Chapter from the book *Analysis of Genetic Variation in Animals*
Downloaded from: http://www.intechopen.com/books/analysis-of-genetic-variation-in-animals
1. Introduction

Effective control of pathogens is typically achieved by the timely interplay between the innate and adaptive immune response of an individual. The first-line of defence of the host immune system is our skin and mucosal membranes, which form a physical and chemical barrier. If a pathogen gains access to the body, for example when there is skin damage, cells of our innate immune system recognise pathogens by means of receptors (acting as sentinels) binding common motifs on pathogen surfaces and mount an immediate non-specific effector response. Often this response leads to the destruction of the pathogens, but also the production of secreted proteins such as interferons (IFNs) or cytokines, which help the interaction between the various cells involved in the immune response. This initial “inflammation” of the infected tissue commonly leads to the recruitment of lymphocytes to the site and triggers a more pathogen-specific response known as the adaptive immune response. In addition, local inflammation triggers activation of the complement system, a group of plasma proteins, which can recruit further immune cells to the site or be directly proteolytic to bacterial surfaces (Janeway 2008).

Not surprisingly, there are multiple genes that code for the molecules of the innate and adaptive host immune response. Between individuals, these genes can show variations within the coding sequence itself and/or within associated regulatory elements such as promoters. Furthermore, many of these genes are part of multicycopy gene families resulting from previous gene duplication events (Kelley and Trowsdale 2005). Gene duplication can lead to new gene copies developing complementary and/or overlapping functions allowing for redundancy in the immune system and the ability to cope with a number of different pathogens (Nei, Gu et al. 1997).

Some of the genetic characteristics of the immune system predate mammalian divergence while others reflect more recent primate and human evolution. However, all living species, including humans, have been under constant bombardment from pathogens over thousands...
or millions of generations and it is likely that this host/pathogen interaction has to a large extent shaped the genetics of both host and pathogen populations.

Remarkably, for some highly mutable and therefore rapidly changing pathogens such as HIV and the Hepatitis C virus (HCV), two small RNA viruses, viral adaptation to the host’s immune response results in characteristic variations in the virus that allow the virus to effectively escape a particular immune response. In contrast, larger, complex DNA viruses, such as Herpes viruses, with a more rigid genome have the coding capacity for additional genes, which interfere specifically with the immune response of the host (Lucas, Karrer et al. 2001). The multitude of escape strategies presented to our immune system, highlights the diversity of modifications needed to successfully combat the pathogens of our modern times. Furthermore, the result of these selective forces on the host over thousands of generations is the extensive host genetic diversity of immune response genes we observe today in all populations.

2. Examples of genetic diversity of the immune response genes

Two genetic systems that are commonly associated with variable infection outcomes are the killer-immunoglobulin like receptors (KIRs) present on NK cells that can determine the activatory or inhibitory propensity of the NK cell as part of the innate immune response and human leucocyte antigens (HLA) that present processed antigens to T lymphocytes, which are part of the adaptive immune response. These two genetic systems are also functionally connected; HLA molecules are ligands for the KIRs and their particular combinations add another layer of complexity to the differential immune response exhibited by individuals to particular pathogens.

2.1 KIR genetic diversity

The region of chromosome 19q13.4 containing the KIR genes is characterised by polymorphism, gene duplication, and linkage disequilibrium leading to KIR haplotypes spanning several hundred kilobases (Martin, Kulski et al. 2004). Genetic variation between haplotypes presents as allelic variation at specific loci and variation in gene content (Martin, Kulski et al. 2004; Pyo, Guethlein et al. 2010).

The KIRs can be separated into KIR2D and KIR3D genes based on the number of extracellular immunoglobulin domains. Further differentiation occurs based on the length of their cytoplasmic domain (short – S and long – L). Those KIR genes with a short cytoplasmic tail that lack an ITIM motif confer an activatory phenotype to the KIR2D and 3D genes while longer cytoplasmic domains with an ITIM motif confers an inhibitory phenotype. Sequencing of the KIR region and of individual genes shows extensive allelic diversity for both activatory and inhibitory genes (Table 1).

