Genetic studies in complex disease

Rosendaal, F.R.; Reitsma, P.H.; Mollica, L.; Lane, D.A.; Margaglione, M.; Merlini, P.; Ardissino, D.

Journal of Thrombosis and Haemostasis, **2**: 342–345

FORUM

Genetic studies in complex disease

F. R. ROSENDAAL* and P. H. REITSMA†
Departments of Clinical Epidemiology and Hematology, Leiden University Medical Center, Leiden, the Netherlands; and †Laboratory for Experimental Internal Medicine, Academic Medical Center, Amsterdam, the Netherlands.

To cite this article: Rosendaal FR, Reitsma PH. Forum on genetic studies in complex disease. *J Thromb Haemost* 2004; **2**: 342.

See also Souto JC. Genetic studies in complex disease: the case pro linkage. *J Thromb Haemost* 2003; **1**: 1676–8; Rosendaal FR. Genetic studies in complex disease: the case pro association studies. *J Thromb Haemost* 2003; **1**: 1679–80.

While we do not wish to restate the points in the debate with Souto [1] concerning various approaches to find novel genetic risk factors for thrombosis, we wish to comment on one statement in Souto's interesting contribution. He rightly points out that only three genetic factors that were recently discovered have been consistently associated with venous thrombotic risk: ABO blood group, factor (F)V Leiden, and prothrombin 20210A. He then states that FV Leiden is particularly illustrative since it was first found through a linkage study. This is not true: FV Leiden was found, following the report on APC-resistance by Dahlbäck [2], nearly simultaneously by three groups, who all used either mixing experiments or direct sequencing and genotyping of the candidate gene (FV) [3–5]. In our paper, we did report on linkage in a single pedigree as a confirmatory experiment [3]. History has it that none of the six established genetic risk factors for venous thrombosis, where we add deficiencies to protein C, protein S and antithrombin to the list, has been found through linkage. We know, since we were there.

References

1 Souto JC. Genetic studies in complex disease: the case pro linkage studies. *J Thromb Haemost* 2003; **1**: 1676–8.
2 Dahlbäck B, Carlsson M, Svensson PJ. Familial thrombophilia due to a previously unrecognised mechanism characterized by poor anticoagulant response to activated protein C: prediction of a cofactor to activated protein C. *Proc Natl Acad Sci USA* 1993; **90**: 1004–8.
3 Bertina RM, Koeleman RPC, Koster T, Rosendaal FR, Dirven RJ, De Ronde H, Van der Velden PA, Reitsma PH. Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 1994; **369**: 64–7.
4 Voorberg J, Roelse J, Koopman R, Büller H, Berends F, Ten Cate JW, Mertens K, van Mourik JA. Association of idiopathic venous thromboembolism with single point-mutation at Arg506 of factor V. *Lancet* 1994; **343**: 1535–6.
5 Greengard JS, Sun X, Xu X, Fernandez JA, Griffin JH, Evatt B. Activated protein C resistance caused by Arg506Gln mutation in factor Va. *Lancet* 1994; **343**: 1361–2.

Correspondence: F.R. Rosendaal, Clinical Epidemiology, C9-P, Leiden University Medical Center, P.O. Box 9600, NL-2300 RC Leiden, the Netherlands.
Tel.: +31 71 5264037; fax: +31 71 5266994; e-mail: f.r.rosendaal@lumc.nl

L. MOLLICA and D. A. LANE
Imperial College London, Hammersmith Hospital Campus, London, UK

To cite this article: Mollica L, Lane DA. Forum on genetic studies in complex disease. *J Thromb Haemost* 2004; **2**: 342–3.

See also Souto JC. Genetic studies in complex disease: the case pro linkage. *J Thromb Haemost* 2003; **1**: 1676–8; Rosendaal FR. Genetic studies in complex disease: the case pro association studies. *J Thromb Haemost* 2003; **1**: 1679–80.

In their articles as part of the *Debate: Genetic studies in complex disease*, Souto and Rosendaal [1,2] make the respective cases for linkage and association studies. It is impossible to disagree with the positive points made about these approaches, which have produced many landmark discoveries. The enthusiasm for these types of study has been facilitated by progress in molecular genetics, which has considerably simplified the detection of various genetic variants. However, despite the respective strengths of linkage and association studies outlined by Souto and Rosendaal, their widespread use has not always been accompanied by proportionate progress. Lack of reproducibility is certainly one of the most dominant outcomes of their application to the genetic basis of many common complex diseases. The field of atherothrombotic disease is no exception, and is flooded by studies reporting inconsistent and contradictory results. Not surprisingly, this has left a legacy of considerable uncertainty [3,4]. Furthermore, simple application of both approaches will always provide an incomplete picture and give only a preliminary indication of the risk associated with a given genotype. Any causative relationship between genotype and disease must be mediated through a phenotype, and this must be fully understood before genetic risks can be applied clinically.

