Copy number variants in genetic susceptibility and severity of systemic lupus erythematosus

T. Ptacek  X. Li  J.M. Kelley  J.C. Edberg
Department of Medicine, Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, AL (USA)

Accepted in revised form for publication by H. Kehr-Sawatzki and D.N. Cooper, 22 August 2008.

Abstract. Systemic lupus erythematosus (SLE) is a systemic autoimmune disorder characterized by the presence of auto-antibodies to nuclear antigens, immune complex deposition, and subsequent tissue destruction. Early studies in twins suggested that SLE has, at least in part, a genetic basis, and a role for class II alleles in the major histocompatibility complex has been known for over 30 years. Through both linkage studies and candidate gene studies, numerous additional genetic risk factors have been identified. The recent publication of two SNP-based genome-wide association studies (GWAS) has resulted in the confirmation of a number of previously identified genetic risk loci and has identified new previously unappreciated loci conferring risk for development of SLE. A role for gene copy number variation (CNV) in SLE has also been appreciated through studies of the complement component 4 (C4) loci and more recent work in the IgG Fc receptor loci. The availability of large SNP-based GWAS datasets will undoubtedly lead to the genome-wide analysis and identification of copy number variants related to genetic susceptibility for development of SLE. We review current studies of CNV in SLE susceptibility that include reports of association between SLE and CNV in C4, IgG Fc receptors, TLR7, and CCL3L1.

Systemic lupus erythematosus (SLE) is a complex genetic disease characterized by an autoimmune response to nuclear antigens, immune complex deposition, and subsequent tissue destruction. SLE affects approximately 1 in 2,000 people with a female-to-male ratio of 9:1 (Hochberg, 1997a). The American College of Rheumatology has established eleven diagnostic criteria for SLE (four of which must be present for participation in a clinical trial) (Tan et al., 1982; Hochberg, 1997b). These criteria include presence of a malar rash, discoid rash, photosensitivity, oral ulcers, arthritis, serositis (pleuritis, pericarditis), renal disorder (persistent proteinuria, cellular casts), neurologic disorder (seizures, psychoses), hematologic disorder (hemolytic anemia, leukemia, lymphopenia, thrombocytopenia), immunologic disorder (anti-DNA, anti-Smith, anti-phospholipid antibodies present), and abnormal antinuclear antibody titer. Consequently, the SLE phenotype is extremely heterogeneous with variable phenotypes even among related individuals.

Although the etiology of SLE is still unclear, both genetic and environmental factors are involved. For example medications, ultraviolet radiation exposure, and infections have all been associated with SLE (Sarzi-Puttini et al., 2005). Candidate gene searches for SLE susceptibility have led to multiple findings in various ethnic groups. Strong and consistent association has been observed with class II alleles of the major histocompatibility complex (MHC). In addition, consistent association of non-MHC variants, some of which are functionally relevant, have been reported (Croker and Kimberly, 2005; Gregersen and Behrens, 2006; Brown et al., 2007). Recent genome-wide association studies (GWAS) have confirmed some previously associated loci and revealed new genetic regions related to the disease. A GWAS, conducted by SLEGEN (International Consortium for SLE Genetics), reported association with multiple sites of the MHC and with polymorphisms of interferon regulatory factor 5.
(IRF5), integrin alpha M (ITGAM), KIAA1542 (PHRF1), the PX domain containing serine/threonine kinase (PTKX), an individual marker on chromosome 1q25, an IgG Fc receptor (FCGR2A), protein tyrosine phosphatase non-receptor type 22 (PTPN22), and signal transducer and activator of transcription 4 (STAT4) (Harley et al., 2008). Another GWAS, conducted by Genentech also reported strong associations in the MHC cluster and the ITGAM-ITGAX region in addition to B lymphoid tyrosine kinase (BLK) (Hom et al., 2008).

One topic currently under intense investigation is the role of copy number variations (CNVs) in the genetic contribution to complex diseases. Copy number polymorphisms arise when an entire gene or gene segment has been duplicated or when a gene is absent in some individuals (Eichler et al., 2007). Extra copies of genes may contribute to overexpression of proteins, and deletion of a gene may lead to a deficiency and functional consequences. Entire gene duplications allow new genes with new functions to evolve while keeping a functional backup copy of the original, ancestral gene (Ohno, 1970).

**Copy number variation in SLE**

Although the broad significance of CNV in disease has only recently been explored in detail, examples of CNVs associated with SLE have been known for more than a decade. Originally, these CNVs were detected as deficiencies of certain proteins, but recent work has shown that the genes coding for these proteins can have a copy number (CN) >2 (duplication), in addition to a CN <2 (deletion or deficiency) (Table 1).

