ARTICLE IN PRESS

Molecular Immunology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

Common and rare genetic variants of complement components in human disease

Elena Goicoechea de Jorge[a,b,1], Alberto López Lera[c,d,1], Rafael Bayarri-Olmos[e,1], Hugo Yebenes[d,g], Margarita Lopez-Trascasa[f], Santiago Rodríguez de Córdoba[d,g,*]

[a] Department of Immunology, Complutense University, Madrid, Spain
[b] Research Institute Hospital 12 de Octubre (imas12), Madrid, Spain
[c] Research Institute Hospital Universitario La Paz (IdiPaz), Madrid, Spain
[d] Ciber de Enfermedades Raras, Madrid, Spain
[e] Department of Clinical Immunology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
[f] Departament of Medicine, Universidad Autonoma de Madrid, Madrid, Spain
[g] Molecular Pathology and Complement Genetics Unit. Centro de Investigaciones Biológicas, CSIC, Madrid, Spain

ARTICLE INFO

Keywords:
Complement
Genetic variants
Disease associations
Polymorphisms

ABSTRACT

Genetic variability in the complement system and its association with disease has been known for more than 50 years, but only during the last decade have we begun to understand how this complement genetic variability contributes to the development of diseases. A number of reports have described important genotype-phenotype correlations that associate particular diseases with genetic variants altering specific aspects of the activation and regulation of the complement system. The detailed functional characterization of some of these genetic variants provided key insights into the pathogenic mechanisms underlying these pathologies, which is facilitating the design of specific anti-complement therapies. Importantly, these analyses have sometimes revealed unknown features of the complement proteins. As a whole, these advances have delineated the functional implications of genetic variability in the complement system, which supports the implementation of a precision medicine approach based on the complement genetic makeup of the patients. Here we provide an overview of rare complement variants and common polymorphisms associated with disease and discuss what we have learned from them.

1. Introduction

The complement system is a key element of innate immunity. It can trigger alarm signals warning about the presence of pathogens, immune complexes or cell remains. Complement discriminates between our own components and pathogenic ones, tagging the latter for elimination by the cells of the phagocyte system, or for direct destruction through cell lysis. Briefly, the actions of complement include the following processes: i) Complement activation by three independent activation pathways, the classical (CP), lectin (LP) and alternative (AP) pathways, which results in the formation of unstable protease complexes, named C3-convertases (AP, C3bBb; CP/LP, C4b2a) that cleave C3 to generate C3b; ii) Convertase-generated C3b can form more AP C3-convertase, providing exponential amplification of the initial activation; and iii) Clustering of C3b molecules around the surface-bound C3-convertase generates the C5-convertase with the capacity to bind and cleave C5, triggering inflammation and initiating formation of the lytic membrane attack complex (MAC). Several regulatory proteins restrict complement activation to the surface responsible for its activation and restrict consumption after activation. The loss of complement regulation leads to the generation of pro-inflammatory components and/or tissue damage. Both situations have pathologic consequences (Ricklin et al., 2010; de Cordoba et al., 2012; Holers, 2008; Ricklin and Lambris, 2013; Sjoberg et al., 2009). In this review, we will describe several common and rare complement variants associated with disease to illustrate how their identification and, eventually, their structural and functional study, has been instrumental to understand the roles of complement in disease.

* Corresponding author at: Centro de Investigaciones Biológicas, Ramiro de Maeztu 9, 28040, Madrid, Spain.
E-mail address: srdecordoba@cib.csic.es (S. Rodríguez de Córdoba).
[1] Equal contribution authors.

https://doi.org/10.1016/j.molimm.2018.06.011
Received 10 April 2018; Received in revised form 6 June 2018; Accepted 8 June 2018

2. Genetic variability in the components and regulators of the Classical Pathway (CP)

The CP is triggered by antibody-antigen immune complexes via C1q recognition of Fc domains in IgM or IgG. Activation of the CP can also occur independently of antibodies, as C1q can bind directly to certain microbial epitopes, apoptotic or necrotic cells and pattern recognition molecules. Independently of the trigger, CP activation tags the activating surfaces for their removal by phagocytosis and CP progression ultimately leads to the destruction of target cells by the assembly of the MAC.

2.1. The C1 complex

The first component of the CP (C1) is a calcium-dependent complex involving three proteins: C1q, C1r and C1s. Under physiologic conditions, the C1 heterotrimer comprises two weakly interacting subunits: C1q, an opsonin containing the binding sites for CP activators, and $C1r_2s_2$, a tetrameric C1s-C1r-C1r-C1s complex with enzymatic activity. Genetic alterations causing deficiency of the C1 complex are strongly associated with the development of autoimmune disease, bacterial infections and the development of glomerulonephritis. The association between C1 complex deficiencies and autoimmunity underlines the fundamental role that the CP has in the elimination of cellular debris; the inability to efficiently clear apoptotic and necrotic cells renders them a source of autoantigens and drives autoantibody production (Navratil et al., 2001).

2.2. C1q

C1q is a 460 kDa hexameric plasma glycoprotein comprised of 18 peptide chains (6 A, 6B and 6C) arranged in a “bouquet of tulips” disposition. It is a pattern recognition molecule capable of binding to a variety of ligands on microbial surfaces, apoptotic or dying cells and to immune complexes (Kishore and Reid, 2000). Interaction between C1q and its receptors gC1qR, calreticulin-CD91 and integrin $\alpha_2\beta_1$, allows clearance of apoptotic cellular debris and triggers subsequent B cell tolerance, macrophage polarization and monocyte-dendritic cell development (Hosszu et al., 2010; Bohlson et al., 2014). Notably, a number of functions of C1q do not involve classical pathway activation, suggesting that C1q plays additional roles in homeostasis like induction of apoptosis in prostate cancer cells overexpressing WOX1 or induction of angiogenesis during wound healing (Kouser et al., 2015; Hong et al., 2009; Bossi et al., 2014). C1q has also been implicated in synaptic pruning in the developing CNS, trophoblast invasion, vascular remodelling and normal placentation during pregnancy (Stevens et al., 2007; Singh et al., 2011).

