Available online at www.sciencedirect.com

SCIENCE DIRECT

Transplant Immunology 14 (2005) 255 – 266

Transplant Immunology

www.elsevier.com/locate/trim

Strategies in analysis of the genetic component of multifactorial diseases; biostatistical aspects

Thomas Barnetche*, Pierre-Antoine Gourraud, Anne Cambon-Thomsen

Unité INSERM 558, Department of Epidemiology Faculté de médecine, 37 allées Jules Guesde, F-31073 Toulouse, France

Accepted 14 March 2005

Abstract

Complex polygenic and multifactorial diseases remain a challenge for human geneticists. Here we aim to remind basic definitions of multifactorial diseases and the genetic related concepts underlying classical methods. Knowledge on pathophysiological process and the genetic information available conditions the design of study. The choice of methodology, between candidate gene approach and genome scan approach, between linkage and association studies, is the most important step. Both methods, linkage analysis and association studies are usually considered as complementary approaches for a given disease. For this reason, in this article, we present the most important classical methodologies in genetic epidemiology of complex disorders. References and examples are given to illustrate.

Keywords: Multifactorial diseases; Association studies; Linkage studies; Immunogenetics; Statistics methods

1. Introduction

In an age when the majority of monogenic human disease genes have been identified, a particular challenge for the coming generation of human geneticists will be resolving complex polygenic and multifactorial diseases. Many studies have been set up to understand etiology of common complex diseases but clear definition of a multifactorial human disease is still missing and their analysis methods are still a matter of debate.

Immunogenetic factors and genes are always excellent candidates in etiology of many complex diseases. The goal of this article is thus to describe the major analysis methods available in genetic epidemiology to study multifactorial disorders. Here we remind of basic definitions of the so-called multifactorial diseases and the genetic epidemiology related concepts underlying classical methods.

2. Definition and study concepts of complex diseases genetics

Apart from monogenic diseases (where only one gene intervenes in etiology), there is a set of complex diseases in which genetic and environmental factors interact. Characteristically, many genes interact in complex ways with multiple environmental factors and with each other [1].

The principal distinction between “simple” monogenic diseases and complex genetic diseases is that the latter do not exhibit classic Mendelian pattern of inheritance. The phenotype of a monogenic disease may be complex (neurological, clinical, or biological signs), but those signs all relate to only one gene playing a major role, with possible environmental factors action. In a multifactorial disease, the major obstacle is that the various components of the phenotypes result from the

* Corresponding author. Tel.: +33 5 61145959; fax: +33 5 62264240.
E-mail address: barnetch@cict.fr (T. Barnetche).

