Reumatol Clin. 2010;6(1):1–2

Reumatología Clínica

www.reumatologiaclinica.org

Editorial

Systemic lupus erythematosus. "What do we know and where are we heading?"

Genética del lupus eritematoso generalizado. ¿Qué se sabe y a dónde se va?

Marta Eugenia Alarcón Riquelme,[a,b]

[a] *Department of Genetics and Pathology, University of Uppsala, Uppsala, Sweden*
[b] *Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA*

ARTICLE INFO

Available online May 28, 2009

During the past 2 years there has been an explosion in genetic findings. The main genes for several complex diseases, such as rheumatoid arthritis, Crohn's disease, etc., have been identified. Complex diseases are those in which the environment interacts in an unknown way with susceptibility genes, leading to the clinical expression of a disease. Compared to rheumatoid arthritis, there is no doubt that in systemic lupus erythematosus (SLE) the identification of susceptibility genes has been more effective. In the years prior to this scientific explosion the results were truly disappointing and geneticists did not know how to continue, but now they have unleashed their imagination.

It has been known for some time that SLE has an important genetic component: more than 8% of women with SLE have a first or second-degree relative with the disease.[1] It all began with the access, costly at first, to microarrays for the detection of bi-allele or SNP *(single nucleotide polymorphisms)* polymorphisms. The polymorphisms are the same that were previously known as restriction enzyme polymorphisms; however, the new methods for the sequencing of genomic DNA have led to the discovery of more than 3 million SNP in the human genome.[2] Initially it was possible to type approximately 10.000 SNP, but currently it is possible to type close to 1 million SNP using only 200 ng of genomic DNA from an individual.[3]

The first studies to appear were reported in diseases with a far smaller genetic component than that of SLE (Diabetes Mellitus type 1 and 2, for example), but in which the disease prevalence was higher, for example rheumatoid arthritis.[4-10] In order to detect susceptibility genes, technology knows no boundaries, leading to a bottleneck for the samples: their number and adequate clinical characterization. But, what is being searched for? If, with a determined number of samples a great number of genes have been found, don't we have it all? The main problem is that association studies using thousands of SNP tend to lead to false positive results (type I error).[11,12] Even more, 5% of all of the data that have led to a statistically significant result will be false, making it necessary to confirm these results. Therefore, even when having a complete genome association study of a 1,000 subjects and 1,000 controls and close to 20,000 significant SNP being identified, these must be confirmed with a new group of subjects; generally, it must be a number 3 times larger than the original study to consider the result as truly valid. That leads to the fact that help is needed on the part of clinical groups in sample collection from subjects. The second problem is that genes have been found in the Caucasian European population in which the disease is less severe. Some research groups have begun to study other population groups, especially individuals with a mix of European and American indigenous, African and Asian ancestry. In the case of these mixed populations, the population stratification phenomenon can be a type 1 error factor and may difficult the precise identification of the susceptibility genes. Even so, current statistical methods combined with the use of thousands of SNP may detect this stratification and correct it by excluding individuals that from a genetic standpoint do not correspond to the group under study, although their physical characteristics may make it appear so.[13] Once more, the main limit is the number of samples, because when correcting stratification in a mixed population, there remains only a portion of the individuals.

Therefore, how can recent findings in SLE be summarized?

For one, there are the genes known for SLE. All of these were discovered in studies of candidate genes, among others, *HLA DRB1*, although it is clear that we still do not know for certain if it is this gene or other genes in linkage disequilibrium with this gene within the region of the major histocompatibility complex region. It has recently been suggested that within the major histocompatibility complex region there are at least 2 SLE susceptibility genes: maybe one of them is *HLA DRB1* and the other, close to *HLA* class I.[14] The genes of the Fc fragment receptors (FR) of immunoglobulins have also been studied for some time, although data regarding them has always been controversial. Recent genome studies have confirmed this for both the major histocompatibility complex as well as the

E-mail address: marta.alarcon@genpat.uu.se

Fc receptors, especially the gene *FCGRIIA*.[15-17] This genome region is complicated because of numerous deletions as well as duplications of the genes localized within.

In addition, the interferon type I gene, *IRF5* was found in a study of candidate genes for the interferon pathway. It is possible that this gene is one of the more important ones in SLE,[18-20] after the major histocompatibility complex. *IRF5* could be the most important gene in mixed populations (north American and Latin American) in a significant manner, because the major histocompatibility complex does not seem to be strongly associated to SLE in these populations. Its origin in these populations seems to be European, later transmitted during admixture. *IRF5* has clearly been confirmed in several studies.[18-20]

Genome studies discovered several genes, among them *ITGAM*, or the alpha -M integrin (also known as Mac-1, CD11b or CR3), a well-known molecule from the SLE physiopathology standpoint, although the molecular details of the mechanisms underlying genetic susceptibility are unknown. In a very interesting way, two separate studies identified 2 genes exclusively expressed on B cells: *BLK* and *BANK1*. The *BLK* gene is a tyrosine kinase and the *BANK1* is an adaptor that mobilizes molecules within the cell to regulate intracellular signaling; in addition, these genes can explain B cell hyperactivity in SLE.[15-17]

The list of genes that have been suggested to play a role in SLE susceptibility, even if long, must still be corroborated for many of these. Meanwhile, we still have to understand how genes interact between themselves or with the environment. This, as well as the mechanisms by which 90% of SLE cases are present in women is what, is left to be discovered.

References

1. Alarcón-Segovia D, Alarcón-Riquelme ME, Cardiel MH, Caeiro F, Massardo L, Villa AR, et al. Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. Arthritis Rheum. 2005;52:1138-47.
2. International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007;449:851-61.
3. Kruglyak L. Power tools for human genetics. Nat Genet. 2005;37(12):1299-300.
4. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science. 2006;314:1461-3.
5. Evans DM, Cardon LR. Genome-wide association: A promising start to a long race. Trends Genet. 2006;22:350-4.
6. Smyth DJ, Cooper JD, Bailey R, Field S, Burren O, Smink LJ, et al. A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. Nat Genet. 2006;38:617-9.
7. The Wellcome Trust Case Control Consortium (WTCCC). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:661-78.
8. Plenge RM, Cotsapas C, Davies L, Price AL, De Bakker PI, Maller J, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet. 2007;39:1477-82.
9. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burtt NP, De Bakker PI, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science. 2007;316:1331-6.
10. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet. 2008;40:638-45.
11. De Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet. 2005;37:1217-23.
12. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005;6:95-108.
13. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38:904-9.
14. Fernando MM, Stevens CR, Sabeti PC, Walsh EC, McWhinnie AJ, Shah A, et al. Identification of two independent risk factors for lupus within the MHC in United Kingdom families. PLoS Genet. 2007;3(11):e192.
15. International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN), Harley JB, Alarcón-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet. 2008;40:204-10.
16. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med. 2008;358:900-9.
17. Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Sanchez E, et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet. 2008;40:211-6.
18. Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G, et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet. 2005; 76(3):528-37.
19. Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, Bauer JW, et al. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet. 2006;38:550-5.
20. Reddy MV, Velázquez-Cruz R, Baca V, Lima G, Granados J, Orozco L, et al. Genetic association of IRF5 with SLE in Mexicans: Higher frequency of the risk haplotype and its homozygozity than Europeans. Hum Genet. 2007;121:721-7.