Search for new thrombosis-related genes through intermediate phenotypes. Genetic and household effects

Juan Carlos Souto

Unitat d'Hemostàsia i Trombosi
Hospital de la Santa Creu i Sant Pau

Abstract

As a complex disease, thrombosis is determined by environmental and genetic factors and by the interaction of these factors. One of the objectives of modern epidemiology is to understand the underlying complexity in complex diseases by means of disentangling the genetic from the environmental factors and quantifying the relative influence of each factor on the disease. Often it is easier and more fruitful to study intermediate phenotypes than the diseases themselves. The first step of these analyses is to determine the relative contributions of genes (heritability), measured environmental factors that are specific to an individual, and environmental factors that are shared by members of a household (household effects), to variation in the complex phenotype. Currently there are published data on 56 phenotypes involved in these kind of genetic studies of hemostasis and potentially related to thromboembolic disease. The main conclusion of these studies is that genetic effects (as measured by heritabilities) are major contributors to the phenotypic variability. Only in 16 of the 56 (29%) of the studied phenotypes, were household effects reported. These results are primordial in the design of further studies consisting in wide-genome explorations in the search for the underlying genes.

Complex disease and complex phenotypes

The majority, if not all, of common diseases are considered complex and they are determined by environmental and genetic factors and by the interaction of these factors. This means that complex diseases do not follow simple mendelian models of inheritance and are rarely determined by a single gene. Some examples of complex diseases are coronary disease, obesity, alcoholism, diabetes, cancer and mental diseases. Certainly, an important goal of modern epidemiology is to understand the underlying complexity in these disorders by means of disentangling the genetic from the environmental factors and quantifying the relative influence of each factor on the disease. Classically, each individual is usually considered to have or no to have a disease and thus the disease status is viewed as a discrete trait. Nevertheless, the modern genetic models assume that there is an unobservable continuous trait, called "liability" "risk" or "susceptibility". If an individual's liability score exceeds some specified threshold, disease results; if the score is below the threshold, the individual is unaffected. For this reason, these models are known as "threshold models".

Practically, all the biological variables that are measured on a continuous quantitative scale are complex, such as weight, height, serum cholesterol levels or plasma factor VIII levels. These variables are referred to as complex phenotypes.

Complex phenotypes as intermediate risk factors for complex diseases

Often it is easier and more fruitful to study intermediate phenotypes than the diseases themselves; for example, blood glucose levels in diabetes or coronary artery calcification in myocardial infarction. Usually, these phenotypes can be objectively and accurately measured both in symptomatic or asymptomatic individuals, providing much more information on the true risk of developing the disease. When we are interested in the genetic basis of a complex disease, a measure of the quantitative intermediate risk factors is meaningful because it is more proximal to gene action than the complex disease itself. Thus, it provides less attenuated genetic signals than when a discrete clinical endpoint such as the presence or absence of a disease is analyzed. As mentioned above, the susceptibility to disease is primarily a complex quantitative process that reflects an unobservable continuous liability. Nevertheless, this variable – liability – can not be measured directly in an individual, and consequently it is statistically less powerful than the intermediate phenotypes.

Variance component methods. Heritability and household effect

There are several relatively new methods in Genetics Epidemiology for studying complex diseases or complex phenotypes. All of them require the study of related individuals (twins, sib-pairs, sibships, nuclear families or extended pedi-

grees, in increasing order of statistical power). When studying common complex diseases (i.e those with life prevalence higher than 10%), the best results are achieved using variance component analyses in extended pedigrees. The first step of these analyses is to determine the relative contributions of genes (heritability), measured environmental factors that are specific to an individual (also called covariates, like age, sex or smoking), and environmental factors that are shared by members of a household, to variation in the complex phenotype. These household effects are due to any environmental characteristic shared by individuals living in the same household. Although, smoking, lack of exercise or common infections may be included, generally it is assumed that the main household effect contributing to the phenotypic variability is diet.

The level of the quantitative, complex phenotype, y , for individual i (y_i) is modeled as a linear function as follows:

$$y_i = \mu + \sum \beta_j x_{ij} + g_i + c_i + e_i$$

where μ is the mean of the main in the population, x_{ij} is the j -th covariate, and β_j is its regression coefficient. The remaining variables in the above formula, g_i , c_i , and e_i , represent the random deviations from μ for individual i that are attributable to additive genetic, household (common environment), and residual error effects, respectively. The residual error component includes true random error, measurement error, and any non-additive genetic components.

The effects of g , c_i , and e_i are assumed to be uncorrelated with one another and normally distributed with mean of zero and variances of σ_g^2 , σ_c^2 , and σ_e^2 . Under these assumptions, the

Table 1. Components of phenotypic variability in hemostasis-related phenotypes exhibiting a household effect

Phenotype	GAIT Project estimates			Other studies estimates		
	h2	c2	Ref.	h2	c2	Ref.
Fibrinogen	34	14	1	-	-	
F. V	44	13	1	-	-	
F. VII	52	NS	1	33	11	4
F. X	43	13	1	-	-	
F. XI	45	16	1	-	-	
TP-TM T*	41	16	3	-	-	
Activated F. XII	-	-		38	23	7
F. XIII B	48	NS	3	41	26	7
F. XIII complex	-	-		29	18	4
Whole blood folate	34	39	3	-	-	
Vitamin B12	35	13	3	-	-	
Platelet count	-	-		57	17	8
Free prot S	46	11	1	-	-	
Total prot S	22	21	1	-	-	
Functional prot S	45	9	1	-	-	
PAI-1	30	14	1	42	12	9

h2 (heritability), c2 (household effect) are given in % of phenotypic variability.