The effects of KIRs on host immune responses are mediated by specific cognate interactions between these receptors and their natural HLA Class I ligands on the surface of target cells, which may result in either activation or inhibition of NK cell cytotoxicity (Lanier 1998). The presence of particular HLA-B/-C alleles will determine the repertoire of inhibitory and activatory ligands that can be utilised by KIRs to modulate NK cell mediated responses. Several KIR and HLA combinations have been associated with both infectious disease outcome (Khakoo, Thio et al. 2004) and autoimmune diseases (Martin, Nelson et al. 2002).
Genetic Variation of Host Immune Response Genes 
and Their Effect on Hepatitis C Infection and Treatment Outcome

Functional model for the interaction between KIR and their HLA ligands has been proposed to account for a suspected hierarchy of inhibitory NK cell responses that may reflect the disease associations (Kulkarni, Martin et al. 2008). In this model, individuals with KIR haplotypes carrying a greater number of activatory genes, and with the appropriate HLA Class I ligands, may exhibit lower inhibition and higher activation (activatory phenotype), which may be beneficial in viral infections but may also correlate to a greater risk of developing autoimmunity.

2.2 HLA genetic diversity

The HLA genetic system within the Major Histocompatibility Complex (MHC) on the short arm of chromosome 6 is the most polymorphic in humans. Similar to the KIR region, the MHC region is characterised by extensive intra- and inter-genic polymorphism (Gaudieri, Dawkins et al. 2000), segment duplications (containing more than a single gene or gene fragment) (Gaudieri, Kulski et al. 1999) and linkage disequilibrium leading to haplotypes containing HLA and other immune-related genes that stretch hundreds of kilobases and even megabases. One major difference between KIR and MHC haplotypes is that within the MHC, for the most part, all individuals have the same number of HLA Class I and II genes but heterozygosity values at each locus is high.

The HLA family comprises class I and class II molecules that can present processed antigen to CD8+ and CD4+ T lymphocytes, respectively. The HLA Class I molecules including HLA-A, -B and –C are expressed on all nucleated cells and share significant similarity reflecting largely overlapping functions but with some important distinctions, including their function as NK cell receptor ligands. The HLA Class II molecules including HLA-DR, -DQ and –DP are expressed on antigen-presenting cells. Orthologues of the HLA Class I and II genes are present in other primates and most vertebrates.

HLA molecules were initially identified via their involvement in self-nonself discrimination in transplantation. Serology-based assays using anti-HLA antibodies were used to match donor and recipient pairs prior to transplantation. These initial HLA typing assays were able to differentiate dozens of HLA Class I and II proteins (Graw, Goldstein et al. 1970). However, it was not until sequence-based typing was developed in the 1990s that the extent of genetic variation within these genes became apparent. Several hundred to thousands of alleles have now been described for many of the HLA Class I and II loci (Table 1).

Much of the variation within HLA genes exists in the peptide binding domains of the molecule. Individuals with different HLA types can therefore present different parts of the pathogen to T lymphocytes and it is thought that the variation observed within the HLA peptide binding domains are driven by positive selection pressures that favour the maintenance of polymorphism in this system (Hughes, Ota et al. 1990). Support for this is provided by several studies that show evidence for heterozygote advantage at the HLA loci following HIV and HCV infection (Carrington, Nelson et al. 1999; Hraber, Kuiken et al. 2007); with heterozygotes presenting a greater variety of peptides to T lymphocytes. Furthermore, studies examining pathogen-load in different human populations suggest pathogen-selection has been one of the driving forces in shaping the extensive variation we observe today in the HLA system (Prugnolle, Manica et al. 2005). However, variation outside the peptide binding domain can also be important as has been shown for HIV (Fellay, Shianna et al. 2007). In some
cases, these variations may alter the effect of microRNAs on transcribed species and ultimately the expression level of HLA on the cell surface (Kulkarni, Savan et al. 2011).

<table>
<thead>
<tr>
<th>Family</th>
<th>Gene</th>
<th>Alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA Class I</td>
<td>A</td>
<td>1,698</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2,271</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>884</td>
</tr>
<tr>
<td>HLA Class II</td>
<td>DRB1</td>
<td>975</td>
</tr>
<tr>
<td></td>
<td>DQB1</td>
<td>158</td>
</tr>
<tr>
<td></td>
<td>DPB1</td>
<td>149</td>
</tr>
<tr>
<td>MIC</td>
<td>MICA</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>MICB</td>
<td>33</td>
</tr>
<tr>
<td>KIR2D</td>
<td>2DL1</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>2DL2</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>2DL3</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>2DL4</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>2DL5</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>2DS1</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2DS2</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>2DS3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2DS4</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>2DS5</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2DP1</td>
<td>22</td>
</tr>
<tr>
<td>KIR3D</td>
<td>3DL1</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>3DL2</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>3DL3</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>3DS1</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>3DP1</td>
<td>23</td>
</tr>
</tbody>
</table>

*Data from International Immunogenetics Project (www.ebi.ac.uk) release date April 2011 for KIR and July 2011 for MIC and HLA.