There are several well-rehearsed potential explanations for the inconsistent literature on genetic risk factors, including poor study design, population stratification and excessive subgroup analysis. Perhaps most important, however, is that the phenotypic effect of a given genotype is likely to be small in complex diseases. Given that the effect of a phenotype on disease is often

Correspondence: David A. Lane, Imperial College London, Hammersmith Hospital Campus, London W12 ONN, UK.
Tel.: +44 (0) 20 8388 2295; fax: +44 (0) 20 8388 2296; e-mail: d.lane@imperialcollege.ac.uk

difficult to demonstrate, the study of a genotype cannot be expected to make this task quantitatively easier. The size of a study is the major determinant of its power. This becomes a critical issue when searching for small increased risk, ~20% or less. Most studies of the risks of haemostatic genetic risk factors in thrombotic disease have been too small to be informative. Conclusions from small studies, especially those failing to show consistent relationships between genotype, phenotype and disease, should be considered with great circumspection. A recent example of how the outcome of a large study (4484 patients and 5757 controls) can overturn conclusions from small studies (~100 patients) is provided elsewhere [5–7]. Inevitably, large study design cannot always be optimal [8], but there is now growing acceptance that they alone provide the increased power that is needed to clarify the role of genes in common disorders such as atherothrombotic disease [9]. Pooling of patients into centrally organized sample and data collections will be necessary to obtain the population size required for these studies. Although this will remove independent research opportunity from investigators with small and medium-sized patient cohorts, it is currently the only feasible approach for tackling the issue of genetic predisposition to common diseases. An example of such an initiative is the Type 1 Diabetes Genetics Consortium (T1DGC), which has been developed to organize international efforts to identify genes that determine an individual's risk of type 1 diabetes (http://www.t1dgc.org).

Although there is current difficulty in drawing firm conclusions from inconsistent results in atherothrombotic diseases, this should neither discourage the search for genetic predisposition to disease, nor imply that molecular genetics is incapable of efficiently approaching diseases with a genetic and environmental complex determinism. Rather, it should provide us a more critical view with which to interpret the literature on genetic risk and from which we can progress.

References

1 Souto JC. Genetic studies in complex disease: the case pro linkage studies. *J Thromb Haemost* 2003; **1**: 1676–8.

2 Rosendaal FR. Genetic studies in complex disease: the case pro association studies. *J Thromb Haemost* 2003; **1**: 1679–80.

3 Lane DA, Grant PJ. Role of hemostatic gene polymorphisms in venous and arterial thrombosis. *Blood* 2000; **95**: 1517–32.

4 Simmonds RE, Hermida J, Rezende SM, Lane DA. Haemostatic genetic risk factors in arterial thrombosis. *Thromb Haemost* 2001; **86**: 374–85.

5 Keavney B, Parish S, Palmer A, Clark S, Youngman L, Danesh J, McKenzie C, Delepine M, Lathrop M, Peto R, Collins R, International Studies of Infarct Survival (ISIS) Collaborators. Large-scale evidence that the cardiotoxicity of smoking is not significantly modified by the apoliprotein E epsilon2/epsilon3/epsilon4 genotype. *Lancet* 2003; **361**: 396–8.

6 Humphries SE, Talmud PJ, Hawe E, Bolla M, Day IN, Miller GJ. Apolipoprotein E4 and coronary heart disease in middle-aged men who smoke: a prospective study. *Lancet* 2001; **358**: 115–9.

7 Stengard JH, Kardia S, Tervahauta M, Ehnholm C, Nissinen A, Sing CF. Utility of the predictors of coronary heart disease mortality in a longitudinal study of elderly Finnish men aged 65–84 years is dependent on context defined by Apo E genotype and area of residence. *Clin Genet* 1999; **56**: 367–77.