CNV has also been associated with autoimmune phenotypes in rodent models that are similar to and used as models of human SLE. Association of the genes that vary in CN with the autoimmune phenotype in murine models is not always recapitulated in human SLE. However, these genes may offer clues to common mechanisms in human SLE and rodent autoimmune phenotypes.

**Complement component 4 (C4)**

The complement system consists of more than 30 plasma and cell-surface proteins and has many roles in the innate immune system, including general defense against infection, clearance of immune complexes and apoptotic debris, and interfacing with the adaptive immune system (Walport, 2001a). Complement's role in clearance of immune complexes and apoptotic debris is thought to be important in the development of or protection against SLE. Genetic deficiency of complement proteins has been known to predispose to development of SLE (Walport, 2001b; Cook and Botto, 2006). For example, homozygous deletions of the early complement components C1, C2, and C3 are strongly associated with development of an SLE-like phenotype (Walport, 2001b). Furthermore, polymorphisms of several genes for complement proteins associated with SLE have already been reported (Walport, 2006).

Aside from homozygous deletions, copy number variation in the C4 locus, a key component in the classical pathway of the complement system, has been extensively studied in the context of SLE. C4 is involved in clearing of immune complexes, apoptotic cells, and infectious agents (Walport, 2001a). Most people have two C4 genes that code for two variants of C4 (C4A and C4B), both of which are found in the class III region of the MHC on human chromosome 6 (Yang et al., 1999). Deficiency of C4 has been reported in SLE as early as the 1970s (Hauptmann et al., 1974; Schaller et al., 1977; Clark and Klebanoff, 1978). The C4 genes have been shown to vary in CN via duplications and deletions of multi-gene cassettes (including RP1), the C4 genes, a CYP21 pseudogene, and TNX) referred to as RCCX modules (Yang et al., 1999). A recent paper analyzing CNVs of RCCX modules demonstrated that low CN predisposed toward SLE while high CN protected against SLE in a population of European Americans (Yang et al., 2007). It is thought that increased CN of the C4 genes protects against SLE by increasing the clearance of immune complexes and apoptotic cell debris, while decreased CN of the C4 genes is a risk factor for SLE due to decreases in the same processes.

**Fcγ Receptor 3B (FCGR3B) and other Fcγ receptors**

Fcγ receptors (FCGRs) are receptors that bind the Fc domain of IgG antibodies, serving as a link between the adaptive and innate immune response. FCGRs regulate immune responses via tyrosine phosphorylation of their activating or inhibitory cytoplasmic domains or of associated accessory or signaling chains (Nimmerjahn and Ravetch, 2007). Genetic variants of Fc receptor genes, especially those in the classical low affinity Fc receptor cluster on human chromosome 1q23 (Fig. 1), are known to have a role in both susceptibility and severity of SLE (Croker and Kimberly, 2005; Alarcon et al., 2006; Brown et al., 2007).

### Table 1. Known CNVs affecting SLE in humans or similar autoimmune phenotypes in rodent models. CN: copy number.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene(s)</th>
<th>Species</th>
<th>Chromosome</th>
<th>CNV</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complement component 4</td>
<td>C4A, C4B</td>
<td>Human</td>
<td>6p21</td>
<td>0 to &gt;4</td>
<td>Deficiency: high risk for SLE; low CN: risk factor for SLE; high CN: protective against SLE</td>
</tr>
<tr>
<td>Toll-like receptor 7</td>
<td>Tlr7</td>
<td>Mouse</td>
<td>XqF5</td>
<td>&gt;2</td>
<td>High CN: increased susceptibility to autoimmune phenotype</td>
</tr>
<tr>
<td>Fcγ receptor 3B</td>
<td>FCGR3B</td>
<td>Human</td>
<td>1q23</td>
<td>0 to &gt;3</td>
<td>High CN: protective against SLE; low CN: risk factor</td>
</tr>
<tr>
<td>Fcγ receptor 3</td>
<td>Fcgr3</td>
<td>Rat</td>
<td>13q24</td>
<td>0 to &gt;3</td>
<td>Low CN: risk factor for crescentic glomerulonephritis</td>
</tr>
</tbody>
</table>

Fig. 1. The human 1q23 classical FCGR gene cluster. Shaded boxes indicate the relative locations and sizes of genes in the cluster, and arrows indicate the direction of transcription. This figure was based on output from the UCSC genome browser (Kent et al., 2002).