Approximately 70 cases of congenital C1q deficiency, either antigenic or functional, have been reported associated with rare variants in the C1Q genes (Roumenina et al., 2011). These C1q deficiencies associate with chronic infections and increased risk of autoimmune pathology, particularly SLE, SLE-like disease and glomerulonephritis (Macedo and Isaac, 2016; Stegert et al., 2015; Botto et al., 1998). In addition, several C1q polymorphisms have been associated with rheumatoid arthritis (RA), cancer and metabolic or neurolopsychiatric disorders like type 2 diabetes or schizophrenia (Abecasis et al., 2004; Dardiotis et al., 2009; Goulielmos et al., 2013; Mosaad et al., 2015; Petry and Loos, 2005; Racila et al., 2006; Trouw et al., 2013; Xu et al., 2009; Zakharyan et al., 2011; Zervou et al., 2011).

2.3. C1r and C1s

C1r and C1s are the enzymatic components of the C1 complex. They circulate in plasma as zymogens in an extended tetrameric Ca^{2+}-dependent complex ($C1r_2s_2$; C1s-C1r-C1r-C1s). Upon binding to the collagen-like domains of C1q, the $C1s_2r_2$ complex undergoes a conformational change that triggers C1r autoactivation, which in turn converts the C1s zymogen into an active serine protease (Venkatraman Girija et al., 2013; Gaboriaud et al., 2014; Arlaud et al., 2002).

Thirty cases with complete C1r/C1s deficiencies have been described to date. Like C1q deficiencies, they associate with increased risk to autoimmune disease, bacterial infections and glomerulonephritis (Grumach and Kirschfink, 2014). Notably, heterozygous gain-of-function mutations in C1r or C1s cause type I periodontal Ehlers-Danlos syndrome (P-EDS), an autosomal dominant disease affecting connective tissue integrity. The rare C1r and C1s variants associated with P-EDS involve the subunit interfaces or inter-domain hinges of C1r and C1s, which may increase C1r autoactivation and lead to CP activation on off-target substrates (Kapferer-Seebacher et al., 2016). Among the disease-associated common C1s variants, there is a polymorphism (rs7311672) that has been reported to confer increased risk for Alzheimer disease (AD) (Jones et al., 2010).

2.4. C1-Inhibitor

C1-Inhibitor (C1INH) is a single chain, highly glycosylated 105 kDa protein containing a long, unique N-terminal domain with no homology in the human genome and a C-terminal SERPIN domain. C1INH irreversibly inhibits proteases from the CP and LP, coagulation, fibrinolysis and contact cascades by forming SDS-stable serpin-protease complexes. In the context of CP regulation, C1INH inhibits the activated C1s, C1r and C2 proteases. C1INH deficiency causes autosomal dominant Hereditary Angioedema (types I and II), a rare disease characterized by spontaneous episodes of edema in the subcutaneous and submucosal layers (Cicardi et al., 2014).

2.5. C4 and C2

C4 is synthesised as a 200 kDa β1-globulin and intracellularly processed into three polypeptide chains (α, β, and γ) with molecular masses of 93 kDa, 78 kDa, and 33 kDa, respectively (Kidmose et al., 2012). C4 exists as two highly homologous isotypes (C4A, for “acidic” and C4B, for “basic”) encoded by two tandem-oriented genes on chromosome 6 (6p.21) and carrying the antigenic determinants corresponding to the Rodgers and Chido blood groups, respectively (Truedsson, 2015). Upon activation by C1s or MASP2, the α-chain of C4 is cleaved to yield C4a and C4b fragments. The nascent C4b fragments become covalently linked to target surfaces through their reactive thioester sites. The C4A and C4B paralogs share 99% sequence identities but exhibit significant functional differences: C4A most efficiently forms a covalent amide bond with amino group-containing antigens, whereas C4B tends to a form a covalent ester bond with hydroxyl group-containing substrates. Such distinctive functionality between C4A and C4B results in different haemolytic activities, covalent affinities to antigens and serological reactivities (Law et al., 1984; Blanchong et al., 2001).

C2, a 83 kDa plasma glycoprotein that circulates in plasma as a zymogen, is organized as a N-terminal portion containing three short consensus repeat (SCR) domains (C2b) and a portion containing a von Willebrand Factor type A domain (vWFA) linked to a C-terminal trypsin-type serine protease domain (C2a) (Mortensen and Jensen, 2016). Upon binding to C4b, C2 is cleaved by activated C1, which releases the C2b fragment. C2a remains bound to C4b as an active serine protease to form the CP C3/C5 convertases C4b2a and C4b2a3b, respectively.

Homozygous C2 deficiency is relatively common, with an estimated prevalence of 1:10.000 in Caucasians, over 90% of them carrying a common 28-base pair deletion (28BP-Del) (Truedsson, 2015). Like the C1 complex deficiencies, the complete absence of C2 is associated with SLE or SLE-like disease and, to a lower extent, with bacterial infections, particularly invasive infection by encapsulated bacteria. Polymorphisms in the C2 locus are associated with AMD. Initial analyses indicated