and lipid homeostasis, body fat composition and simultaneous biological explorations of the function of several organs. Many of the non-invasive and invasive phenotyping techniques that are readily available in the rat, including for example blood pressure measurements, remain difficult or impossible to apply in other species, including the mouse. The laboratory rat is also a practical model system for repeated sampling of large volumes of blood and urine required for metabolomic studies.

1.2. Genetic mapping panels in rats and mice

The fundamental basis of genetic studies of biological and molecular regulations is to determine causality between genotypes of genetic markers across the genome and one or several phenotypes of interest. In practice, causality can be assessed through statistical analyses of phenotype variation and either segregation of alleles (genetic linkage) or frequency of alleles (genetic association) in a cohort of genetically heterogeneous individuals. Applying these approaches in experimental systems is convenient because one can consider both genetic background differences and contrasting phenotypic features between inbred strains in order to design the optimal genetic cross to map genetic loci controlling a phenotype of interest. Based on this information, classical genetic crosses (F2 cross, backcross) can be arranged to produce a cohort of genetically heterogeneous hybrids where alleles and phenotypic traits segregate (Fig. 1). Panels of recombinant inbred (RI) strains available in rats and mice provide convenient systems for extensive phenotypic screenings in genetic studies. In contrast to classical crosses, these