Table 1. Number of alleles of variable host immune-related genes

Given the central role of HLA in the adaptive and innate immune response, it is not surprising that the outcome of hundreds of diseases have been associated with certain HLA alleles and more generally with genes within the MHC region. However, the intrinsic properties of the region, particularly the extensive linkage disequilibrium, have made the delineation of specific disease genes or alleles difficult (Dawkins, Leelayuwat et al. 1999).

2.3 Diversity of other genes involved in host immune response

Other genes within the MHC that also exhibit allelic variation include the MIC genes (MICA and MICB) that are ligands for the activatory NK cell receptor NKG2D. This gene family also has multiple gene copies in the genome with extensive diversity (Table 1). Interestingly, the HLA-B and -C genes and MICA (and to a lesser extent MICB) are within a region of
approximately 200kb in the MHC and are in strong linkage disequilibrium such that specific combinations of alleles for these genes are commonly inherited together. Again, this represents a problem when geneticists try to disentangle disease associations with these genes. On the other hand, the genetic proximity of these genes is likely to have functional consequences for NK cell and T lymphocyte function. This highlights the importance of assessing clusters of related genes as opposed to studying genes in isolation. In-vitro and in-vivo studies that can better understand the immunological relevance of specific allelic combinations of these genes are therefore needed.

Other immune-related genes within the MHC are positioned between the HLA Class I and II genes (sometimes referred to as the HLA Class III region). Genes within this region include the central mediating cytokine Tumour Necrosis Factor alpha gene (TNF-α), which exhibits variations within the regulatory portion of the gene; this may lead to different expression and secretion levels of the cytokine. Furthermore, variations within the TNF-α gene and surrounding region also form haplotypes (Allcock, Windsor et al. 2004) reinforcing the need to functionally assess the relevance of combinations of variations within a cluster rather than a single polymorphism.

The IFN genes involved in antiviral immune responses are also an example of a complex gene family including IFN-α, -β, -γ and -λ, each with several subtypes and in some cases allelic variation. The IFN-λ family includes the genes IL28A, IL28B and IL29 on chromosome 19. Genetic variation of the IL28B gene (encoding IFN-λ3) has become of recent interest in the field of Hepatitis C due to the identification of variations upstream of this gene associated with HCV infection and treatment outcome in recent genome-wide association studies (GWAS) (discussed later). Additional variation is also observed for toll-like receptors (TLRs), all of which are associated with the host’s immune response.

3. Hepatitis C infection: A case example demonstrating the effect of host and viral genetic diversity on infection outcome

Hepatitis C is a global health issue with more than 170 million people worldwide suspected of carrying the virus. Variations in the global distribution of HCV reflect different exposure risks. In developed nations emergence of new cases is predominantly associated with the use of intravenous drugs while in developing nations the use of unscreened blood or blood products and increasing use of intravenous drugs are the leading causes of new cases (Shepard, Finelli et al. 2005) (Lavanchy 2011) (Dore, Law et al. 2003).

HCV is a single-stranded RNA virus and following infection about 30% of individuals spontaneously clear the virus within the first six months of infection. The majority of individuals develop chronic infection of which about 20% develop cirrhosis and 3% develop hepatocellular carcinoma (Venook, Papandreou et al. 2010) (NIH 2002). In developed countries, HCV infection is a major indicator for liver transplantation. Current standard of care for HCV involves the use of pegylated-interferon alpha in combination with ribavirin (pegIFN-α/RBV) but this treatment regime is effective in only about 50% of cases and is typically associated with a plethora of side-effects. There is no protective or therapeutic vaccine available for HCV.
Like other complex diseases, HCV infection and treatment outcome is the result of interactions between various host and viral determinants. The classical approach of heritability estimates using family and twin studies indicate a significant genetic contribution to HCV infection outcomes (Goncales, Fernando et al. 2000; Fried, Kroner et al. 2006). In a study including 3993 haemophilia individuals infected with HCV, the concordance rate between 257 sibling pairs was two-fold higher than randomly paired subjects for spontaneous and treatment-related clearance of HCV. The heritability estimate from the study was approximately 0.3 (heritability of 1 indicates total phenotype variation due to genetic differences). These findings indicate that genetic factors play an important role in HCV disease outcome. This review will focus on genetic variability in host and viral determinants to provide an insight into the possible mechanisms underlying HCV infection and treatment outcome.

3.1 Host genetic diversity and natural history of HCV infection

Upon infection, HCV triggers the host innate and adaptive immune responses. Accordingly, many studies have shown that polymorphisms in genes involved in the host’s innate and adaptive immune response are associated with HCV infection outcome (reviewed in (Rauch, Gaudieri et al. 2009) (Schmidt, Thimme et al. 2011) ; Figure 1 and Table 2). Nevertheless, we still lack clear understanding of the underlying mechanisms by which a person develops chronic infection or progresses to end-stage liver disease.