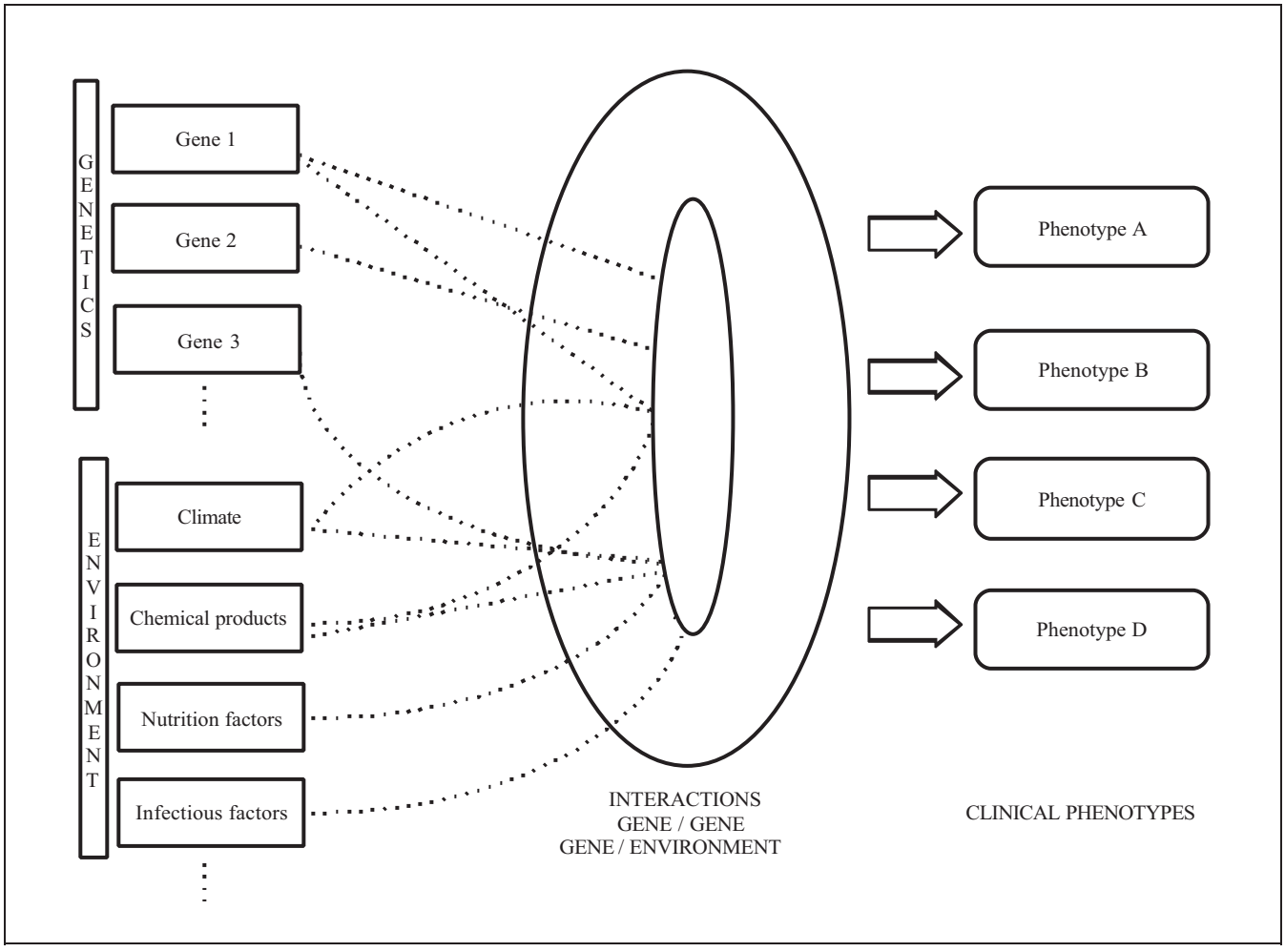
0966-3274/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.trim.2005.03.015

interaction between multiple genes and others factors. The difficulty is to correlate the various components of the phenotype to each individual factor. Each gene has a weak effect, thus studies require large sample size. Studies generally lack statistical power, and the complexity of interactions between environmental factors and phenotypes are difficult to take into account.

Thus the study of complex diseases genetics remains difficult and presents many additional challenges for genetic epidemiology. Consequently, extensions of methodologies designed for monogenic inheritance do not necessarily suit. The complex interplay between genes and between genes and environment is schematically presented in Box A.

Box A

Main features of factors involved in aetiology of complex disease



The parameters to consider for designing a genetic study of a complex disease are at different levels:

1) The knowledge of pathophysiological process:
   - If known or if specific hypotheses on the role of particular genes are available, the *candidate gene* approach is applied. It consists in choosing markers near the candidate gene.
   - If unknown, a *genome scan* approach may be set up using markers distributed over the genome.
2) The methodological strategy most suitable according to the kind of genetic information available:
   - *Linkage studies* which consist in finding a physical linkage between a marker and a gene. These studies deal with family design.
   - *Association studies* which consist in finding a correlation between several factors and the disease. These studies imply in general unrelated individuals or a population.

Genetic linkage refers to co-transmission of polymorphic genetic markers or traits, physically linked on the same chromosome. Linkage studies are focused on the co-segregation of genes, or markers, situated at different loci. One tests the