**FCGR3B** codes for the neutrophil-specific, low-affinity FCGR3 that is glycosylphosphatidyl-inositol (GPI) linked to the membrane (Nimmerjahn and Ravetch, 2007). Although it is GPI linked, unlike the other FCGRs which are transmembrane proteins, FCGR3B is a functional regulator of neutrophil activation (Huizinga et al., 1990a; Hundt and Schmidt, 1992; Edberg and Kimberly, 1994; Kocher et al., 1997). Functional alleles of **FCGR3B** alter the binding of IgG that subsequently results in different quantitative levels of neutrophil activation in donors homozygous for neutrophil antigen 1 (NA1) versus NA2 (Croker and Kimberly, 2005; Brown et al., 2007). Evidence of increased copy number of **FCGR3B** was originally suggested through the detection of three allele patterns of the receptor in three individuals (NA1, NA2, and SH, now known as HNA-1a, HNA-1b, and HNA-1c, respectively) (Koene et al., 1998). NA-null individuals have been known for many years, and clinically it can result in isoinnate neutropenia in children born to NA-null mothers (Schnell et al., 1989). The NA1 and NA2 allotypes have been localized to FcγRIIIb on human neutrophils (Stroncek, 2002; Bux, 2008). Molecular and genetic characterization of individuals homozygous for **FCGR3B** deficiency was first reported in 1990 in a healthy individual and in a patient with SLE (Clark et al., 1990; Huizinga et al., 1990b). A high copy number of **FCGR3B** has more recently been reported to be protective against glomerulonephritis and lupus nephritis in humans, while low CN was a risk factor (Aitman et al., 2006). **FCGR3B** CNV has also been found to be similarly associated with SLE (but not organ specific lupus) and other autoimmune disorders (Fanciulli et al., 2007). The frequencies of the three copy number variant **FCGR3B** alleles are different across ethnic groups (Kissel et al., 2000), and these differences may contribute to ethnic and geographical differences in SLE (Lau et al., 2006).

While there is no rodent homologue of **FCGR3B**, rodent Fgr3 may be considered to be a functional homologue of the human gene. Fgr3 is expressed in neutrophils and participates in neutrophil activation (Coxon et al., 2001), and CNV of **Fgr3** has been reported to be associated with risk for glomerulonephritis in rats (Aitman et al., 2006).

CNV of **FCGR2C** has been reported to be linked to CNV of **FCGR3B** (de Haas et al., 1995) and to predispose to idiopathic thrombocytopenic purpura (Breunis et al., 2008). However, CNV of the other **FCGR** genes in SLE remains unexplored. Further characterization of the **FCGR** gene cluster on human chromosome 1q23 and their relationship to SLE promises to be a fruitful area for future studies.

**The Yaa locus and toll-like receptor 7 (TLR7)**

The BSXB strain of mice spontaneously develops an autoimmune phenotype similar to human SLE, which includes production of anti-nuclear antibodies. Unlike human SLE, the autoimmune phenotype was enhanced in males rather than females. This observation led to the discovery of a Y-linked autoimmune accelerator locus, referred to as **Yaa** which increases the severity of the autoimmune phenotype in this mouse model (Murphy and Roths, 1979; Izui et al., 1995).

While the identity of the **Yaa** gene is not known, CNVs of **Yaa** were recently reported and duplications of **Yaa** were shown to further increase susceptibility to autoimmunity in mice (Pisitkun et al., 2006). The authors attempted to identify the causative gene by quantifying mRNA and genomic DNA within the **Yaa** locus. Of the genes with amplified mRNA and genomic DNA, Pisitkun et al. (2006) proposed **Tlr7** as the causative gene because of its biological properties including the binding of RNA. B-cell hyperactivity has been previously reported in **Yaa** mice. The authors also demonstrated that **Tlr7** interacts with Bruton’s tyrosine kinase (Btk) and that Btk signaling is necessary for the autoimmune response to nucleolar antigens in **Yaa** mice. **Tlr7** is part of a family of the toll-like receptors that recognize nucleic acids and nucleic acid-like structures (Takeda et al., 2003). It is possible that **Tlr7** has some reactivity to nuclear material from apoptotic debris and that increased CN of **Tlr7** would cause increased autoimmune response to nuclear material.

CNVs of the human **TLR7** gene have recently been examined in SLE, but no association between **TLR7** CN and the SLE phenotype was observed (Kelley et al., 2007). It is possible that **Tlr7** is not the autoimmunity-predisposing gene in the **Yaa** locus. The **Tlr8** gene is contiguous to **Tlr7** and is therefore within the same interval identified by Pisitkun et al. (2006). It is also possible that **Tlr7** is the causative gene in the **Yaa** locus and that this is an example of a disparity between the genetics of human SLE and murine models of autoimmunity (Mestas and Hughes, 2004; Vyse, 2008).