---

**Fig. 1.** Host genetic factors and viral adaptation associated with natural history of HCV infection outcome. The metabolic and fibrogenic genes are part of a seven gene signature used to predict liver fibrosis (Huang, Shiffman et al. 2007).
### Table 2. Immune response gene polymorphisms associated with HCV infection and treatment outcome

<table>
<thead>
<tr>
<th>Gene family</th>
<th>Gene/Allele</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA class I &amp;II</td>
<td>HLA-A<em>02, -A</em>11, -B<em>07, -DQB1</em>03, -DRB1<em>04, -DRB1</em>11, -DRB1*08</td>
<td>Spontaneous clearance</td>
</tr>
<tr>
<td></td>
<td>HLA-A<em>23, -A</em>01, -B<em>08, -B</em>61, -B<em>38, -Cw</em>01, -Cw*03</td>
<td>Persistence</td>
</tr>
<tr>
<td>ISGs</td>
<td>PKR -168 CT genotype</td>
<td>Spontaneous clearance</td>
</tr>
<tr>
<td></td>
<td>OAS-I polymorphism in 3'UTR- GG genotype</td>
<td>Persistence</td>
</tr>
<tr>
<td></td>
<td>MxA -88T</td>
<td>SVR</td>
</tr>
<tr>
<td></td>
<td>OAS-I - rs3213545 T, rs1169279 A and rs2859398 C alleles</td>
<td>SVR</td>
</tr>
<tr>
<td></td>
<td>SOCS3-4874 AA genotype</td>
<td>Treatment</td>
</tr>
<tr>
<td>KIR</td>
<td>KIR2DL3-HLA-C ligand</td>
<td>Spontaneous Clearance</td>
</tr>
<tr>
<td>IFNs</td>
<td>IL-28B - rs12979860 CC genotype</td>
<td>Spontaneous Clearance and SVR</td>
</tr>
<tr>
<td></td>
<td>IL-28B - rs8099197 TT genotype</td>
<td>SVR</td>
</tr>
<tr>
<td></td>
<td>IFN-γ -764G</td>
<td>SVR</td>
</tr>
<tr>
<td>TLR</td>
<td>TLR-7 variants</td>
<td>Persistence</td>
</tr>
<tr>
<td>Cytokines</td>
<td>IL-10-592AA genotype</td>
<td>Spontaneous Clearance</td>
</tr>
<tr>
<td></td>
<td>TNFα -863C/-308G haplotype (black subjects), -863C</td>
<td>Persistence</td>
</tr>
<tr>
<td></td>
<td>IL-18 -607A, -137C</td>
<td>SVR</td>
</tr>
<tr>
<td></td>
<td>IL-10-1082 GG genotype</td>
<td>SVR</td>
</tr>
<tr>
<td></td>
<td>IL-10 - ACC promoter haplotype</td>
<td>Progression to liver fibrosis</td>
</tr>
<tr>
<td></td>
<td>IL-10 -592AA, -819TT genotypes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-12B - 3'UTR 1188C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-18 - 607A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-6 -572C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-18 -607C/<em>, IL-18 -137G/</em>, IFN-γ 874T/*, IL-10 -1082A/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-10 -824T, -1087A, -1087/-824 haplotypes AT and AC</td>
<td></td>
</tr>
</tbody>
</table>

SOCS3 = suppressor of cytokine signalling 3. Underlined HLA alleles appear in more than one study.

3.1.1 Candidate gene studies

3.1.1.1 Immune response genes associated with HCV outcome

The innate immune system has a critical role in detecting pathogens and triggering the immune response. TLRs activate the innate immune system by detecting pathogen associated molecular patterns (PAMPs) such as viral RNA. TLR7 and TLR8 are receptors for single-stranded RNA and are localised on the endosomal membrane. Ligation of these TLRs can activate intracellular pathways resulting in the production of IFNs and subsequently several anti-viral IFN stimulated genes (ISGs). Variations in TLR7 and some ISGs have been shown to be associated with HCV infection outcome (reviewed in (Rauch, Gaudieri et al. 2009). One such study has shown the association between variations within TLR 7 and HCV infection outcome (Schott, Witt et al. 2008). In this study, variants at position 32 and 2403 of TLR7 were associated with viral persistence and IFN-α treatment outcome. Other genetic variations within the ISGs MxA, OAS-1 and PKR have been associated with HCV infection outcome (Knapp, Yee et al. 2003) (Table 2).