8 Humphries SE, Hawe E, Sukhbir D, Miller GJ, Talmud PJ. In search of genetic precision. *Lancet* 2003; **31**: 1908–9.

9 Keavney B, Parish S, Lathrop M, Peto R, Collins R. Large-scale evidence that the cardiotoxicity of smoking is not significantly modified by the apoliprotein E epsilon 2/epsilon3/epsilon 4 genotype. Author reply. *Lancet* 2003; **31**: 1909–10.

M. MARGAGLIONE

Unita' di Aterosclerosi e Trombosi, I.R.C.C.S. 'Casa Sollievo della Sofferenza', S. Giovanni Rotondo; and Genetica Medica, Università di Foggia, Italy

To cite this article: Margaglione M. Forum on genetic studies in complex disease. *J Thromb Haemost* 2004; **2**: 343–4.

See also Souto JC. Genetic studies in complex disease: the case pro linkage. *J Thromb Haemost* 2003; **1**: 1676–8; Rosendaal FR. Genetic studies in complex disease: the case pro association studies. *J Thromb Haemost* 2003; **1**: 1679–80.

The importance of inheritance in human diseases was recognized nearly a century ago. Since then, two contrasting points of view have prevailed, the 'Mendelist and the biometricist view of the world' [1]. Mendelian diseases are traits in which differences are due to different alleles at a single (or few) locus. Linkage analysis has had a remarkable role in leading to the identification of genes responsible for mendelian diseases. However, it has been proven to be a far less reliable tool when applied to non-mendelian (also known as multifactorial or complex) diseases. On the other hand, complex diseases are at the cross-roads between inherited and non-inherited factors, and are the result of the contribution of many different genes to the total variability, with no particular gene having a large effect. In complex diseases, association studies have found hundreds of potential associations with common gene variants, no matter whether the gene variants identified directly affect the risk or are simply a marker. Unfortunately, association studies have shown a long series of false-positive results and very few associations have been consistently replicated [2].

Two different working models underlie the different procedures for studying the contribution of genes to diseases [3,4]. On one hand, the probability of sharing a segment of a chromosome among affected individuals, linkage analysis, leads to the possibility that a specific region harbors one (or more) pathogenic gene(s). Then, one has to demonstrate the biological plausibility of the identified gene (reverse genetics). On the other hand, in association studies we need to have an *a priori* hypothesis on the role of the investigated gene (forward genetics). As far as venous thromboembolism is concerned, the latter approach gave remarkable results leading to the identification of important risk factors, such as the factor V Leiden and the factor II A20210 allele [5,6]. It is not unlikely that future association

Correspondence: M. Margaglione, Unita' di Aterosclerosi e Trombosi, I.R.C.C.S. 'Casa Sollievo della Sofferenza', S. Giovanni Rotondo; and Genetica Medica, Università di Foggia, Italy.

Tel.: +39 0881 733842; fax: +39 0881 732188; e-mail: ate.tro@operapadrepio.it

studies may be able to find new gene variants affecting susceptibility to venous or arterial thrombosis. Linkage analysis showed, at the best, a series of loci that could explain a portion of the interindividual variability of intermediate phenotypes [7,8]. However, we must look ahead. The analysis of the genome sequence revealed a high percentage of annotated genes with unknown functions (42%). The percentage was higher if one also considers hypothetical genes (59%!) [9]. In other words, we know of the existence of a huge number of genes, but for most of them we ignore what they do. We need to know their function. It is conceivable that a number of new genes may be functional for normal hemostasis and allelic variants may play a significant role in common, complex diseases, such as atherosclerosis, venous and arterial thrombosis. In the meanwhile, we have to rethink our approach in the search for the genetic component of complex diseases. As we are facing a spectacular and rapid advance of technology, the association of computational and molecular technology may offer new models to approach this search. An example of this is the identification of a new gene associated with common forms of ischemic stroke [10]. The deCODE genetics of Iceland has identified as a susceptibility locus the gene encoding phosphodiesterase 4D (PDE4D) by merging linkage analysis and association studies, investigating a high number of individuals and adopting strict epidemiological criteria. Traditional linkage analysis and association studies have been successful in locating loci and identifying functional variants, but they are probably coming at an end.

It is conceivable that in the future the identification of genes affecting susceptibility to complex diseases will require new methods of population-based genetic investigation taking advantage of high-throughput genomic methods and broad community collaboration.