NS: not significant

*TP-TM T: thromboplastin-thrombomodulin mediated time. For details see reference 3.

PAI-1: plasminogen activator inhibitor type 1

variance of the phenotype (σ_y^2) results from the sum of these three components:

$$\sigma_y^2 = \sigma_g^2 + \sigma_c^2 + \sigma_e^2$$

Finally, the heritability (h^2) of the phenotype is the proportion of the total phenotypic variability attributable to genetic effects: $h^2 = \sigma_g^2 / \sigma_y^2$ and the household effect (c^2) is the proportion of the total variability due to shared household effects: $c^2 = \sigma_c^2 / \sigma_y^2$

Analysis of thrombosis and hemostasis phenotypes as complex phenotypes

Recently, the GAIT (Genetic Analysis of Idiopathic Thrombophilia) Project started a systematic study of the hemostasis pathways and thrombosis using the above-mentioned variance component methodology. This study includes a total of 40 phenotypes (1-3). While GAIT approach was unique, it is not the only study of the genetics of hemostasis. There are at least 26 genetic studies in the literature that use different designs (mainly twin studies) to analyze 20 of the 40 phenotypes included in GAIT, and 16 additional hemostasis-related phenotypes. Among these 26 studies, those of Freeman et al (4) and Rosendaal et al (5) are worthy of mention. These two studies were also based on the analyses of variance components in extended families and thus they had considerable statistical power and reliability.

Currently there are published data on 56 phenotypes

involved in these kind of genetic studies of hemostasis and potentially related to thromboembolic disease. The main conclusion of these studies is that genetic effects (as measured by heritabilities) are major contributors to the phenotypic variability. Only in 16 of the 56 (29%) of the studied phenotypes, were household effects reported. They account for 9% to 39% of the phenotypic variability. No significant household effects were found with the remaining 40 phenotypes. In contrast, the genetic effects seem to be the major contributors to the variability in the hemostasis-related plasma traits. Table 1 lists the estimates of heritabilities for these 16 phenotypes, which exhibited significant household effects. In all cases, except the whole blood folate, genetic effects were clearly higher than the household effects.

Interestingly, the liability to thrombosis had a heritability of 61% but did not show a significant household effect (2).

Apart from their theoretical interest, these genetic studies are very important because any phenotype that shows a significant and sizeable heritable component justifies the search for the underlying genes. Fortunately, we can learn a great deal about these complex phenotypes and diseases because of the availability of sophisticated techniques, mainly quantitative trait linkage analysis. Also, using the powerful genome-wide scan technique that searches for linkage of these phenotypes with hundreds of highly polymorphic DNA markers (microsatellites), we can anticipate a lot of important data in the field of hemostasis and thrombosis. In fact, the GAIT Project (6) has recently reported the result of a genome scan and much more important data are certain to appear and to contribute to our knowledge of hemostasis.

References

1. Souto JC, Almasy L, Borrell M, Garí M, Martínez E, Mateo J, Stone WH, Blangero J, Fontcuberta J. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation* 2000;101:1546-1551.
2. Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, Coll I, Felices R, Stone W, Fontcuberta J, Blangero J. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT Study. *Am J Hum Genet* 2000;67:1452-1459.
3. Souto JC, Almasy L, Borrell M, Stone WH, Blanco-Vaca F, Soria JM, Blangero J, Fontcuberta J. Thromboplastin-thrombomodulin-mediated time and serum folate levels are genetically correlated with the risk of thromboembolic disease: results from the GAIT Project. *Thromb Haemost* 2002;87:68-73.
4. Freeman MS, Mansfield MW, Barrett JH, Grant PJ. Genetic contribution to circulating levels of hemostatic factors in healthy families with effects of known genetic polymorphisms on heritability. *Arterioscler Thromb Vasc Biol* 2002;22:506-510.
5. Rosendaal FR, Hasstedt SJ, Bauer K, Broze GJ, Long GL, Scott BT, Callas P, Bovill EG. Heritability of clotting factors, coagulation inhibitors and activation peptides. *Blood* 2001;98:789a.
6. Soria JM, Almasy L, Souto JC, Bacq D, Buil A, Faure A, Martínez-Marchán E, Mateo J, Borrell M, Stone W, Lathrop M, Fontcuberta J, Blangero J. A quantitative-trait locus in the human factor XII influences both plasma factor XII levels and susceptibility to thrombotic disease. *Am J Hum Genet* 2002;70:567-574.
7. De Lange M, Snieder H, Ariens RAS, Spector TD, Grant PJ. The genetics of haemostasis: a twin study. *Lancet* 2001;357:101-105.
8. Garner C, Tatu T, Reittie JE, Littlewood T, Darley J, Cervino S, Farrall M, Kelly P, Spector TD, Thein SL. Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood* 2000;95:342-346.
9. Hong Y, Pedersen NL, Egberg N, de Faire U. Moderate genetic influences on plasma levels of plasminogen activator inhibitor-1 and evidence of genetic and environmental influences shared by plasminogen activator inhibitor-1, triglycerides, and body mass index. *Arterioscler Thromb Vasc Biol* 1997;17:2776-2782.