Not surprisingly, polymorphisms in genes encoding cytokines and chemokines have been associated with HCV outcome. A recent meta-analysis on IL-10 gene polymorphisms and HCV infection outcome found the IL-10 (-1082) GG genotype was more frequent in subjects with persistent HCV infection and correlated with higher IL-10 serum levels (Zhang, Zhang et al. 2010). IL-12 plays an important role as a pro-inflammatory cytokine in T helper 1 lymphocyte differentiation and a study on 123 chronically infected and 72 spontaneous resolvers showed that viral persistence was associated with homozygosity for a 3’UTR variant of IL-12 and correlated with lower production of IL-12 (Houldsworth, Metzner et al. 2005).

3.1.1.2 NK cell KIRs and HLA ligands

Several groups have investigated the level at which the genetically polymorphic NK receptors and ligands influence HCV infection outcome. Two studies have confirmed that homozygosity for the NK cell inhibitory receptor KIR2DL3 and its ligand, HLA-C alleles belonging to the C1-group (defined by asparagine at residue 80 of the extracellular domain), results in a higher probability of resolution of HCV infection (Khakoo, Thio et al. 2004; Vidal-Castineira, Lopez-Vazquez et al. 2010). It has been hypothesised that this combined genotype results in relatively weak inhibition of NK cells and therefore protects against HCV by rendering NK cells more easily activated than in other subjects. This would be consistent with the interaction between KIR2DL3 and C1-group HLA-C alleles being relatively weak compared to the alternative combinations KIR2DL2 + C1 and KIR2D1 + C2 (Moesta, Norman et al. 2008). Furthermore, other HLA-C and KIR as well as HLA-B and the 3D KIR gene interactions are also likely to be involved in HCV resolution as has been shown for HIV infection (Gaudieri, DeSantis et al. 2005).

3.1.1.3 HLA

Differences in the strength and breadth of CD4+ and CD8+ HCV-specific T cell immune responses is an important correlate of HCV infection outcome (Lechner, Wong et al. 2000; Lauer, Lucas et al. 2005). Host HLA molecules govern the cellular T lymphocyte response to HCV and accordingly HLA class I and class II alleles have been associated with HCV outcome. Genetic association studies have shown that HLA class I alleles HLA-B*27, -A*11, -
Cw*01, -A*03 and -B*57 are associated with viral clearance, while HLA-B*38, -Cw*07, -B*08 and -A*01 are associated with viral persistence or chronic infection. HLA class II alleles, which present antigen to CD4+ T cells, are also shown to be associated with viral clearance or persistence. HLA-DRB1*04, -DRB1*11, -DQB1*03 and -DRB1*01 are associated with viral clearance, while HLA-DRB4*01 and -DRB1*07 are associated with chronic infection (reviewed in (Rauch, Gaudieri et al. 2009; Schmidt, Thimme et al. 2011) (Table 2).

Studies of HCV single source outbreaks provides the opportunity to examine the influence of HLA by removing or reducing confounding effects such as gender and age. One study on individuals from an Irish HCV single source outbreak including 227 women (141 chronic and 86 cleared) that were exposed to HCV contaminated immunoglobulin showed the association of HLA-B*27 with spontaneous resolution (McKiernan, Hagan et al. 2004). This study highlights the contribution of the host's immune CD8+ T cell response and HLA genes to HCV outcome. However, the relationship between HLA and HCV outcome should be considered in the context of viral variation. It is likely that viral immune escape mutations present in the incoming virus will affect HCV infection outcome as has been shown using single source outbreak cohorts (Merani, Petrovic et al. 2011) (Salloum, Oniangue-Ndza et al. 2008). Accordingly, viral adaptation to the host's HLA-restricted immune response will be an important correlate of infection outcome (Figure 1).

3.1.2 GWAS

The initial association between polymorphisms upstream of IL28B and HCV infection outcome were identified using GWAS. GWAS are used to identify single nucleotide polymorphisms (SNPs) associated with disease outcome without a priori knowledge. The method screens a large number of host variants in hundreds or thousands of case/control subjects. The SNPs are pre-selected based on how well they “tag” certain areas or genes due to the linkage disequilibrium pattern across the genome, such that not all of the more than three million base differences (on average) between two unrelated individuals need be sampled. Although tagging SNPs may not necessarily be the causative variation they are likely to be near putative causative variations that can be further investigated.