References

1 Risch NJ. Searching for genetic determinants in the new millennium. *Nature* 2000; **405**: 847–56.
2 Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002; **4**: 45–61.
3 Souto JC. Genetic studies in complex disease: the case pro linkage studies. *J Thromb Haemost* 2003; **1**: 1676–8.
4 Rosendaal FR. Genetic studies in complex disease: the case pro association studies. *J Thromb Haemost* 2003; **1**: 1679–80.
5 Bertina RM, Koeleman RPC, Koster T, Rosendaal FR, Dirven RJ, De Ronde H, Van der Velden PA, Reitsma PH. Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 1994; **369**: 64–7.
6 Poort SR, Rosendal FR, Reitsma PH, Bertina RM. A common genetic variation in the 3′-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* 1996; **88**: 3698–703.
7 Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, Coll I, Felices R, Stone W, Fontecuberta J, Blangero J. Genetic susceptibility to thrombosis and its relationship with physiological risk factors: the GAIT study. *Am J Hum Genet* 2000; **67**: 1452–9.
8 Soria JM, Almasy L, Souto JC, Buill A, Martinez-Sanchez E, Mateo J, Borrell M, Stone W, Lathrop M, Fontecuberta J, Blangero J. A new locus on chromosome 18 that influences normal variation in activated protein C resistance phenotype and factor VIII and its relation to thrombosis. *Blood* 2003; **101**: 163–7.
9 Venter JC, Adams MD, Myers EW *et al.* The sequence of the human genome. *Science* 2001; **291**: 1304–51.
10 Gretarsdottir S, Thorleifsson G, Reynisdottir ST, Manolescu A *et al.* The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. *Nat Genet* 2003; **35**: 131–8.

P. MERLINI and D. ARDISSINO

To cite this article: Merlini P, Ardissino D. Forum on genetic studies in complex diseases. *J Thromb Haemost* 2004; **2**: 344–5.

Since the introduction of user-friendly and high-throughput techniques for performing genetic studies in complex diseases, a large number of (published and unpublished) studies have been carried out in an attempt to find polymorphisms or variants of putative genes and establish their relationships with myocardial infarction.

However, as none of the single nucleotide polymorphisms (SNPs) of genes encoding proteins involved in atherothrombosis has yet been consistently associated with an increased or decreased risk of myocardial infarction or stroke, it is not surprising that doubts are beginning to arise as to whether the main reason for this lack of success is methodological [1]. Are we using an inadequate tool? Association studies are appealingly quick and cheap, have successfully identified traditional risk factors and, Rosendaal points out [2], have been extensively used in genetic research—but extensively is not the same as successfully. On the other hand, although linkage studies have been successfully used in the study of monogenic diseases, their value in identifying new culprit genes associated with myocardial infarction has not yet been proved.

Rosendaal makes a very good point that should not be forgotten: it is not a question of deciding between association and linkage studies, but between 'proof of concept' studies and 'fishing expeditions'.

We know quite a lot about the pathophysiology of myocardial infarction, and have identified a number of factors that are almost certainly involved, such as inflammation, thrombosis and lipid metabolisms. However, the proteins involved in these mechanisms are encoded by thousands of genes, and we still do not know enough about other mechanisms involving an equally large number of genes. The result is that, although there are many pathophysiological candidates, we need to identify the most important and complete our understanding of them. Positional candidates are welcome. Linkage studies can help us to focus on the most likely regions and thus enable us to concentrate on what our knowledge indicates as being the most likely candidates; the epidemiological relevance of such candidates can then be confirmed (or not) by association studies. Linkage and association approaches are therefore not mutually exclusive but complementary, and have been recently integrated

Correspondence: D. Ardissino.
E-mail: ardis001@planet.it

with other approaches such as gene expression profiling (transcriptome analysis) and proteomics. By studying the tissues or cells (atherosclerotic plaques, circulating blood cells) directly involved in both normal and diseased states, as well as at different time points during the course of disease, we can identify new candidates deserving further investigation in functional and epidemiological studies; furthermore, the availability of new databases correlating gene function and known SNPs could lead to a major breakthrough. In our opinion, we are now beginning to come out of the fog. Rather than being simply a means of finding genes, genetic studies are much more complex and require the integration of different methodologies and approaches, in the same way that the development of a new drug does not depend on Phase III trials alone but has to go through many other stages before reaching the clinical arena.

References

1 Souto JC. Genetic studies in complex disease: the case pro linkage studies. *J Thromb Haemost* 2003; **1**: 1676–8.
2 Rosendaal FR. Genetic studies in complex disease: the case pro association studies. *J Thromb Haemost* 2003; **1**: 1679–80.