**Other CNV loci in SLE**

It is likely that CNVs in other loci will be important contributors to both susceptibility to and severity of SLE and other autoimmune diseases. Indeed, recent studies in Crohn’s disease, an autoimmune inflammatory disease of the gastrointestinal tract, have demonstrated a role for CNV...
in the β-defensin gene cluster on human chromosome 8 (Fellermann et al., 2006). The defensins are endogenous antibiotic and cytotoxic peptides that form a chemical barrier at the epithelial surface, and their relative deficiency may lead to bacterial adherence to the mucosa, slow invasion, and secondary mucosal inflammation. The chromosomal region containing the defensin genes is characterized by segmental duplication, and there is extensive CNV in the defensin genes. Patients with Crohn’s disease are characterized as having lower copy numbers of the defensin genes. Patients with Crohn’s disease are characterized as having lower copy numbers of the defensin genes suggesting a functional role for these genes in the development of the disease (Fellermann et al., 2006). In SLE, a recent report has suggested that CNV in the chemokine (C-C motif) ligand 3-like 1 (CCL3L1) locus is associated with risk of disease development (Mamtani et al., 2008). In summary, there are already tantalizing clues that CNVs will be important in understanding the genetic basis of SLE and other autoimmune phenotypes.

Current genome-wide association studies and CNVs

Although impressive progress has been made in defining candidate CNVs associated with SLE, there are still many open questions and technical challenges. At present, given the availability of two independent GWAS in SLE using single nucleotide polymorphisms (SNPs), the use of the GWAS data for analysis of CNV is most promising. CNVs can be detected through either direct scoring or through identification of SNP markers known to be in linkage disequilibrium with CNVs. While there are certainly many methodological issues to be overcome in doing such analyses, these data would permit the first genome-wide assessment of CNVs in SLE. Early successes in using SNP-based arrays for assessment of CNVs in complex genetic phenotypes include recent reports in amyotrophic lateral sclerosis (ALS) (Blauw et al., 2008; Cronin et al., 2008), schizophrenia (Stefansson et al., 2008), ischemic stroke (Matarin et al., 2008), and autism (Marshall et al., 2008). While the currently available GWAS in SLE have not yet reported CNV data, it is likely that such analyses will be completed soon and will yield additional insights into the role of genetic variants in the development and prognosis of SLE.

Significance of CNVs in basic and clinical SLE studies

There are a number of different mechanisms by which CNVs could be causative variants in disease susceptibility and severity (Beckmann et al., 2007). CNVs might have functional consequences through direct structural disrupting/alteration of gene(s). Certainly, CNV would be expected to alter gene dosage, and this might amplify the impact of SNP variants that may or may not occur in regions of CNV. For example, it is well established that CCL3L1 gene dose (i.e., CNV) and CCR5 genotype combinations have a substantial impact on cellular immunity, HIV-AIDS pathogenesis, and responsiveness to antiretroviral therapy in patients with HIV-AIDS (Dolan et al., 2007; Ahuja et al., 2008). In SLE, similar interactions between CCL3L1 CNV and CCR5 genotypes have been suggested (Mamtani et al., 2008). It is also interesting to note that CNVs in CCL3L1 may also be associated with other autoimmune diseases such as rheumatoid arthritis (McKinney et al., 2008).

Another area of potential interest for the role of CNVs is drug response. As suggested by the example of CCL3L1 and antiretroviral treatment in HIV-AIDS (Gonzalez et al., 2005; Ahuja et al., 2008), it is likely that there are CNVs that will influence treatment of SLE and other autoimmune diseases. One provocative example could be the use of intravenous immunoglobulins (IVIGs) as a treatment for immune deficiency, autoimmune, and inflammatory disorders (Negi et al., 2007). Aside from the expense, a significant problem in its clinical use is the variability in its therapeutic efficacy in different individuals. Mechanisms of IVIG’s immunoregulatory effects are complicated and still remain unclear, although recent studies suggest a significant role for IgG Fc receptors. Both FCGR2B and FCGR3A have been proposed as key factors in the mechanism of action of IVIG’s anti-inflammatory properties (Samuelsson et al., 2001; Siragam et al., 2006; Park-Min et al., 2007). In this context, recently identified CNVs in FCGR3A and FCGR2C (Breunis et al., 2008) and potential CNVs in other FCGR genes may result in different levels of FcGR protein expression in different individuals, thus enhancing or reducing regulatory functions of IVIG.

Conclusion

Demonstration of the importance of CNVs in SLE will undoubtedly yield new insights into the pathogenic mechanisms in SLE and other complex autoimmune diseases. Although remarkable progress has been made in both initial GWASs and target gene studies in SLE, there are still significant gaps in our understanding of the genetic basis of the disease. Further, more dense screens in multiple ethnic groups will be a necessary follow-up to the initial GWAS successes in SLE. It is still not known if there will be common CNV variants that associate with disease and/or whether there will be less common/rare variants that are found in affected individuals as observed in the recent genome-wide CNV analysis in ALS (Blauw et al., 2008). Current and future CNV candidate regions/gene will require methodological development for independent replication studies. Finally, integrating our knowledge about SNPs and CNVs will be required to gain a complete understanding of the genetic basis of SLE and related autoimmune diseases.

Acknowledgements

We thank Dr. Robert P. Kimberly for his continued support.
References