Rauch et al utilised a GWAS to identify host genetic factors associated with HCV spontaneous clearance (Rauch, Kutalik et al. 2010). The SNP rs8099917 upstream of the gene IL28B was the strongest predictor for HCV clearance (OR=2.31, CI=1.74-3.04). Similarly, Thomas et al also showed a SNP upstream of IL28B was associated with spontaneous resolution of HCV in different ethnic groups (European OR=2.6, CI=1.9-3.8 and African OR=3.1, CI=1.7-5.8); in this case the SNP was rs12979860 (Thomas, Thio et al. 2009). Individuals homozygous for the C allele at rs12979860 were more likely to clear HCV infection than individuals with the CT or TT genotype. The protective effect of the CC genotype was seen in subjects belonging to Caucasian as well as African-American ethnicity. Although the two studies identified different SNPs, both flanked the IL28B gene and are likely to reflect a protective IL28B haplotype.

IL28B encodes for the type III IFN-λ3. IFN-λ stimulates an intracellular cascade that turns on IFN-α/β like anti-viral responses. In addition, IFN-λ plays a vital role in inhibiting HCV manifestations by interfering with virus replication (Robek, Boyd et al. 2005). Furthermore, genetic variations upstream and flanking IL28B may correlate to different expression levels and many groups are now trying to understand the role of IFN-λ in HCV infection.
3.2 Host and viral factors influencing treatment response

The current standard of treatment for HCV is pegIFN-α/RBV. However, this treatment regimen is associated with several severe side-effects and sustained virological responses (SVR) is observed in 30-80% subjects. Several factors have been shown to be predictive of SVR, which include HCV genotype (GT) and pre-treatment viral load. Subjects infected with GT 1 have a low SVR rate of less than 50%, while subjects infected with GTs 2 or 3 have higher SVR rates of approximately 80% (Selzner and McGilvray 2008).

Various approaches such as candidate gene studies and GWAS have identified genes associated with different response rates to pegIFN-α alone or combination therapy (Table 2). The most widely studied genes encode for the IFN genes and ISGs. A study on subjects in the Hepatitis C Antiviral Long-Term Treatment Against Cirrhosis (HALT-C) trial demonstrated associations between variants in IFN-α pathway genes and treatment outcome. Specifically, variants in the Interferon a receptor R1 (IFNAR1) (-22T>G), IFNAR2 (-33T>C), Janus kinase 1 (JAK1) (+112G>T) were associated with SVR. The Tyrosine kinase 2 (TYK2) -2256 A allele was also associated with SVR (Welzel, Morgan et al. 2009). Su et al examined the association in 12 ISGs in 374 treatment naïve HCV subjects. Three SNPs in the IFN induced 2'-5' oligoadenylated synthetase -like gene (OASL) (rs3213545 T, rs1169279 A and rs2859398 C alleles) were found to be associated with SVR (Su, Yee et al. 2008). The T allele in the promoter (–88) of MxA was also associated with SVR (Suzuki, Arase et al. 2004).

Although these association studies highlight the relevance of IFN and ISG variants in HCV treatment, additional gene expression studies have been carried out to directly correlated expression levels with treatment outcome. Chen et al identified 18 differentially expressed genes in liver biopsy samples obtained before the commencement of IFN-α treatment. Most differentially expressed genes in treatment responders and treatment non responders were IFN sensitive genes (including OAS2 and 3 and MxA) and two (ISG15/USP18) belonged to the IFN regulatory pathway (Chen, Borozan et al. 2005). Up regulation of ISG15/USP18 was identified as a predictor of IFN therapy outcome. ISG15, a ubiquitin-like protein, and USP18, part of the protease pathway, are thought to be important in host innate defense immunity.

Four independent groups utilised GWAS to evaluate the association between SNPs in the human genome and HCV treatment outcome (Ge, Fellay et al. 2009; Suppiah, Moldovan et al. 2009; Tanaka, Nishida et al. 2009; Rauch, Kutalik et al. 2010). The SNP rs8099917 located 10kb upstream of the IL28B gene (discussed above) was identified in all four studies as showing a significant association with HCV treatment outcome. SNP rs12979860 located 3kb upstream of IL28B was shown by Ge and colleagues (Ge, Fellay et al. 2009) to also be associated with HCV treatment outcome. The four studies included different ethnic groups but genotyping and analytical methods were similar, although the type of commercial SNP set used by the groups varied and may account for the lack of detection of SNP rs12979860 in some of the studies (Rauch, Rohrbach et al. 2010). The same SNPs upstream of IL28B were associated with HCV treatment and infection outcome.

The GWAS performed by Ge et al utilised subjects in the IDEAL study; a large randomized control trial comparing the effectiveness of different forms of pegIFN-α (Ge, Fellay et al. 2009). The North American study included 1137 subjects from three ethnics groups Caucasian, African American and Hispanic. Seven SNPs were reported to be associated with treatment outcome with the SNP rs12979860 having the strongest correlation to treatment outcome with the CC genotype showing a two-fold greater rate of SVR than genotype TT in
all ethnic groups. The C allele frequency is highest in south East Asian populations and accordingly they achieve the highest SVR with combined therapy than subjects from European or African background. Furthermore, the frequency of the protective C allele is significantly higher in individuals from European ancestry than African American and explains half of the difference in response between the two ethnic groups.

Interferon inducible gamma 10kDa protein (IP-10), which belongs to the chemokine family has also been identified as a predictor of treatment response. The low plasma level of IP-10 was shown to be associated with a decrease in HCV RNA and SVR in early or first phase of treatment (Askarieh, Alsiö et al. 2010). In another study by Lagging et al, they correlated variants in three SNPs related to \textit{IL-28B} with pre-treatment plasma levels of IP-10 (Lagging, Askarieh et al. 2011). These studies suggest assessment of IP-10 levels and testing of \textit{IL-28B} variants can help predict treatment response in a clinical setting.

Cytokines profiles have been studied with respect to viral clearance and persistence during the current treatment regime. Polymorphisms in \textit{IL-12}, \textit{IL-18}, \textit{IL-10} cytokines have been shown to be associated with HCV treatment outcome (Wan, Kung et al. 2009). The findings of most of these studies are inconsistent reflecting the differences in study design and likely ethnic differences.

Similar to host genetic factors, genetic variability within the viral genome is likely to influence treatment response. Studies have indicated that amino acid substitutions in the core region (70 and 91), which forms the nucleocapsid of HCV, are related to treatment outcome. Similarly, amino acid substitutions in the interferon sensitivity determining region (ISDR) located in the non-structural 5A region (NS5A) of the viral genome is found to be associated with favorable treatment outcome (Maekawa and Enomoto 2009).

Better understanding of the complex interactions between genes and environmental factors may significantly help in predicting the response to current HCV therapy in the future.

\subsection*{3.3 Host and viral factors influencing disease progression}

Chronic liver disease is a known morbidity of HCV infection with 20\% of chronic HCV cases developing these complications, of which approximately 5\% go on to develop hepatocellular carcinoma. Studies have shown demographic as well as epidemiological factors are associated with progression to fibrosis. Some subjects progress rapidly to early stage of fibrosis, while some do not develop fibrosis. Thus genetic factors predisposing to fibrosis are crucial in pathogenesis of hepatic fibrosis.

Pro-inflammatory and anti-inflammatory genes have been shown to be associated with liver disease progression in HCV. The anti-inflammatory cytokine IL-10 is likely to have an immuno-modulatory role in fibrosis. Promoter polymorphisms in \textit{IL-10}, which affect the level of IL-10 production, have been shown to be associated with susceptibility to liver fibrosis. In a study of Japanese chronic HCV subjects, the IL-10 -824 T allele, -1087 A allele and -1087/-824 haplotypes AT and AC were shown to be risk factors for progression of hepatic fibrosis (Ishida, Ikebuchi et al. 2011). Another study investigated the combined effect of SNPs in \textit{IL-10}, \textit{IL-18} and IFN-\(\gamma\) genes on 77 chronic HCV infected patients. Subjects carrying 3-4 high risk genotypes were associated with greater risk of developing liver cirrhosis (Bouzgarrou, Hassen et al. 2011).
Risk factors identified in these studies would help in understanding the molecular mechanisms involved in pathogenesis of chronic liver disease in HCV infected individuals. Such studies may also provide new insights for developing drug targets thus reducing the burden of HCV worldwide.

4. HCV diversity and viral adaptation

As HCV is an obligate intracellular virus with a small RNA genome, it evolves by escaping/adapting to the host’s immune response. The mechanisms utilised by HCV to establish persistent infection are not fully understood but involve several strategies (Nolan, Gaudieri et al. 2006).

HCV has a high mutation rate due to the lack of proof-reading activity of its RNA dependent RNA polymerase. This characteristic along with the high rate of replication of HCV results in the circulation of genetically related variants/quasispecies within the infected individuals. In addition, HCV is classified into six major GT and more than 50 different subtypes based on the sequence variability of the virus. GTs 1 and 3 are the most prevalent GTs and also account for the majority of HCV infections worldwide (Chayama and Hayes 2011).

The efficacy of HCV-specific T lymphocyte responses is therefore compromised by the mutability of the viral genome within HLA Class I and II-restricted epitopes and this represents a potent escape strategy of HCV (Timm, Lauer et al. 2004; Thimme, Lohmann et al. 2006) as well as other viruses such as HIV (Goulder and Watkins 2004). These mutations may exist in the infecting virus or arise in-vivo can affect HLA-peptide binding and therefore influence the selection of peptides presented by an individual. These mutations may also impair HLA-peptide and T lymphocyte interaction and reduce the efficiency of the T lymphocyte-mediated immune response. The relevance of HCV immune escape mutations in chronic infection was first demonstrated in chimpanzees (Weiner, Erickson et al. 1995; Erickson, Kimura et al. 2001; Grakoui, Shoukr y et al. 2003) and subsequently in humans (Cox, Mosbruger et al. 2005; Tester, Smyk-Pearson et al. 2005) (reviewed in (Bowen and Walker 2005). Evidence for T lymphocyte escape and reversion has also been recently demonstrated in acute infection (Timm, Lauer et al. 2004) and 18-22 years after a common-source outbreak (Ray, Fanning et al. 2005). Flanking epitope escape mutations have also been described and alter proteasomal epitope processing and subsequent peptide presentation (Seifert, Liermann et al. 2004). Overall, this emphasizes the complexity of multiple effective mechanisms of T lymphocyte immune escape in HCV.

Studies utilizing population-based genetic approaches, which examine the association between specific viral polymorphisms and the HLA types of individuals within a host population, have demonstrated viral escape (adaptation) to HLA-restricted immune pressure evident at the population level for both HIV and HCV (Gaudieri, Rauch et al. 2006; Timm, Li et al. 2007). The approach is based on the premise that the host’s HLA molecules regulate immune responses by presenting specific viral epitopes to T lymphocytes and viral polymorphisms within or flanking these epitopes that compromise the efficacy of these T lymphocyte responses (i.e. allow viral escape) would be identified as HLA-specific HCV polymorphisms (viral adaptations). Accordingly, HCV can escape host immune responses via mutations that confer a selective advantage resulting in the fixation (or dominance) of specific strains within the host and to some extent drive HCV evolution.
Given the extent of sequence diversity between the GTs (Figure 2), it would be anticipated that the sequence context of epitopes restricted for specific HLA molecules would be altered and potentially disrupted for other GTs. The initial population-based genetic study on HCV GT 1 (Gaudieri, Rauch et al. 2006) was extended to examine HCV adaptation to HLA-restricted immune responses in GT 3 (Rauch, James et al. 2009). The study reported that the immune escape profiles differ between the two main circulating GTs 1 and 3 reflecting the extensive variation between the GTs and the observation that individuals can clear one GT but can be re-infected with other GTs with limited protection.

Fig. 2. HCV diversity in the (a) NS3 and (b) NS5B regions for Caucasian chronic HCV populations.

5. Summary

Host and viral genetic determinants within genes related to the host immune response are important in determining infection outcome, including HCV outcome. Studies accounting for variability in both host and pathogen are likely to provide a more complete understanding of the complex host/pathogen interaction and its variable outcomes. Specifically, an understanding of how viral adaptation to host HLA-restricted immune responses (and perhaps to NK cells given recent evidence of “footprints” in HIV genome due to KIR genotypes; (Alter, Heckerman et al. 2011)) affects infection outcome will provide an insight into vaccine design.
Given our current knowledge, the HCV virus would be particularly amenable to a T lymphocyte vaccine strategy given that immunity in those that control the virus consists of broad T lymphocyte responses and viral structural constraints that prevent escape from all mounted T lymphocyte responses. Identifying the viral T lymphocyte targets of a successful host immune response is crucial to vaccine design and should consider the genetic diversity of the virus as well as the HLA genetic system of the host population for which the vaccine is designed.

6. References


Genetic Variation of Host Immune Response Genes and Their Effect on Hepatitis C Infection and Treatment Outcome


Analysis of Genetic Variation in Animals includes chapters revealing the magnitude of genetic variation existing in animal populations. The genetic diversity between and within populations displayed by molecular markers receive extensive interest due to the usefulness of this information in breeding and conservation programs. In this concept molecular markers give valuable information. The increasing availability of PCR-based molecular markers allows the detailed analyses and evaluation of genetic diversity in animals and also, the detection of genes influencing economically important traits. The purpose of the book is to provide a glimpse into the dynamic process of genetic variation in animals by presenting the thoughts of scientists who are engaged in the generation of new idea and techniques employed for the assessment of genetic diversity, often from very different perspectives. The book should prove useful to students, researchers, and experts in the area of conservation biology, genetic diversity, and molecular biology.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: