



# MONASH University

## **Interferon Lambda Polymorphisms: Validation and Characterisation**

***Sara Bonanzinga***  
***(B. Sc. (Hons))***

A thesis submitted for the degree of Master of Biomedical Science at  
Monash University in 2016  
Department of Microbiology

## Copyright notice

© The author (2016). Except as provided in the Copyright Act 1968, this thesis may not be reproduced in any form without the written permission of the author.

*I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.*

## Abstract

Polymorphisms in the interferon lambda (IFN $\lambda$ ) gene cluster have been shown to influence spontaneous and treatment induced clearance of hepatitis C virus (HCV) infection. The most studied single nucleotide polymorphism (SNP), rs12979860, lies within the first intron of IFN $\lambda$ 4, though is still commonly referred to as IFN $\lambda$ 3. Several association studies have been performed, with little insight into the mechanisms behind the observations.

Assay validation provided clear evidence that plasma samples provide sufficient genomic DNA to detect IFN $\lambda$ 3 genotypes, as well as being robust under lengthy storage conditions allowing retrospective testing.

Association studies previously performed have focused on individuals infected with HCV genotype 1. To explore this further, the relationship between IFN $\lambda$ 3 genotype and HCV subtypes 1a and 1b was investigated, and it was determined that the benefit of having a CC genotype was greater in individuals infected with HCV 1b.

Nationwide distribution of IFN $\lambda$ 3 genotype was determined, and it was found that the predominant rs12979860 genotype was CT, followed by CC, then TT. This profile is similar to that in North America and Western European countries, likely a reflection of migration and ethnic diversification. In Queensland, Tasmania, and South Australia, the CT genotype accounts for more than 50% of individuals. In contrast, the predominant genotype in New South Wales was CC (50%), and in both the Northern Territory and Victoria, CC and CT were roughly equal in distribution.

The structure of the IFN $\lambda$  gene cluster and its products was studied in an attempt to find why multiple copies of very similar genes have been retained. It was found that substitution of C for the ancestral T at rs12979860 changes the predicted local DNA structure and creates a potential binding site for the transcription factor E2F, which can act as both a repressor and activator. In addition, a tyrosine residue (Y160) found only in the IFN $\lambda$ 2 amino

acid sequence, may be functionally significant because it confers the potential for additional post-transcriptional regulation.

It was predicted that differences in local DNA structure produced by rs12979860 alleles might affect transcription. Enzyme linked Immunosorbent Assays (ELISAs) were used to measure the plasma concentrations of IFN $\lambda$ 2 and IFN $\lambda$ 3. A difference in the IFN $\lambda$ 2:IFN $\lambda$ 3 ratio was observed: the T allele correlated with a decrease (CC > CT > TT) due mainly to a decrease in IFN $\lambda$ 2.

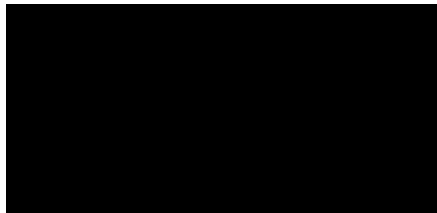
Thymidine phosphorylase (TP) is an enzyme that is known to be induced by interferons, suggesting that its activity may be influenced by IFN $\lambda$  genotype. It was found that the presence of the T allele correlated with an increase (CC < CT < TT) in whole blood TP activity, implying that the T allele and IFN $\lambda$ 3 are associated with stronger inflammatory responses.

The work described in this thesis encompasses a range of different techniques, and combined a number of different disciplines; molecular biology, epidemiology, genetics, and biochemistry. It has shown that the rs12979860 polymorphism in the IFN $\lambda$  region probably has a far greater impact than is currently appreciated.

## Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature: .....



Print Name: ...Sara Bonanzinga.....

Date: 3.10.2016.....

## Acknowledgements

To my supervisor, Professor Hans Netter, thank you for taking me on as your student. Even as an off-campus student, I have found you to be approachable and encouraging. It has been fortuitous that you came to VIDRL in the final stages of my candidature, and I thank you for your supervision.

My co-supervisor Professor Scott Bowden; I thank you. Thank you for the suggestion and encouragement to undertake further studies, and for the opportunity to complete my Masters degree while continuing to work in your laboratory. Your ongoing support has been incredible. It has been a busy few years, but it has been worth it. I have a great sense of accomplishment, which I wouldn't have without you.

An enormous amount of gratitude goes to Tim Shaw, who was not only an immense help in the direction of the gene duplication chapter, but also performed the HPLC assay. I appreciate all the time you have spent teaching me in the past few years, from your love of proteins and genetics, to your endless search for that little bit extra. The biggest life lesson I walk away with is to not believe dogma unless I understand it.

During the early months of my candidature, a lot of time was spent talking and laughing with Dr Jacinta Holmes. Thank you for continuing the routine IFN $\lambda$ 3 diagnostic testing whilst I was on maternity leave, and keeping the numbers ticking over. IFN $\lambda$ 3 was a common subject matter in our respective studies, but it is on a personal level that our friendship was secured.

To the people I work with in the Research and Molecular Development division at VIDRL, I thank you. When you go out into the 'real world', you realise that there are a lot of people who are unhappy with their work life. I am not one of those people; I love my job, and I am grateful to be able to work with the group of people that I do. In particular, I would like to thank those who have worked in the Molecular Microbiology section of our division. So to Liz, Lilly, Jean, Kathy, and Rachel – I thank you.

And of course, I thank my family. I am lucky to have an incredibly supportive husband, who has never doubted my ability to accomplish what I set out to do. And I thank my amazing daughter, who will not understand for many years just how much I love coming home to her excitable voice yelling "Mummy's home!!"

## Table of Contents

List of Figures.....	viii
List of Tables.....	x
List of Abbreviations.....	xi
Chapter 1: Introduction .....	1
1.1: Hepatitis C Virus.....	1
1.1.1: HCV Classification.....	1
1.1.2: HCV Disease Progression .....	2
1.1.3: HCV Transmission.....	3
1.1.4: HCV Diagnosis and Monitoring.....	5
1.1.5: HCV Treatment.....	6
1.2: Single Nucleotide Polymorphisms and Genome-Wide Association Studies.....	7
1.3: Interferon Lambda SNPs and HCV Treatment Response .....	9
1.3.1: Discovery of a Correlation Between Interferon Lambda 3 and HCV Response.....	9
1.3.2: Identification of IFN $\lambda$ 4.....	13
1.3.3: Treatment Responses .....	16
1.3.4 Spontaneous HCV Clearance .....	19
1.3.5: Metabolic Associations.....	20
1.3.6: Interferon Stimulated Genes.....	21
1.3.7: Association with Outcome in Transplant Patients .....	21
1.4: Association of Interferon Lambdas in non-HCV Settings .....	23
1.5: Summary.....	25
Chapter 2: Validation of an Interferon Lambda 3 Allelic Discrimination Assay.....	26
2.1: Introduction .....	26
2.2: Methods.....	28
2.2.1: Sample Selection and Extraction Methods.....	28
2.2.3: IFN $\lambda$ 3 Genotyping and Allelic Discrimination .....	30
2.2.4: White Blood Cell Count .....	32
2.2.5: Longevity of Genomic DNA.....	32
2.2.6: Freeze-Thaw Analysis.....	32
2.3: Results .....	33
2.3.1: Sample Extraction Comparisons .....	33
2.3.2: White Blood Cell Count .....	37

2.3.3: Longevity of Genomic DNA.....	37
2.3.4: Effect of Multiple Freeze-Thaw Cycles.....	40
2.4: Discussion .....	42
2.4.1: Extraction Methods.....	42
2.4.2: Longevity Studies .....	46
2.4.3: Effect of Multiple Freeze-Thaw Cycles.....	48
Chapter 3: Optimisation of the IFN $\lambda$ 3 Assay .....	50
3.1: Introduction .....	50
3.2: Methods.....	52
3.2.1: DNA Extraction .....	52
3.2.2: IFN $\lambda$ 3 Genotyping: ABI Allelic Discrimination Protocol.....	52
3.2.3: Reduced Volume PCR Reaction using In-House Assay .....	52
3.2.4: Master mix Reagent: SensiFAST Genotyping Kit .....	53
3.2.5: TIB MolBiol LightMix® Kit IFN $\lambda$ 3 rs12979860 Assay .....	54
3.3: Results .....	55
3.3.1: Reduced volumes.....	55
3.3.2: SensiFAST Genotyping Kit .....	57
3.3.3: TIB MolBiol LightMix Kit .....	62
3.4: Discussion .....	66
3.4.1: Reduced Volumes .....	66
3.4.2: SensiFAST Genotyping Kit .....	67
3.4.3: TIB MolBiol LightMix Kit .....	69
3.5: Conclusion.....	71
Chapter 4: Association of IFN $\lambda$ 3 SNPs with Different HCV Genotype 1 Subtypes .....	72
4.1: Introduction .....	72
4.2: Methods.....	74
4.2.1: Patients .....	74
4.2.2: IFN $\lambda$ 3 Genotyping.....	74
4.2.3: HCV Genotyping .....	74
4.2.4: HCV Detection .....	74
4.2.5: Statistical Analysis .....	75
4.3: Results .....	75
4.4: Discussion .....	80
4.5: Summary.....	85
Chapter 5: IFN $\lambda$ 3 Genotype Distribution Throughout Australia.....	87
5.1: Introduction .....	87
5.2: Methods.....	89



5.3: Results .....	89
5.4: Discussion: .....	92
Chapter 6: Gene Duplication, Phylogeny and Function of the Interferon $\lambda$ Family Members .....	96
6.1 Background.....	96
6.2: Part 1: Theoretical .....	96
6.2.1 Aims and Methods.....	96
6.2.2 Results and Discussion.....	97
6.2.3 IFN $\lambda$ Phylogeny: Homologs of IFN $\lambda$ s and their Receptors in other Vertebrates..	101
6.2.4 Structures of the <i>IFN<math>\lambda</math></i> genes and the genes that encode their receptors.....	106
6.2.5 Comparison of Individual IFN $\lambda$ genes and products.....	108
6.2.6 Influence of rs12979860 on DNA and RNA structure .....	111
6.2.7 Summary and Conclusions of Part 1 .....	114
6.3: Part 2: Experimental.....	115
6.3.1 Methods .....	116
6.3.2 Results .....	117
6.3.3 Discussion .....	122
6.4 Conclusion.....	124
Chapter 7: Association Between Thymidine Phosphorylase, Inflammation and Interferon $\lambda$ ....	125
7.1 Introduction.....	125
7.1.1 Regulation of TYMP expression.....	128
7.1.2 Physiological role of TP: Regulation of Mitochondrial Replication.....	129
7.1.3 Induction of TP Activity by Interferons and Other Pro-Inflammatory Stimuli .....	130
7.1.4 Hypothesis and Aim .....	131
7.2 Materials and Methods.....	131
7.2.1 Blood Samples.....	131
7.2.2 Assay for Thymine Phosphorylase Activity .....	131
7.3 Results.....	132
7.4 Discussion.....	135
Chapter 8: Final Discussion .....	137
References .....	141

## List of Figures

Figure 1.1: Illustration of rs12979860 SNP.....	11
Figure 1.2: Human chromosome 19.....	15
Figure 1.3: Global allele distribution and frequency for rs12979860.....	17
Figure 2.1: Average ct values for each extraction method.....	34
Figure 2.2: Allelic Discrimination and real time PCR output.....	36
Figure 3.1: Results of post-read analysis following SensiFAST master mix.....	60
Figure 3.2: Results of post-read analysis following SensiFAST master mix on known IFN $\lambda$ 3 CC samples.....	62
Figure 3.3: Melt-curve analysis from LightCycler software.....	64
Figure 4.1: SVR according to HCV-1 subtypes and IFN $\lambda$ 3 genotypes.....	80
Figure 5.1: rs12979860 genotype distribution patterns according to States and Territories in Australia.....	90
Figure 5.2: Global allele frequency of rs12979860 (including Australia).....	91
Figure 6.1: Alternative outcomes of gene duplication.....	97
Figure 6.2: Orientation of IFN $\lambda$ genes (and pseudogenes) on human chromosome 19.....	98
Figure 6.3: Detailed map of the IFN $\lambda$ locus.....	100
Figure 6.4: Evolutionary succession and co-occurrence tree.....	103
Figure 6.5: Predicted interaction networks.....	104
Figure 6.6: Phylogeny of representative IFN $\lambda$ genes.....	105
Figure 6.8: Size comparison of IFN $\lambda$ genes and the associated receptor genes, <i>IL10RB</i> and IFN $\lambda$ R1.....	106

Figure 6.9: Long-term activation of IFN $\lambda$ 2 and IFN $\lambda$ 3 may be limited by receptor availability.....	107
Figure 6.10: Amino acid conservation comparison between IFN $\lambda$ 1-4.....	108
Figure 6.11: Alignment of the amino acid sequences of the human type III IFN family.....	110
Figure 6.12: Alignment of IFN $\lambda$ 2 and IFN $\lambda$ 3 amino acid sequences and the corresponding model of IFN $\lambda$ 2 protein.....	111
Figure 6.13: Structure of the <i>IFN<math>\lambda</math>4</i> first intron.....	112
Figure 6.14: Effect of rs12979860 SNP on local DNA structure.....	114
Figure 6.15: Competition of IFN $\lambda$ 2 and IFN $\lambda$ 3 for receptor complexes.....	116
Figure 6.16: IFN $\lambda$ 2 and IFN $\lambda$ 3 concentrations according to IFN $\lambda$ 3 rs12979860 genotype, and the ratio between them.....	119
Figure 6.17: IFN $\lambda$ 2 and IFN $\lambda$ 3 ELISA results re-standardised and compared according to IFN $\lambda$ 3 rs12979860 genotype.....	121
Figure 7.1: TP pathway diagram.....	125
Figure 7.2: Radial phenogram of TP .....	126
Figure 7.3: Alignment of amino acid sequences of coelacanth and human TP.....	127
Figure 7.4: TP proximal promoter region.....	128
Figure 7.5: HPLC assay for TP activity.....	132
Figure 7.6: Measurement of TP activity.....	133
Figure 7.7: Kinetics of TP activity in blood.....	134
Figure 7.8: Relationship between TP activity and IFN $\lambda$ 3 genotype.....	135

## List of Tables

Table 2.1: Primer and Probe sequences for IFN $\lambda$ 3 genotyping assay.....	31
Table 2.2: White blood cell count in four components of blood.....	37
Table 2.3: Effect of long-term storage of plasma at -20°C.....	38
Table 2.4: ct values of buffy coat derived DNA samples .....	39
Table 2.5: ct values following freeze-thaw cycles.....	41
Table 3.1: Mastermix volumes for SensiFAST trial.....	53
Table 3.2: PCR parameters for melt curve analysis.....	54
Table 3.3: ct values of initial reduced PCR reactions.....	56
Table 3.4: ct values of reduced PCR reactions.....	57
Table 3.5: The ct values obtained from the amplification plot.....	59
Table 3.6: A further 4 IFN $\lambda$ 3 genotype CC specimens with SensiFAST master mix.....	61
Table 3.7: Comparison of TaqMan allelic discrimination and LightMix protocols.....	65
Table 4.1: IFN $\lambda$ 3 and HCV genotype distribution.....	76
Table 4.2: Distribution of IFN $\lambda$ 3 genotypes within HCV 1 subtypes.....	77
Table 4.3: Association of HCV genotype and IFN $\lambda$ 3 genotype with gender.....	78
Table 5.1: Breakdown of ethnicity according to States.....	92
Table 6.1: Comparison of product sizes of <i>IFN<math>\lambda</math></i> genes and their receptors.....	106
Table 6.2: Usable results from the two ELISA assays.....	118

## List of Abbreviations

3PNPLA3	Phospholipase Domain-Containing Protein
ABI	Applied Biosystems
AIDS	Acquired Immune Deficiency Syndrome
ALT	Alanine Transaminase
AST	Aspartate Aminotransferase
BMI	Body Mass Index
CSF	Cerebrospinal Fluid
ct	Cycle Threshold
DNA	Deoxyribonucleic Acid
dNTP	deoxynucleotide triphosphate
EDTA	Ethylenediaminetetraacetic Acid
GWAS	Genome-Wide Association Study
HBV	Hepatitis B Virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
IDEAL	Individualized Dosing Efficacy vs. Flat Dosing to Assess Optimal Pegylated Interferon Therapy
IFN	Interferon
IFN $\alpha$	Interferon Alpha
IFN $\lambda$	Interferon Lambda
ISG	Interferon Stimulated Genes
IU	International Units
Jak-STAT	Janus kinase/Signal Transducers and Activators of Transcription
LDL	Low-Density Lipoprotein
LDLR	Low-Density Lipoprotein Receptor
LiPA	Line Probe Assay
NAFLD	Non-Alcoholic Fatty Liver Disease
NANBH	non-A, non-B hepatitis
NTC	No Template Control
NVR	Null Virological Responder
PBMC	Peripheral Blood Mononuclear Cell
PCR	Polymerase Chain Reaction
pegIFN $\alpha$	Pegylated Interferon alpha
RBV	Ribavirin
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism
SVR	Sustained Viral Response

TdR	Thymidine
TdR	deoxypyrimidine
TK2	TdR kinase
Tm	Melting Temperature
TP	Thymidine Phosphorylase
TTP	thymidine triphosphate
UdR	Deoxyuridine
VIDRL	Victorian Infectious Diseases Reference Laboratory
VR	Virological Responder

## Chapter 1: Introduction

### 1.1: Hepatitis C Virus

Hepatitis C virus (HCV) causes a large global disease burden, with approximately 150-180 million chronically infected individuals worldwide, and 499,000 dying annually from HCV related liver disease [1, 2]. HCV infections are often asymptomatic and around 25% of those infected clear the virus spontaneously with the remainder developing chronic infection. In chronic infection, the virus continues to replicate in the liver and, after a period of at least 10 years, can lead to fibrosis, and then to cirrhosis; disease progression may be accelerated if additional risk factors apply, such as alcohol consumption. Ultimately, the disease may progress to hepatocellular carcinoma (HCC). Australia has an estimated 230,000 people living with chronic HCV infection, including 50,000 with moderate to severe liver disease [3]. HCV related cirrhosis is the leading indication for liver transplantation in Australia [4] and the United States of America [5]. Recently, the World Health Organization has adopted a global health strategy, which takes steps to halt viral hepatitis transmission, and eliminate the public health threat by 2030 [6].

#### 1.1.1: HCV Classification

HCV is classified in the *Hepacivirus* genus in the *Flaviviridae* family, which also includes the flaviruses and pestiviruses [7]. The *Flaviviridae* have similar genomic structure and replication strategies; they are positive, single-stranded, enveloped RNA viruses, which replicate in the cytoplasm of the host. They infect a range of hosts, and are generally able to replicate in both vertebrate and invertebrate cells [8]. HCV is unusual among the *Flaviviridae*

as it does not have an arthropod vector, or an animal reservoir. Mosquito-borne illnesses attributed to this family of viruses include Dengue Fever and Yellow Fever, while other members of the family use ticks as vectors, such as in tick-borne encephalitis.

HCV was formally recognised as the causative agent of the then-called non-A, non-B hepatitis (NANBH) in 1989 and was the first virus identified based on a molecular approach [9]. Prior to this, the viruses responsible for hepatitis A and hepatitis B had been identified but the existence of a third viral hepatitis agent had been recognized; it was known to be transmitted percutaneously as it was seen in both injecting drug users and in blood transfusion recipients [5]. Once the virus had been identified, specific antibody (anti-HCV) assays were developed, allowing for routine testing and improving the safety of the blood supply [10] which significantly reduced the rate of new incidences of HCV infection. Since the discovery of HCV, nucleotide sequence data has accumulated. Sequence analysis of HCV isolates show that the virus possesses considerable heterogeneity, allowing it to be classified into a number of genotypes. HCV classification has most recently been expanded to include seven genotypes (1-7) and 67 subtypes [11]. Previous to this publication, there were six known genotypes for HCV and these remain the major categories as there is only a single isolate of HCV genotype 7 [11, 12].

### **1.1.2: HCV Disease Progression**

Two distinct phases of disease are recognised following HCV infection; acute clinical hepatitis, and chronic infection. Symptomatic acute HCV infection is uncommon with the majority of people infected with HCV developing a chronic infection [13], in which the virus remains in the body, often asymptotically. This is vastly different to otherwise healthy



adults whom acquire hepatitis B virus (HBV) infection, with 90% clearing the virus and developing protective immunity [14]. Once HCV infection has become chronic, it is likely to be the immune response, rather than the virus itself, which causes liver damage [15].

### **1.1.3: HCV Transmission**

Prior to the screening of donor blood for anti-HCV, introduced during the early 1990s, between 5 – 10% of HCV infections were attributed to transfusions; this mode of transmission has since been eliminated. Before the introduction of a mandatory screening program, recipients of multiple transfusions were placed at high risk with approximately 94% of people with severe haemophilia A in Melbourne testing positive to anti-HCV [16]. In countries where donors are remunerated for their blood, there are higher rates of anti-HCV in the donor blood compared to those adopting a voluntary blood donation system [17]. Of greater global concern, there remain some countries in which donor screening for diseases including HIV, HBV, and HCV are not performed routinely [18]. This has a flow-on effect to other countries; immigrants to Australia are not necessarily screened for infections such as HCV, and studies have shown that 3.4% of immigrants from some sub-Saharan African countries test positive to anti-HCV [19].

Approximately 90% of new HCV infections in Australia are in people with a recent history of injecting drug use [20]. This makes control of the disease problematic, as the majority of injecting drug users belong to a largely marginalised group within the society. As HCV infection is often asymptomatic with a slow pathogenesis, injecting drug users may not be seeking treatment as they do not have adequate information regarding the long term

consequences of hepatitis C [21]; those who have a regular health care practitioner are more likely to initiate a course of treatment, as are those whose doctors inform them about the possibility of progression to cirrhosis or HCC [22]. It has also been reported that people who are committed to cease using illicit drugs are more willing to undergo treatment for their HCV infection [22]; those who undertake therapy whilst still taking drugs are more likely to be non-compliant to follow up treatment than those who abstain from drug use or persons who have no history of drug use [23]. Although drug use is a former contraindication for HCV treatment, educating and treating this patient group will likely result in the greatest decline of new HCV infections. There has been a series of longitudinal studies following injecting drug users in Melbourne; the data acquired from this has not only allowed modelling of HCV transmission in these groups [24], but has also led to the concept of “treatment as prevention” [25]. The idea of this approach being that an individual is treated for HCV infection, and their HCV RNA negative status reduces transmission to their injecting partners, which in turn reduces further transmission to their injecting partners. Furthermore, if the primary HCV positive injecting partners are also treated (‘treat your friends’); this reduces the frequency of new infections to uninfected individuals, reinfection following clearance, and can potentially reduce HCV prevalence amongst injecting drug users further than treating individuals at random [25].

Prison inmates constitute another population in which HCV infection rates are higher than the general population. Modes of transmission inside prisons, including drug use, piercings, tattooing, and physical assaults, have allowed the prevalence of HCV infection within the Victorian prison system to be around 39% [26]. With such a high prevalence rate, it is imperative to remember that people do not remain incarcerated for life. Education and

counselling of people within this group is important, so that once their sentences are complete they can seek further treatment, reducing liver disease progression and the potential for an ongoing risk to the greater community. Historically, opportunities to treat within the prison system have been largely limited by the lengthy treatment regimen. With the recent availability of the direct-acting antivirals (DAAs), this limitation has been largely minimised, with treatment times as short as 12 weeks.

Mother-to-baby transmission of HCV has been shown to be an uncommon event with the risk of virus acquisition being around 5% in HCV RNA-positive mothers and approximately 10% when the mother has HIV co-infection [27, 28]. Current guidelines do not include treatment of babies.

#### **1.1.4: HCV Diagnosis and Monitoring**

Diagnosis of HCV infection involves both serological and molecular testing. Anti-HCV assays are highly sensitive, but are unable to distinguish between a chronic or resolved infection. Once an infection is resolved, the anti-HCV remains detectable, but HCV RNA is cleared. In order to diagnose a current infection, HCV RNA must be detected by nucleic acid testing; there are both qualitative and quantitative assays available for diagnostic use. Quantitative HCV RNA assays have become increasingly sensitive, and able to detect virus levels as low as 12 International Units (IU)/ml. Quantification of HCV RNA becomes important when monitoring a patient's response to treatment, as well as being a pre-treatment indicator of achieving virus loss (a sustained viral response (SVR)). There is no vaccine to prevent HCV infection but antiviral therapy is available. The aim of HCV therapy is to achieve a SVR, which

is defined as a sample in which reverse transcription PCR is unable to detect HCV RNA in the serum taken 24 weeks post treatment (SVR 24). Several studies now use SVR 12 (12 weeks after cessation of therapy) after it was shown that it has a strong correlation with SVR 24 [29].

#### **1.1.5: HCV Treatment**

Until recently, combination pegylated interferon alpha (pegIFN $\alpha$ ) and Ribavirin (RBV) was the standard treatment (standard of care) for patients with chronic hepatitis C. It involves a weekly injection of pegIFN $\alpha$ , with oral intake of RBV twice daily. In patients with acute HCV infection, treatment with IFN $\alpha$ -2b prevents progression to a chronic infection, with a cure rate of 98% [30]. Duration of therapy is dependent on HCV genotype; patients infected with genotypes 2 and 3 receive 24 weeks of treatment and patients infected with non-genotype 2-3 HCV (predominantly genotype 1 in Australia) receive 48 weeks. Patients infected with genotype 2 or 3 achieve SVR rates of 70-80% while those infected with genotype 1 only achieve SVR rates of 45-50%. The treatment regime can be demanding, and comes with a plethora of side effects, primarily associated with pegIFN $\alpha$ , including flu-like symptoms (fever, chills, and aches), cough, rash, hair loss, and depression. Because of these debilitating effects, many patients are unable to complete treatment. Altering dosage of medication is also necessary in some patients, such as in the case of RBV-induced anaemia, where between 9-13% of patients require dose reductions [31]. A recent study has identified two single nucleotide polymorphisms (SNPs) on human chromosome 20 that predict the chance of acquiring RBV-induced anaemia during early treatment [32].

The overall aim of HCV therapy is to achieve an SVR. There are a number of viral and host factors that predict the likelihood of achieving an SVR, including the baseline viral load, HCV genotype, ethnicity, gender, and body mass index (BMI).

## **1.2: Single Nucleotide Polymorphisms and Genome-Wide Association Studies**

A SNP is defined as a genetic variation at a single nucleotide (containing any of the nitrogenous bases adenine, cytosine, guanosine, or thymine) at a particular locus, either between individuals, or in paired chromosomes, leading to variation of genotypes within a population. The completion of the human genome project established the map of the human genome sequence in 2001 [33], with more than 1.4 million SNPs initially identified. The paper pointed to the potential of running Genome Wide Association Studies (GWAS) and linkage disequilibrium studies; this would eventually lead to identifying the underlying genetics for some disease phenotypes [33].

A GWAS aims to identify a correlation between SNPs and a particular phenotype, often related to disease or metabolic associations with treatment. They are performed with the use of a microarray chip with up to 1 million SNPs. Depending on the platform used, the SNP in each sample can be detected by differential hybridization (ie, allelic discrimination with pre-loaded SNP alleles), or by detecting the allele present at the loci using bead technology [34]. The GWAS can identify whether a SNP is associated with a disease state, by comparing allele frequencies in a group of affected individuals with that of a healthy population.

Linkage disequilibrium is the occurrence of alleles at different loci being inherited in non-random frequencies. This is of particular interest in studies of disease risk and susceptibility; if the occurrence of a particular allele is shown to correlate with a disease state, it may serve as a marker for predicting a disease, or in determining appropriate dosage of therapies. A follow-up review [35] from one of the authors of the original report outlined many of the advances made in the decade following the original publication. A major advance has been the increased sensitivity of applications, including detecting chromosome imbalances and identifying candidate genes for conditions such as autism [36]. The advances in technology and affordability has now broadened the number of known human SNPs to over 44 million [37]. Most common SNPs are inherited in haplotype blocks; the presence of one allele almost always predicts the presence of a particular allele at a correlating locus [38].

In bi-allelic SNPs, three alternative genotypes are present within a population; an individual may be homozygous for the minor or major allele, or has a heterozygote status. A minor allele is characterised as the least common allele in any given population; as genetic inheritance varies between different ethnic groups, it is possible that a minor allele in one population may be the major allele in another population. In regions of low rates of historical recombination, populations exhibit low haplotype diversity; Gabriel and colleagues [39] detected the greatest haplotype diversity in African (or African American) populations, followed by European then Asian populations. In addition, they reported that of the 28% of haplotypes identified solely in one population group, 90% were found in the African groups.

To date, many GWAS have been performed on populations of European ancestry, as SNP chips are able to account for up to 10 proxy genotypes, but for the African populations, the existing SNP chips only match genotype subsets [40]. Need and Goldstein [40] highlight the disparity of GWAS as performed by ethnicity; at the time of their publication, 86% of the studies performed were exclusively on people of European descent. A notable reason for the population bias is that the regions which have sufficient funding for GWAS are primarily European. In a multicultural society in which medication rules are to be determined according to genetic background and risks, it is important to include a wide range of the representative ethnicities in any study.

When a GWAS is performed, and a SNP which correlates with the phenotype is identified, there are two possible forms of association; it may identify the causal allele, or alternatively, be in linkage disequilibrium with the causal allele. Both provide data which allows genetic-based treatment or diagnosis, but the former also provides a site to which treatment may be directed. Risk loci have been identified for a number of disorders, and have led to alternative therapeutic approaches [35].

### **1.3: Interferon Lambda SNPs and HCV Treatment Response**

#### **1.3.1: Discovery of a Correlation Between Interferon Lambda 3 and HCV Response**

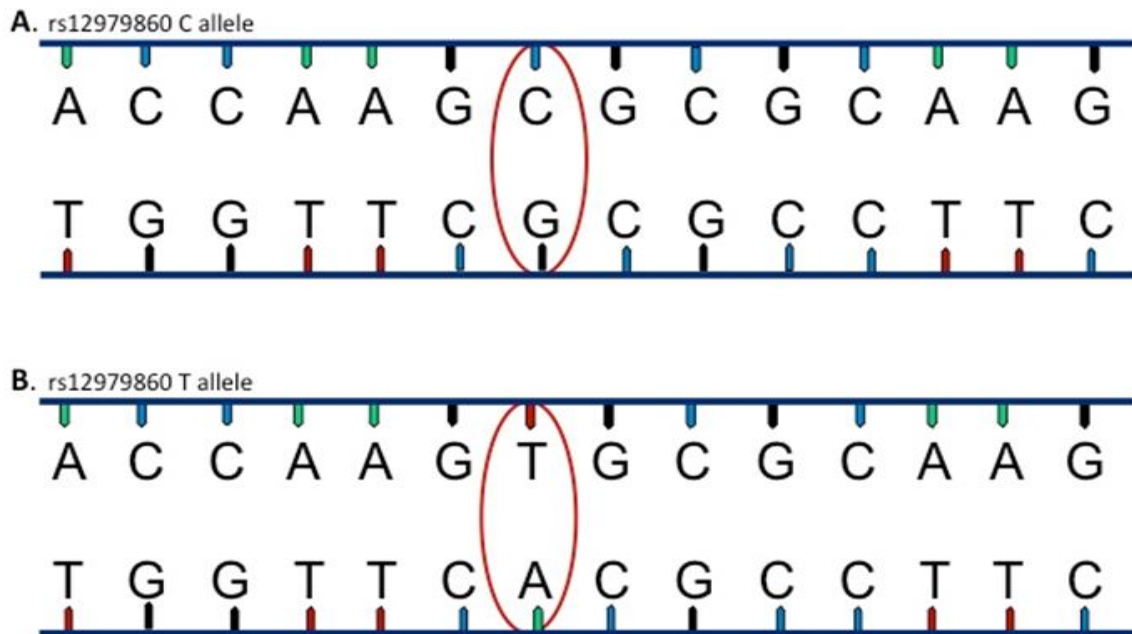
In an attempt to identify host factors contributing to clearance of HCV, a number of institutions performed their own respective GWAS. In 2009 – 2010, four separate studies identified a SNP involved in HCV clearance, all in the region of the interferon lambda (IFN $\lambda$ ) gene cluster on chromosome 19 [41-44]. Although each of the studies identified SNPs within

the same region, the causal variant remains elusive. At the time of these landmark studies, it was accepted that the type III IFN family consisted of IFN $\lambda$ 1, IFN $\lambda$ 2 and IFN $\lambda$ 3 (previously designated IL29, IL28A, and IL28B, respectively). Each of these genes contains five exons, which is distinct from type I IFNs which do not contain introns [45]. Like other IFN families, type III IFNs induce an immune response upon viral infection [46]. They signal via the Janus kinase/Signal Transducers and Activators of Transcription (JAK-STAT) pathway, and switch on a number of interferon stimulated genes (ISGs) [47]. While type I IFN (IFN $\alpha$ ) and type III IFN (IFN $\lambda$ ) have a similar gene expression profile, IFN $\alpha$  often induces gene expression to a higher magnitude than IFN $\lambda$  [48]. In addition to this, and perhaps bringing in some relevance to the association with response to HCV treatment, expression of IFN $\lambda$  is augmented in influenza-infected cells when pre-treated with IFN $\alpha$  [49]. IFN $\lambda$ 3 expression requires the engagement of a receptor complex, consisting of IFN $\lambda$ 3R and IL10RB to activate the JAK-STAT pathway. While IL10RB is ubiquitously expressed, IFN $\lambda$ R1 is limited to epithelial cells, hepatocytes, and dendritic cells [50].

Ge and colleagues [41] performed a GWAS on 1671 chronically infected HCV genotype 1 patients, who had undergone 48 weeks treatment of pegIFN $\alpha$  -2b or pegIFN $\alpha$  -2a and RBV, with the majority being from the “Individualized Dosing Efficacy vs. Flat Dosing to Assess Optimal Pegylated Interferon Therapy” (IDEAL) study [51]. The patient subgroup was assessed with an SVR phenotype in response to the therapy. Seven SNPs with an association with SVR were identified on chromosome 19, near the IFN $\lambda$ 3 gene. The SNP with the highest association ( $p=1.37 \times 10^{-28}$ ) was rs12979860, 3kb upstream of the IFN $\lambda$ 3 gene. It is bi-allelic with three possible genotypes; CC (good response), CT (intermediate response), and TT



(poor response) (Figure 1.1). They reported IFN $\lambda$ 3 genotype CC increased the likelihood of SVR by 2-3 fold across the ethnic groups studied (Caucasian, African American, and Hispanic), and also reported that the favorable genotype is more frequent in Caucasians, compared to Hispanics and African Americans (39%, 34%, and 16%, respectively) [41].



**Figure 1.1: Illustration of a SNP; variation at the nucleotide level results in different genotypes. In rs12979860, there are two alternate alleles; C (A), or T (B).**

The three remaining independent GWAS identified rs8099917 as the SNP with the highest association to SVR in HCV genotype 1 patients [42-44]. This SNP is located 8.9kb upstream from the IFN $\lambda$ 3 gene, and is also bi-allelic. Genotype TT is associated with a good response, and genotypes GT and GG are associated with a poor response to pegIFN $\alpha$ /RBV therapy. As with rs12979860, a heterozygote has a slightly improved SVR, although not clinically significant. It is important to note that in each of these studies, the rs12979860 SNP was not included in the analysis, nor were people of African descent included in the study populations.

Suppiah et al [43] used a Caucasian Australian cohort of 293 genotype 1, chronic hepatitis C patients, of which 131 responded to HCV antiviral therapy, and 162 did not. Their GWAS was performed using data for over 300,000 SNPs. Several SNPs were identified which suggested an association to treatment response; the SNP with the highest predictive value was rs8099917 ( $p=7.06 \times 10^{-8}$ ). A confirmation cohort with samples from Europe and Australia (an additional 555 patients) was used for further investigation. They carried 172 SNPs to this, and found 14 which had an association with HCV clearance. Again, their data specified rs8099717 to have the highest association ( $p=9.25 \times 10^{-9}$ ).

Tanaka et al [44] used a cohort of 154 Japanese HCV genotype 1 patients; 82 null virological responders (NVR), and 72 responders (VR). NVR was defined as a less than 2  $\log_{10}$  decline in HCV RNA at 12 weeks pegIFN $\alpha$ /RBV treatment, and HCV RNA detectable at 24 weeks. They reported two SNPs which correlated highly with NVR; rs8099917 had the highest association (possible genotypes TT, GT, GG;  $p=3.11 \times 10^{-15}$ ), followed by rs12980275 (possible genotypes AA, AG, GG;  $p=1.93 \times 10^{-13}$ ). The associations were replicated in an additional cohort ( $n=172$ ), with the same results (rs8099917;  $p=2.68 \times 10^{-32}$ , rs12980275;  $p=2.84 \times 10^{-27}$ ). The poor response alleles (G for rs12980275, and G for rs8099917) were found in higher frequencies in the NVR group, and patients homozygous for the poor response allele were found exclusively in the NVR group. They also showed that expression levels of IFN $\lambda$ 3 in PBMCs were lower in patients with the minor alleles.

The fourth GWAS was performed by Rauch and colleagues [42] in a Caucasian Swiss population. The study included 1362 patients; 1015 chronically infected with HCV, and 347 who had spontaneously cleared the virus. They looked at associations for both response to

treatment, and to spontaneous clearance. The top SNP identified was rs8099917. The G allele was associated with a progression to chronicity ( $p=6.07 \times 10^{-9}$ ), as well as failure to achieve a SVR ( $p=3.11 \times 10^{-8}$ ). Therefore, this SNP has been identified as playing a role in clearing HCV both naturally, and with the aid of pegIFN $\alpha$  and RBV treatment. Viral factors were also assessed in this study, and the effect of the SNP was strongest in HCV genotypes 1 and 4 patients, where treatment is often challenging.

Following the initial GWAS studies, many researchers have examined IFN $\lambda$ 3 SNPs in their studies, leading to an abundance of publications on this topic. This included expanding the clearance associations, initially identified in HCV genotype 1 patients, to those infected with HCV genotypes 2 and 3 [52]. Although correlations exist between IFN $\lambda$ 3 genotype and a vast degree of factors involved with HCV infection, the causative association of these polymorphisms on disease outcome is presently unknown.

Many studies have shown variations between CC and non-CC genotypes (rs12979860), suggesting that perhaps the C allele offers a recessive protective effect [53]. Carriers of the T allele, in particular homozygous TT individuals, are more likely to progress to chronic HCV infection than CC individuals [54].

### **1.3.2: Identification of IFN $\lambda$ 4**

In 2013, Prokunina-Olsson et al [55] identified a dinucleotide variation rs368234815 (initially named ss469415590) (TT or  $\Delta$ G) upstream from rs12979860 on chromosome 19. The

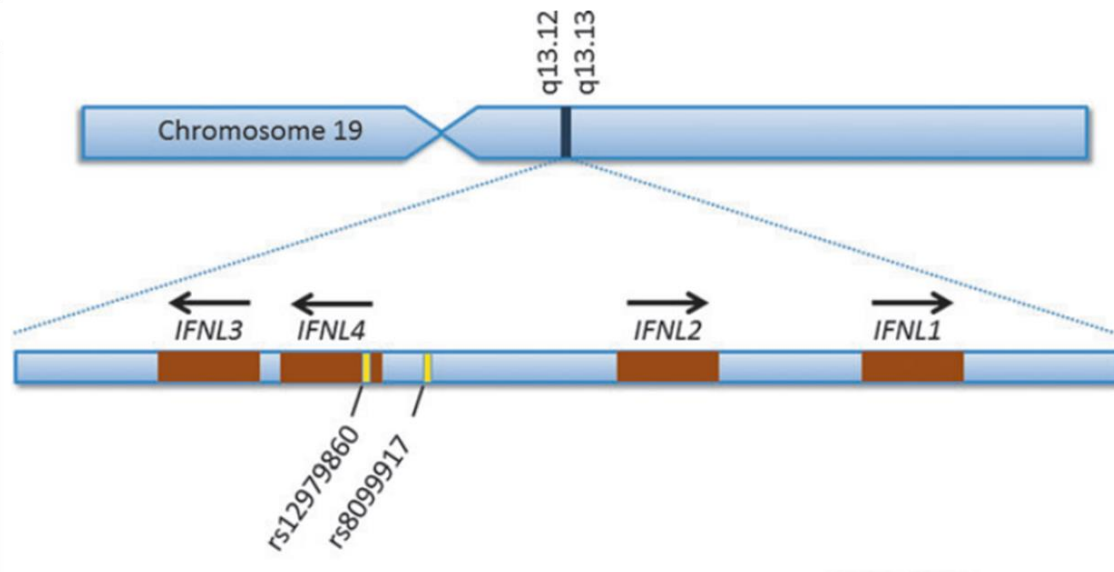
protein from this newly identified gene, designated IFN $\lambda$ 4 (and a member of the type III IFN gene family), is unfavourable for individuals infected with HCV, and is only expressed when a person carries at least one copy of the rs368234815 [ $\Delta$ G] variant. A beneficial insertion allele (rs368234815 [TT]) causes a frame-shift mutation, and halts the production of the IFN $\lambda$ 4 protein. This beneficial variant is common in Asians (93%), Europeans (68%), and in the minority of Africans (23%), once again supporting previous observations of differences in ethnic backgrounds in the response to HCV therapy [55] .

The dinucleotide variant rs368234815 is in strong linkage disequilibrium with IFN $\lambda$ 3 rs12979860 in Asian and European populations, but less so in African populations. This association between IFN $\lambda$ 4 and treatment response for HCV was interrogated by a number of subsequent studies to establish if there was any further advantage to using rs12979860. It was found to offer no further predictive value in Caucasian patients with HCV genotypes 1 or 4 [56, 57]. In contrast, for African-American individuals, rs368234815 was a better indicator for both treatment response and spontaneous clearance than rs12979860 [55]. The correlation is lesser with IFN $\lambda$ 3 rs8099917 across all ethnic groups analysed, and studies have shown that IFN $\lambda$ 4 is invariably a better predictor of treatment outcome for patients infected with HCV than rs8099917 [55, 57].

It has been suggested that this IFN $\lambda$ 4 SNP is a causal variant to the effect seen by rs12979860 and others, although the precise mechanism remains unclear [58]. A number of previous studies have shown that increased ISG levels are detrimental to viral clearance, including pre-treatment levels independent of IFN $\lambda$ 3 genotype [59-61], and associated with IFN $\lambda$ 3 genotype [62-64]. By having a higher baseline ISG level, the individuals have a limited

capacity to produce endogenous IFN production and are more refractory to exogenous IFN therapy [61]. The production of IFN $\lambda$ 4 in  $\Delta$ G carriers reduces the host's ability to clear HCV; Terczynska-Dyla et al [65] showed that by having a functioning IFN $\lambda$ 4 gene, carriers of the rs368234815  $\Delta$ G allele have increased activation of ISGs, a correlation previously observed in unfavourable IFN $\lambda$ 3 genotypes [62, 64]. They concluded that the increase of ISG expression in HCV patients is directly resulting from production of the IFN $\lambda$ 4 protein [65].

The different predictive values for IFN $\lambda$ 4 between IFN $\lambda$ 3 SNPs rs12979860 and rs8099917 have been attributed to the proximity to the gene itself, with rs12979860 located within the first intron of IFN $\lambda$ 4 itself, and rs8099917 being further downstream (Figure 1.2) [58].



**Figure 1.2: Human chromosome 19, with the location and orientation of the IFN $\lambda$  gene family, including “IFN $\lambda$ 3” SNP locations rs12979860 and rs8099917 [66]**

In concordance with the higher frequency of  $\Delta$ G alleles found in African populations compared to Asian/Caucasian populations, Key et al [67] performed sequence alignments

from various mammals, and deduced  $\Delta G$  to be the ancestral genotype. Due to the level of sequence similarity among mammals, they concluded it is a functionally relevant interferon.

### **1.3.3: Treatment Responses**

The initial studies identifying the correlation between IFN $\lambda$ 3 genotype and HCV response were performed when dual therapy of pegIFN $\alpha$ /RBV was the standard treatment [41-44]. Although the chances of achieving an SVR are higher in IFN $\lambda$ 3 rs12979860 CC patients, those who do not respond to treatment are at greater risk of serious disease than their non-CC counterparts.

IFN $\lambda$ 3 genotypes follow a much defined geographical distribution; it is likely that these differences have arisen through selection pressure rather than a founder effect, though there is no concrete evidence in support of either view. The protective allele is almost fixed throughout South-East Asia, intermediate in Europe, and is the minor allele in Africa (Figure 1.3) [53]. This is also reflective of the HCV genotypes more prevalent in these geographically distinct areas; in Africa genotype 4 is the predominant strain, which is notoriously difficult to treat. In contrast, South-East Asia has a high proportion of HCV 6c-I; patients with this genotype experience a greater chance of SVR. Fittingly, this is the same pattern seen in rates of HCV clearance, with persons of African ancestry more likely to develop a chronic HCV infection, and to fail therapy than patients of other ethnic backgrounds [68-70]. However, once a chronic disease is established, people of African descent have a slower rate to fibrosis compared to Caucasian patients [71, 72]. In addition, Asian patients with a chronic HCV infection have a higher chance of developing HCC than their Caucasian and

African counterparts [73]. These patterns of disease progression mirror IFN $\lambda$ 3 genotype frequencies; those ethnicities that have favourable IFN $\lambda$ 3 genotype have a stronger immune response. Of interest, this correlation between IFN $\lambda$ 3 genotypes and fibrosis is not limited to HCV infection alone, but is also found in patients infected with HBV and non-alcoholic fatty liver disease patients [74]. The fibrosis, cirrhosis, and HCC develop as a result of the immune response; if the body is less equipped to respond to an infection, the capacity to eliminate the disease will be lessened.



**Figure 1.3: Global allele distribution and frequency for rs12979860 [53]**

Adverse psychological effects whilst on IFN therapy have been well documented [75-77]. Though a number of symptoms are seen across all groups, often increasing with time whilst on treatment, three were seen to be associated with the IFN $\lambda$ 3 C allele; loss of energy, lack of sleep, and reduced appetite [78]. Interestingly, a significant loss of energy and appetite were seen in IFN $\lambda$ 3 CT patients, followed by CC and then TT. This is an interesting

observation, as many associations can be clearly defined by the carriage of a beneficial allele, rather than the heterozygote specifically.

Due to the time of discovery, and cohorts available for retrospective analysis, many studies regarding correlation of IFN $\lambda$ 3 and IFN $\lambda$ 4 genotypes to HCV therapy were based on a treatment regime of pegIFN $\alpha$ . The relevance of these nucleotide variations has been widely questioned with the advent of newer therapies, including triple therapy and interferon-free regimes [79-81]. With the uptake of the newer therapies, a number of groups performed analyses to determine whether there was any benefit in testing for these variations. Although the treatment response rates using triple therapies are superior to dual therapy across all IFN $\lambda$ 3 genotypes, there are still significant differences between IFN $\lambda$ 3 and IFN $\lambda$ 4 genotypes.

In one study of 146 Japanese HCV genotype 1b patients undergoing Telaprevir (a direct acting HCV NS3 protease inhibitor) based triple therapy, 100% of patients with rs8099917 TT (favourable allele) showed an end of treatment response, compared to 85% of non-TT patients [82]. In another Japanese study, patients with IFN $\lambda$ 4 TT genotype had an additive predictive value in SVR when undergoing triple therapy (pegIFN/RBV/Telaprevir) compared to dual therapy ( $p < 0.0001$ ) [83]. This effect is also relevant with interferon-free regimes; a correlation has been observed between IFN $\lambda$ 4 and Sofosbuvir/RBV in HCV 1 patients [84], with a slower early viral decline seen in  $\Delta$ G carriers.

A cost analysis by Camma et al [85] revealed HCV genotype 1, IFN $\lambda$ 3 CC patients may be better off on a dual therapy approach; not only do these patients have a higher probability



of clearing the virus than their non-CC counterparts, it also reduces the adverse reactions related to the first generation protease inhibitors, Boceprevir and Telaprevir, and the risk of development of resistance mutations [86]. Whilst on Boceprevir, a larger percentage of patients require dose reduction and cessation due to anaemia when compared to those on dual pegIFN/RBV therapy (21% and 2%, vs 13% and 1% respectively) [87].

#### **1.3.4 Spontaneous HCV Clearance**

IFN $\lambda$ 3 genotype not only has a role in patients on antiviral therapy for HCV, but is also significantly associated with spontaneous clearance of the virus [42, 53, 54]; that is, patients whom test positive to anti-HCV, but have undetectable HCV RNA without having undertaken a course of treatment. This association is seen across both European and African ethnicities [53].

In 1978-79, a total of 2867 women were exposed to HCV via anti-D immunoglobulin contaminated with HCV genotype 1b [88, 89]. As a single-source outbreak, it offers a unique insight into the pathogenesis of the disease. By eliminating any variables due to the viral strain, it allows host responses to be analysed as a separate entity. Tillman et al [90] reported that those women with CC genotype (rs12979860) were more likely to clear HCV spontaneously, with 66% of CC patients resolving the infection without therapy, compared to 24% and 3% for CT and TT patients, respectively. Jaundice was more common in patients with the CC genotype while non-CC individuals who were jaundiced during acute infection were more likely to spontaneously clear the virus than those without jaundice. In contrast, the authors found that jaundice had no significant association with the likelihood of

spontaneous viral clearance in those with a CC genotype [90]. The findings are consistent with those patients having a CC genotype developing a more potent early multi-specific immune response. Previous studies have shown jaundiced individuals are more likely to clear HCV [13, 91]; which links prior clearance indicators with IFN $\lambda$ 3 genotypes.

In contrast to the European study [88, 89], while assessing the differences in spontaneous clearance rate in HCV infected Chinese blood donors, Rao et al [92] identified four SNPs with a significant association. The SNP with the highest association was rs8099917, and there was no correlation between spontaneous clearance and the rs12979860 genotype. Both rs8099917 TT and rs12979860 CC genotypes are highly represented in the Chinese population, seen at a rate of 93% and 92% respectively [93]. This observation further indicates there are more factors involved with spontaneous clearance in different ethnic populations than the IFN $\lambda$ 3 genotype alone.

#### **1.3.5: Metabolic Associations**

IFN $\lambda$ 3 genotype has been shown to influence a number of metabolic traits in patients infected with HCV, including steatosis, cholesterol levels, and insulin resistance. The presence of steatosis has been found to be lower in rs12979860-CC than non-CC [94], while other groups have found this trend to have no statistical significance [95]. Steatosis is also significantly associated with BMI, to which there is no correlation to IFN $\lambda$ 3 genotype [95]. Higher total cholesterol levels are seen in patients with IFN $\lambda$ 3 CC genotype [95]. CC homozygotes have lower triglyceride levels [95]

IFN $\lambda$ 3 CC is associated with higher serum low-density lipoprotein (LDL) levels in HCV genotype 1 patients, but not in HCV genotype 3 patients, nor in a healthy control group [96]. A SNP in the low-density lipoprotein receptor (LDLR), rs14158, was found to be in linkage disequilibrium with IFN $\lambda$ 3 rs12979860. Although the predictive value for treatment response in HCV patients is not as high as the IFN $\lambda$ 3, it is proposed that it augments the predictive power [54]. Carriers of the IFN $\lambda$ 3 T allele (rs12979860) are more likely to have insulin resistance; this finding was reported in non-diabetics, and held true across several HCV genotypes [95].

#### **1.3.6: Interferon Stimulated Genes**

In chronic HCV patients, carriers of the IFN $\lambda$ 3 T allele have higher ISG expression than found in homozygous CC patients [62-64]. This is in contrast to expression levels in normal uninfected liver; IFN $\lambda$ 3 CC individuals have highest levels of hepatic ISG expression, followed by CT, then TT with the lowest expression levels [97]. Further from being associated with IFN $\lambda$ 3 genotype, higher ISG levels were associated with non-responders and also patients who relapsed. This suggests that ISG expression is not determined by IFN $\lambda$ 3 genotype; both are independently associated with treatment response [63].

#### **1.3.7: Association with Outcome in Transplant Patients**

The literature describing the significance of IFN $\lambda$ 3 genotype in a liver transplant setting has created some controversy. A common observation is that the frequency of the good

response IFN $\lambda$ 3 genotype is lower in transplant recipients in HCV infected patients than in the donors [98-102]. This is consistent with those progressing to severe liver failure being more likely to be non-CC, given both the greater spontaneous clearance and more successful treatment rate of IFN $\lambda$ 3 CC patients. When comparing liver transplant non-HCV patients, the frequencies of IFN $\lambda$ 3 genotype differ, with a greater number of CC recipients seen as compared to HCV patients. Furthermore, IFN $\lambda$ 3 genotype had no effect on the outcome of the transplant in non-HCV settings. [98].

Allam et al [98] observed that in non-HCV patients IFN $\lambda$ 3 genotype has no significance in graft survival. In contrast in the HCV group have a higher probability of graft survival if the recipient is genotype CC or CT. They also observed that the donor genotype had no association with graft survival, regardless of HCV status of the recipient.

IFN $\lambda$ 3 TT patients are more likely to have an earlier recurrence of HCV following transplantation compared to either CC or CT [98, 99]. The opposite is seen in livers from CC donors, with recurrence and viral loads increasing in comparison to non-CC donors [99, 101], perhaps suggesting that the protective allele is reversed in donors. The greatest risk of severe recurrence occurs in non-CC recipients who receive a liver from a CC donor [103]. When both recipients and donors carry the IFN $\lambda$ 3 CC genotype, SVR rates of 83% - 100% [100, 104] have been reported, but are reduced to just 16% - 25% if neither recipient nor donor is genotype CC [104, 105]. If a CT recipient received a liver from a CC donor, an SVR rate of 61% was reported, in contrast to 39% for a CC recipient receiving a liver from a CT donor [104]. Lange et al [106] reported slightly different observations, with only one patient who received a liver from a non-CC donor not achieving SVR; they suggested the beneficial

donor genotype may compensate for the recipient's poor predictors of response. This contrasts with the previous examples, in which donor and recipient genotypes appear to have an additive effect.

Studies into graft versus host disease (GVHD) have been performed for a number of years, with the intention of identifying biomarkers correlating with a phenotypic effect. GVHD is a rare but severe complication following liver transplantation, and has a high mortality rate due to the graft tissue recognizing the recipient as 'foreign' [107, 108]. The impact on GVHD with SNPs in a range of both proinflammatory and anti-inflammatory cytokines has been assessed, and are reviewed by Markey *et al* [109]. In short, both the donor tissue and host recipient generate cytokines, and therefore each contribute to the pro- and anti-inflammatory response following transplantation. Further studies, including that with IFN $\lambda$ 3 SNPs, will eventually lead to a more personalized donor/recipient match, thereby reducing the likelihood of graft rejection or GVHD.

#### **1.4: Association of Interferon Lambdas in non-HCV Settings**

A number of groups have looked for an association between IFN $\lambda$ 3 genotype and disease progression in HBV infection [110-112]. Severity of illness or treatment response in chronic hepatitis B patients is not associated with IFN $\lambda$ 3 genotype [111, 113], nor is there a difference between HBV infected individuals and healthy controls [111]. There is however, an association between IFN $\lambda$ 3 levels in serum and chronicity; those with a self-limiting HBV infection or healthy controls had a higher level than those with a chronic infection [112].

In non-alcoholic fatty liver disease (NAFLD), patients with an IFN $\lambda$ 3 rs12979860 CC genotype had more severe liver damage than those with a non-CC genotype [114]. This study excluded patients infected with HCV, thereby bypassing the previously reported associations between HCV treatment response and IFN $\lambda$ 3 genotype. Interestingly, there was no significant correlation between severity of liver disease and IFN $\lambda$ 3 rs8099917. The association seen with rs12979860 was further enhanced by the presence of phospholipase domain-containing protein (3PNPLA3) rs738409 GG. This 3PNPLA3 SNP has been shown to be associated with hepatic fat content in both alcoholic liver disease [115] and NAFLD [116].

In the setting of HIV, IFN $\lambda$ 3 CC is related to a higher all-cause mortality rate than non-CC patients. Parczewski et al [117] reported this observation in a cohort of 507 Caucasian individuals, including HIV mono-infected, and HIV/HCV co-infected patients. This observation was made following the initiation of anti-retroviral therapy, and does not hold true prior to the start of treatment. When restricted to AIDS-related deaths, the trend still applies, but there is no statistical significance between death rates and IFN $\lambda$ 3 genotype.

A correlation between IFN $\lambda$ 3 genotype and childhood disease has been found; Gaudieri et al [118] reported that the carriage of the T allele of the rs12979860 SNP increases the likelihood of developing allergies, including asthma, and food allergies.

The TT genotype at rs8099917, which confers a good response in terms of HCV clearance, has been associated with an increased risk of schizophrenia in the Chinese Han population [119]. Although there are probably multiple causative agents for schizophrenia, this

observation may give some credence to suggestions there is a viral association with the disease [120].

### **1.5: Summary**

IFN $\lambda$ 3 genotype is associated with modulating the inflammatory response to a number of stresses. The fact that there is a clear association between various SNPs and successful treatment for HCV infection has allowed routine SNP genotype testing to become a useful tool when planning a treatment regime. The host genetic information combined with the viral genotype has provided a diagnostic tool, which in the era of pegIFN $\alpha$ /RBV therapy has helped patients to either commit to undergoing treatment, or to delay until better treatments became available. More recently, with the introduction of direct acting antivirals, IFN $\lambda$ 3 genotype may be of reduced relevance in economically advantaged countries with easy access to new drugs with higher response rates. However, information gained from IFN $\lambda$ 3 genotyping and genotyping for human leukocyte antigen C and its ligands, the killer immunoglobulin-like receptors, has been shown to enhance prediction of clearance [121]. Additionally, IFN $\lambda$ 3 genotyping can be used in combination with other non-invasive markers to develop an algorithm which predicts rate of fibrosis and cirrhosis progression [122]. Alternatively, IFN $\lambda$ 3 genotyping may remain relevant in some countries to treat those individuals infected with genotype 2 or 3 with a pegIFN $\alpha$ /RBV treatment regime as a matter of cost-effectiveness.

## **Chapter 2: Validation of an Interferon Lambda 3 Allelic Discrimination Assay**

### **2.1: Introduction**

Successful nucleic acid extraction is a vital component of molecular testing and is now routinely performed in diagnostic and research facilities. A number of commercial kits are available for both RNA and DNA extraction for several different sample types. Commercial kits have allowed for better quality assurance and reproducibility than historical methods applied from research laboratory based protocols. Some methods are labour intensive and technically demanding, whilst automated extraction methods have largely reduced hands-on time while maintaining a high quality throughput. Many extraction methods are best suited to specific applications. It is important to assess the efficiency of an extraction protocol on each specimen type, as well as on each analyte. Factors including yield, specificity, turn-around time, and ease of use must be considered before implementing an extraction protocol for specific PCR applications.

Various studies have shown that successful genotyping of polymorphisms within the human genome require DNA to be extracted from cellular blood components, such as the buffy coat. Buffy coat, containing most of the white blood cells, is visible after centrifugation of anticoagulated blood and is found as a thin layer between the denser red blood cells and plasma fraction. These cells contain high numbers of nucleated leukocytes; when the cells are lysed host chromosomal DNA is easily accessible. This makes the white blood cells a suitable substrate for human single



nuclear polymorphism (SNP) analysis. However, in some instances, buffy coat cannot be obtained due to blood collection protocols (e.g. serum instead of plasma) and further processing at institutions (e.g. specimen reception spins samples but only removes the plasma fraction).

The aim of this chapter was to assess the practicality and efficiency of using plasma or serum samples to determine the IFN $\lambda$ 3 rs12979860 genotype. If plasma and serum specimens contain sufficient genomic DNA, the need to collect, process and store buffy coat specimens for IFN $\lambda$ 3 genotyping may be eliminated.

Plasma and serum samples are simple to process, and less blood is required to be collected from patients. These are typically the specimens of choice for other HCV related assays, such as viral load and genotyping, meaning that no additional or special collection is necessary.

There are a range of methods for the extraction of DNA from blood and body fluids, including silica matrices and magnetic bead-based purification, many of which are commercially available. It is essential to evaluate which extraction method is most suited to the downstream application. This study investigated the efficiency of four different extraction kits; two were used for extracting DNA from buffy coat samples, and three for plasma samples. A spin column extraction method (QIAGEN Melbourne, Australia; cat # 51306) was chosen for both buffy coat and plasma extractions, as they are commonly used in the laboratory, and are known to be a reputable product. For the extraction of DNA from buffy coats, a protein precipitation method was included, to determine whether this significantly impacted the quality of DNA output (Quantum Scientific, Brisbane, Australia; cat #2300720). In

the case of extracting DNA from plasma samples, two additional extraction methods currently in use for other applications within the laboratory were chosen due to being different methodologies. The Corbett (Qiagen) X-tractor gene uses an automated vacuum system (Qiagen, Melbourne, Australia; cat # 950107), while the Abbott m2000sp uses an automated magnetic bead application (Abbott, Sydney, Australia; cat# 06K12-24). The potential benefit of the latter two plasma extraction methods is the number of specimens which are able to be processed simultaneously, on an automated platform; the Corbett X-tractor gene is capable of extracting 96 samples within a run, while the Abbott m2000sp can process 48.

Additionally, the longevity and viability of using stored plasma and serum samples for IFN $\lambda$ 3 genotyping was assessed, to determine whether stored samples are suitable for retrospective testing.

## **2.2: Methods**

### **2.2.1: Sample Selection and Extraction Methods**

Sixty samples were selected for analysis due to having received multiple EDTA tubes from single time points, allowing for two collections of the buffy coat fraction of the blood, as well as the plasma component. Plasma was collected from the EDTA tube and stored in a 2ml tube at -20°C until required. Using the same transfer pipette, the buffy coat was collected and stored in an additional 2ml tube in a dedicated box at -20°C. An additional 14 plasma specimens were included in the extraction comparison, to make a total of 74 for that component.

**2.2.1.1: Buffy Coat Extraction: Qiagen DNA MiniKit**

Stored buffy coat samples were thawed and DNA was extracted using the Qiagen QIAamp DNA MiniKit (QIAGEN, Melbourne, Australia; cat # 51306) according to manufacturer's instructions, with an input volume of 200µl. The extraction incorporates a spin column with a silica membrane; following cell lysis, the DNA bound to the matrix undergoes a series of washes prior to eluting in 100µl elution buffer.

**2.2.1.2: Buffy Coat Extraction: ArchivePure Blood Kit**

DNA was extracted from buffy coat samples using the ArchivePure DNA Blood Kit (Quantum Scientific, Brisbane, Australia; cat #2300720) according to the manufacturer's protocol. In short, 1ml of whole blood was added to a cell lysis solution, followed by protein precipitation, and then DNA recovery with 100µl of a DNA hydration solution.

**2.2.1.3: Plasma Extraction: Qiagen DNA MiniKit**

Stored plasma was thawed and DNA was extracted using Qiagen QIAamp DNA MiniKit as described above in 2.2.1.1.

**2.2.1.4: Plasma Extraction: Corbett (Qiagen) X-tractor**

DNA extraction from plasma was performed on the Corbett X-tractor according to manufacturer's directions, using the Corbett (Qiagen) DX reagent kit (Qiagen, Melbourne, Australia; cat # 950107). Aliquots of 180µl of plasma were dispensed

into the lysis plate prior to initiation of the extraction. After a series of lysis and wash steps, the final eluate (100µl) is captured in a 96-well elution plate; the eluate was removed and placed in 1.5ml micro centrifuge tubes for storage purposes.

#### **2.2.1.5 Plasma Extraction: Abbott m2000sp**

Plasma samples were processed in the Abbott m2000sp automated extraction system with the Sample Preparation System<sub>DNA</sub> kit (Abbott, Sydney, Australia; cat# 06K12-24) according to manufacturer's directions. The system utilizes a magnetic bead method, with an initial input of 200µl of plasma. The final elution of 100µl is aliquoted into a 96-well plate; the eluate was removed and stored in 1.5ml microcentrifuge tubes until required.

#### **2.2.3: IFNλ3 Genotyping and Allelic Discrimination**

IFNλ3 rs12979860 genotyping was performed using custom designed TaqMan probes as described in Tillman *et al* [90], and is ordered from Life Technologies (Melbourne, Australia). A single tube containing the combined primer and probe mix (Table 2.1) is diluted to a working concentration of 900nM for the primers, and 200 nM for the probes. The C allele is detected via the VIC fluorescent dye, with a peak absorption wavelength of 538nm, and the T allele is detected via the FAM dye, which has a peak absorbance wavelength of 494nm.

Forward Primer	GCCTGTCGTGTACTGAACA
Reverse Primer	GCGCGGAGTGCAATTCAAC
VIC probe (C allele)	TGGTTC <u>G</u> CGCCTTC
FAM probe (T allele)	CTGGTTC <u>A</u> CGCCTTC

**Table 2.1: Primer and Probe Sequences used in IFN $\lambda$ 3 Genotyping Assay.**  
**C allele is detected with VIC probe, and T with FAM probe. The SNP is underlined and italicised.**

The master mix consists of 12.5 $\mu$ l TaqMan Genotyping MasterMix (Life Technologies, Melbourne, Australia; cat# 4371355) and 1.25 $\mu$ l probe/primer mix, with 11.25 $\mu$ l DNA sample/control from various extraction methods (see 2.2.1).

The assay was performed on Applied Biosystems 7500 Fast Real Time System, using SDS software, version 1.4. The allelic discrimination consists of a 'pre-read' to analyse baseline fluorescence (60°C, 1 minute), an amplification step (10 minutes at 95°C, followed by 40 cycles of 92°C for 15 seconds and 60°C for 1 minute). Amplification curves and ct values (i.e., the cycle threshold at which fluorescence is detected) are analysed for each well.

A 'post-read' for determining genotype (60°C for 1 minute) was performed following the amplification step. The background fluorescence from the pre-read is subtracted from the post-read data to obtain a result. A genotype is automatically assigned to each sample; when an allele is detected, the ct is recorded, and if it is not detected, it will state 'undetermined'. Therefore, all heterozygotes will have a ct for each

allele, while homozygotes will have a ct recorded for one allele, with 'undetermined' noted in the channel for the alternative allele. In addition, the automatic determination of alleles is made with a quality assurance of 90% or greater; any samples which fall below the 90% threshold is considered 'undetermined'.

#### **2.2.4: White Blood Cell Count**

Four randomly selected samples were selected for analysis using the Coulter LH500. Each sample was processed so four components, each of 300µl, could be analysed separately; an aliquot of whole blood was taken, then the tube re-spun to enable collection of plasma, buffy coat, and finally red fraction following removal of other components. Saline was added to the buffy coat fraction, as it was too viscous for the instrument; this was taken into account when calculating cell counts.

The probe on the instrument was inserted into the tubes containing the various samples, and the blood counts were compared across each component.

#### **2.2.5: Longevity of Genomic DNA**

Twenty-four plasma samples which had been tested for IFNλ3 genotype on initial arrival were stored for two years at -20°C. The DNA from the 24 stored samples was then re-extracted using the QIAamp DNA MiniKit, and re-tested for IFNλ3 genotype to determine if there had been any deterioration in the quality of the genomic DNA.

#### **2.2.6: Freeze-Thaw Analysis**

Three plasma samples were selected to undergo multiple freeze-thaw cycles to determine the effect on the IFNλ3 allelic discrimination assay. The samples were

selected to include one of each genotype; a CC, CT, and a TT. Each plasma sample was aliquoted into ten 2ml tubes, and then placed in a -20°C freezer. One aliquot of each genotype would remain frozen until the completion of the freeze-thaw cycles. The remaining nine aliquots for each were completely thawed, and then returned to the freezer, resulting in two freeze-thaw cycles. This process of freezing and thawing was repeated until there was a representative sample for each genotype up to and including 10 cycles. Once the multiple freezing and thawing had completed, the DNA was extracted using the Qiagen DNA MiniKit, and subsequently assayed for IFN $\lambda$ 3 genotype.

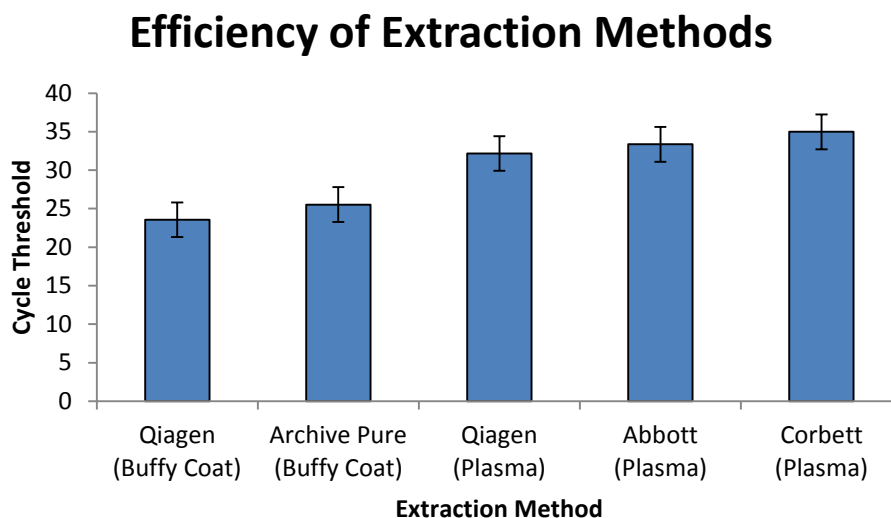
## **2.3: Results**

### **2.3.1: Sample Extraction Comparisons**

To evaluate the efficiency of various extraction methods, comparative studies were performed using four different kits. Both buffy coat extraction methods worked effectively; the average ct values from the QIAamp kit were slightly lower than the ArchivePure kit (22.77 and 24.07 respectively) (Figure 2.1), indicating better efficiency. The automatic allelic determination was superior using extracted material from the QIAamp kit, with a 94% efficiency, compared with 84% for the Archive Pure method.

To test whether a sufficient amount of genomic DNA can be obtained from plasma samples, the corresponding plasma component from the buffy coat extractions were tested (section 2.2.1). Of the three plasma extraction kits, the QIAamp kit proved the most efficient, followed by the Abbott, then the Corbett kit (average ct values are

32.39, 34.11, and 34.89, respectively)(Figure 2.1) ). The Corbett extraction resulted in a poor amplification profile, and also failed to assign genotypes to 49% of the samples in the allelic discrimination analysis. The Abbott and QIAamp kits were comparable (97% and 100% allele call efficiency, respectively). The Qiagen method performed best across each sample type; buffy coat extractions gave a superior signal to those of the plasma extractions. The average ct between DNA from buffy coat samples and plasma samples differed by 10 cycles (Figure 2.1). The assay is not quantitative, but demonstrates that plasma derived genomic DNA can still be detected, which indicates there is sufficient genomic DNA within these blood components.



**Figure 2.1: Average ct values for each extraction method. The lower the ct value, the more efficient the extraction.**

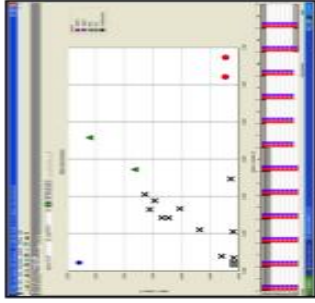
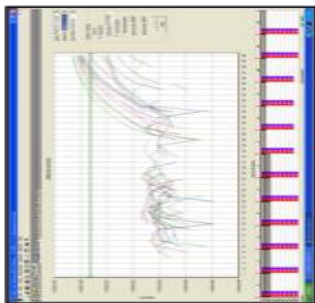
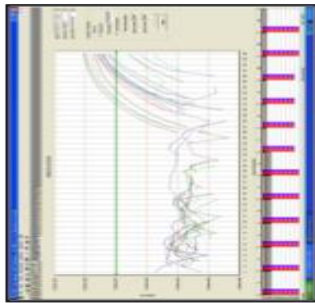
The output from the ABI 7500 allows a graphical view of both the amplification plot and the allelic discrimination plot (Figure 2.2), allowing a clear visual comparison between the various extraction methods. Samples are represented on the plot by a



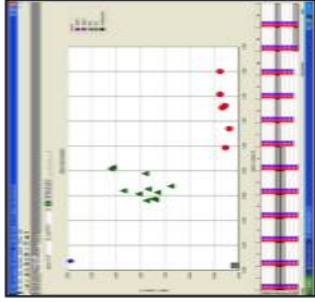
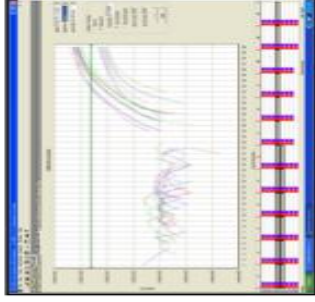
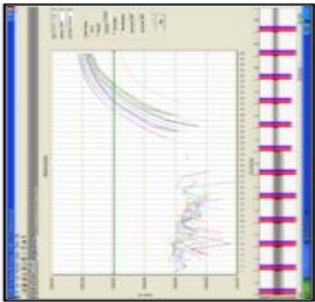
blue diamond along the Y-axis if only a T allele is detected, or as a red circle along the X-axis if only a C allele is detected. In the heterozygote samples, a green triangle is used to show where both alleles are detected.

**Figure 2.2 (next page): Allelic Discrimination and Real time PCR for the identification of IFN $\lambda$ 3 genotype in samples derived from different specimen types and extraction methods. The sample type and extraction method are indicated. Amplification curves of C allele (A) and T allele (B), and allelic discrimination plot (C) are viewed separately but analysed together. Samples for which genotypes can not be assigned are depicted with an 'X' on the allelic discrimination plot.**

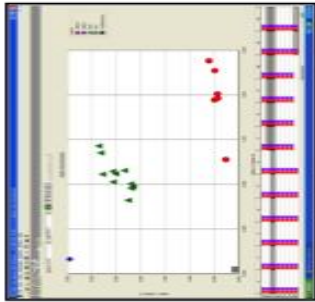
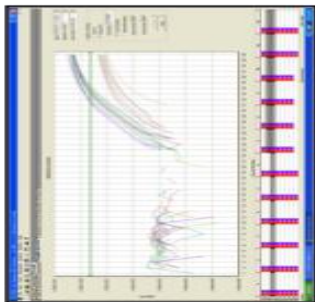
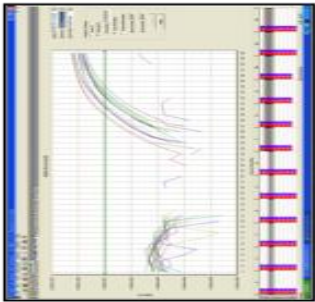
Corbett  
Plasma



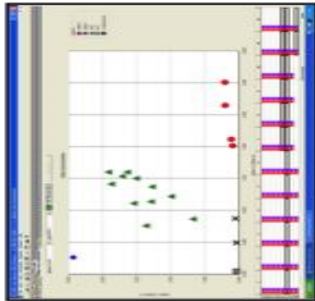
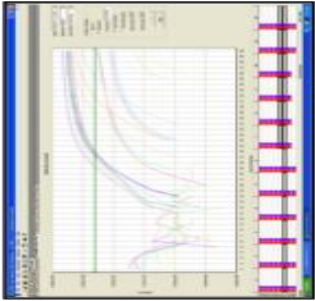
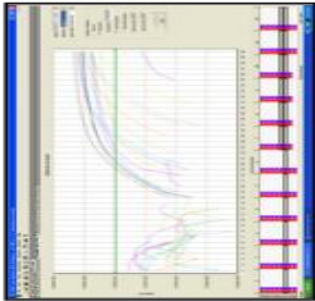
Abbott m2000  
Plasma



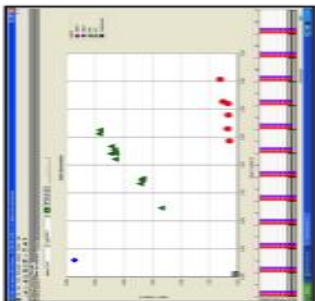
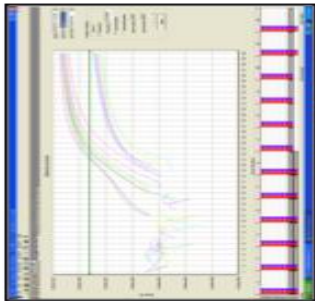
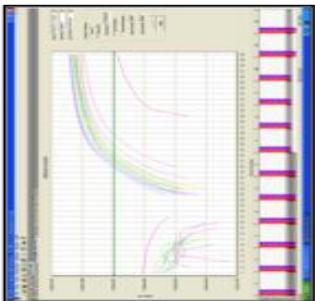
Qiagen DNA Mini Kit  
Plasma



Archive Pure  
Buffy Coat



Extraction Kit: Qiagen DNA Mini Kit  
Specimen Type: Buffy Coat



C Allele Curve

A

T Allele Curve

B

Allelic  
Discrimination  
Plot

C

### 2.3.2: White Blood Cell Count

A small pilot study consisting of four samples was carried out to compare white blood cell count across different fractions of the blood, with each component containing a total volume of 300µl. As demonstrated in Table 2.2, the buffy coat contains 50 -100 times more leukocytes than the plasma, showing that although the numbers are reduced, white blood cells are still present at a measurable concentration in the plasma fraction.

Sample ID	White Blood Cell Count (x10 <sup>9</sup> /L)			
	Whole Blood	Plasma	Buffy Coat	Red
13551133	3.6	0.1	4.4	3.2
13551299	5.3	0.1	5.3	0.6
13551301	5.5	0.1	11.2	0.9
13551305	7.9	0.1	8.9	1.5

**Table 2.2: White blood cell count in four components of blood samples for four IFNλ3 specimens.**

### 2.3.3: Longevity of Genomic DNA

In order to test the longevity of the samples for IFNλ3 testing, DNA was extracted from 24 plasma samples (see section 2.2.1.3) which had been previously tested, and then stored at -20°C for two years, and retested for IFNλ3 genotype. All genotype results were concordant with the initial testing, performed two years prior on the

fresh samples. The results, shown in Table 2.3, showed there was very little deviation in ct values, with the average difference being 0.015 cycles. This demonstrates that plasma samples can be used to retrospectively test for IFN $\lambda$ 3 genotype, as the genomic DNA within the sample does not degrade significantly over the time period.

Well	Sample ID	Allele	2011	2013	Genotype
			Ct	Ct	
A1	11513176	C ALLELE	33.6822	33.0492	CC
A1	11513176	T ALLELE	Undetermined	Undetermined	
A2	11513181	C ALLELE	Undetermined	Undetermined	TT
A2	11513181	T ALLELE	32.9624	34.1868	
A3	11513556	C ALLELE	32.7041	32.0879	CC
A3	11513556	T ALLELE	Undetermined	Undetermined	
A4	11513557	C ALLELE	31.9768	32.0194	CC
A4	11513557	T ALLELE	Undetermined	Undetermined	
A5	11516607	C ALLELE	29.6622	28.9323	CC
A5	11516607	T ALLELE	Undetermined	Undetermined	
A6	11513606	C ALLELE	32.9587	33.3687	CC
A6	11513606	T ALLELE	Undetermined	Undetermined	
A7	11513672	C ALLELE	36.9156	34.8106	CT
A7	11513672	T ALLELE	35.3148	33.2847	
A8	11513674	C ALLELE	33.902	35.7937	CT
A8	11513674	T ALLELE	32.3419	34.5321	
A9	11513677	C ALLELE	35.3132	34.4824	CC
A9	11513677	T ALLELE	Undetermined	Undetermined	
A10	11513680	C ALLELE	32.8528	33.1123	CT
A10	11513680	T ALLELE	31.3742	31.6598	
A11	11514649	C ALLELE	Undetermined	Undetermined	TT
A11	11514649	T ALLELE	28.1759	28.2793	
A12	11514651	C ALLELE	34.0186	33.9833	CT
A12	11514651	T ALLELE	33.0004	32.6205	

**Table 2.3: Effect of long-term storage of plasma at -20°C for IFN $\lambda$ 3 SNP testing. Samples were re-extracted after two years storage, and the allelic discrimination assay was performed again. The ct values were compared between timepoints, and differed by an average of 0.015 cycles. NOTE: This table only details 12 of the 24 samples tested.**

As buffy coat specimens are occasionally used as a back-up sample, the effect of long-term storage on these specimens was also assessed. As used in the initial

testing, the extractions were performed using the Qiagen MiniKit, with an input volume of 200µl (see section 2.2.1.1). The ct values were almost identical after extracting and genotyping after two years storage at -20°C (Table 2.4). Two samples run in 2011 had higher ct values than expected for buffy coat specimens (depicted in red text in Table 2.4). Once these outliers were removed from analysis, the average ct difference was 0.7792 cycles. As was the cases in the plasma samples, the quality of the genomic DNA within the buffy coat samples did not degrade substantially over time.

	Lab ID	Allele	ct (2011)	ct (2013)
1	10593117	C	31.6947	24.0016
		T	32.2862	22.8353
2	10593500	C	22.7119	23.5658
		T	23.275	22.3154
3	11500713	C	23.5137	23.5197
		T	22.6852	22.2861
4	11501109	C	23.1326	21.2913
		T	Undetermined	Undetermined
5	11501656	C	25.9342	22.5403
		T	Undetermined	Undetermined
6	11503042	C	22.7237	23.4867
		T	22.2805	22.3247
7	11503045	C	20.6816	21.2483
		T	Undetermined	Undetermined
8	11504607	C	35.5949	21.5049
		T	Undetermined	Undetermined
9	11505020	C	22.8692	22.3373
		T	Undetermined	Undetermined
10	11505425	C	25.8352	23.3093
		T	24.0822	22.1497

**Table 2.4: Ct values of buffy coat derived DNA samples following two years storage at -20°C. Represented in red are the two samples in which an unusually high ct value was given in the initial testing.**

**2.3.4: Effect of Multiple Freeze-Thaw Cycles**

To assess the impact of multiple freeze/thaw cycles, three plasma samples were selected (see section 2.2.6). These samples had been previously characterized as CC, CT, or TT IFN $\lambda$ 3 genotype, then stored in 10 separate 200 $\mu$ l aliquots, and subjected to up to 10 freeze/thaw cycles, and re-tested. These samples were designated CC1-CC10, CT1-CT10, and TT1-TT10, with the number reflecting the number of freeze/thaw cycles the aliquot has undergone. For the IFN $\lambda$ 3 CT and TT samples, the ct value was the same across the freeze/thaw cycles, whereas the ct value of the CC specimen improved slightly. These results showed that multiple freeze/thaw cycles do not affect integrity of the sample in the IFN $\lambda$ 3 assay (Table 2.5).

Sample ID	Allele	Result	Sample ID	Allele	Result	Sample ID	Allele	Result
CC1	C ALLELE	28.9135	CT1	C ALLELE	35.9027	TT1	C ALLELE	Undetermined
CC1	T ALLELE	Undetermined	CT1	T ALLELE	34.8055	TT1	T ALLELE	30.7038
CC2	C ALLELE	29.0145	CT2	C ALLELE	36.1337	TT2	C ALLELE	Undetermined
CC2	T ALLELE	Undetermined	CT2	T ALLELE	34.5696	TT2	T ALLELE	33.039
CC3	C ALLELE	28.1231	CT3	C ALLELE	35.8878	TT3	C ALLELE	Undetermined
CC3	T ALLELE	Undetermined	CT3	T ALLELE	34.9854	TT3	T ALLELE	32.0159
CC4	C ALLELE	28.5729	CT4	C ALLELE	36.4089	TT4	C ALLELE	Undetermined
CC4	T ALLELE	Undetermined	CT4	T ALLELE	34.4191	TT4	T ALLELE	32.2521
CC5	C ALLELE	25.5143	CT5	C ALLELE	35.4937	TT5	C ALLELE	Undetermined
CC5	T ALLELE	Undetermined	CT5	T ALLELE	34.4365	TT5	T ALLELE	31.5172
CC6	C ALLELE	Undetermined	CT6	C ALLELE	35.5543	TT6	C ALLELE	Undetermined
CC6	T ALLELE	Undetermined	CT6	T ALLELE	34.7895	TT6	T ALLELE	32.0766
CC7	C ALLELE	23.8484	CT7	C ALLELE	35.3994	TT7	C ALLELE	Undetermined
CC7	T ALLELE	Undetermined	CT7	T ALLELE	34.75	TT7	T ALLELE	31.4647
CC8	C ALLELE	23.6587	CT8	C ALLELE	35.6215	TT8	C ALLELE	Undetermined
CC8	T ALLELE	Undetermined	CT8	T ALLELE	35.694	TT8	T ALLELE	29.6987
CC9	C ALLELE	24.0679	CT9	C ALLELE	Undetermined	TT9	C ALLELE	Undetermined
CC9	T ALLELE	Undetermined	CT9	T ALLELE	Undetermined	TT9	T ALLELE	31.2967
CC10	C ALLELE	Undetermined	CT10	C ALLELE	38.0411	TT10	C ALLELE	Undetermined
CC10	T ALLELE	Undetermined	CT10	T ALLELE	35.8012	TT10	T ALLELE	Undetermined

Table 2.5 Ct values following freeze-thaw cycles:

CC1-CC10, CT1-CT10, and TT1-TT10 represent IFN $\lambda$ 3 CC, CT, and TT samples respectively, undergoing 1-10 freeze/thaw cycles. For example; CC1 is the CC sample which has had 1 freeze/thaw cycle; CC10 is the CC sample which has had 10 freeze/thaw cycles.

The cycle threshold remained virtually the same across the freeze-thaw cycles, showing the integrity remains the same.

## **2.4: Discussion**

### **2.4.1: Extraction Methods**

Extraction with the Qiagen DNA MiniKit was the most efficient method to isolate chromosomal DNA for either the buffy coat samples or plasma samples (Figure 2.1). This is the extraction method presently used most widely in our laboratory. It is easy to use, and as specified in the kit insert, is able to be used for a variety of sample types including plasma, serum, cerebrospinal fluid, urine, and tissues. This makes it a highly versatile extraction method, and many laboratory members are competent users of the kit.

The ArchivePure extraction kit gave comparable results to the Qiagen extraction method (Figure 2.1), however, it was more time consuming, and required additional plasticware. The kit is only suitable for the extraction of DNA from blood products, meaning it could not be utilized for other sample types commonly tested within the laboratory, such as urine and CSF samples used in different applications. Whilst this does not impact the result on the IFN $\lambda$ 3 genotype determination, it does play a role in cost-effectiveness to a diagnostic laboratory.

The Abbott m2000sp gave comparable results to the Qiagen method for plasma samples (Figure 2.1). The instrument, which uses magnetic bead technology, has the capability of extracting up to 48 samples at a time, taking approximately 3.5 hours. The benefit of using this machine rather than the Qiagen is the reduction in hands-on time; using the manual Qiagen kit requires the user to be in the laboratory for the entire procedure, aside from a 10 minute incubation period, but the extraction can comfortably be performed in 90 minutes



for 24 samples. However, a disadvantage of using the Abbott m2000sp is that the extraction kits are designed for batches of 24 reactions; this reduces the cost-effectiveness of the extraction, as any extractions performed that are not a multiple of 24 results in a waste of reagents. The Abbott m2000sp is primarily used as an extraction platform for the quantitative analysis of a range of viruses, and the extraction kit is one component of the total reagents.

The final extraction method utilized in this study was the Corbett X-tractor (Qiagen). This platform had the greatest potential; the automated extraction is carried out on a 96-well plate format. With a brief hands-on time required to aliquot the samples into the lysis plate, and load the instrument with the plastic ware and reagents, it takes less than two hours to obtain the DNA from 96 patient samples. It uses a vacuum extraction method, on a silica capture plate. The final elution is dispensed into removable tubes as part of a 96-well plate. This means a multi-channel pipette may be utilized to transfer DNA from the elution plate into the PCR plate. Due to the potential advantages of using this platform, it was disappointing to find it was the least efficient of the methods. With a successful allelic determination rate of only 50% (Figure 2.2), it was shown that this extraction method was unsuitable for IFN $\lambda$ 3 genotype testing.

Although genomic nuclear DNA is commonly isolated from buffy coat or tissue samples, this study has shown that plasma samples can be used to determine IFN $\lambda$ 3 genotype. The knowledge that cell free DNA and RNA circulates in plasma is not new, and has been reported as early as 1948 [123]. There was, however, little research in the following 30 years involving circulating nucleic acids [124], perhaps as a result of the use of more

intrusive procedures required for diagnostic testing of different diseases. Circulating cell free DNA is recognized as a useful resource for testing of cancer patients, with varying explanations as to its presence. While some researchers suggest it is primarily a result of apoptosis, others have the view it is due to cell secretion [125].

Cell free DNA has been utilized in more recent years as an alternative to invasive procedures such as amniocentesis and chorionic villus sampling during pregnancy. Maternal plasma and serum carry cell-free foetal DNA, allowing for rhesus blood type, and foetal gender to be accurately determined [126]. Single gene disorders are also able to be identified, on condition that the mutation is not carried by the mother [127].

Given the widespread use of plasma and serum samples as a diagnostic tool in cancer patients and in prenatal testing, it is not surprising that these sample types are suitable in determining the IFN $\lambda$ 3 genotype. Although a small number of samples were analysed to determine the number of white blood cells within the plasma fraction of the blood (Table 2.1), the fact that cell-free DNA circulates throughout the plasma and serum makes determination of the white blood cell number in these samples less important.

Although the majority of samples received by our laboratory for IFN $\lambda$ 3 testing are plasma or whole blood samples, serum samples are not uncommon. These serum samples have been successfully used to determine IFN $\lambda$ 3 genotype (data not shown), which is in agreement with the aforementioned observations. It could perhaps be presumed that serum would be a less successful specimen type than plasma due to the lack of cells. However, it appears the opposite is true. Lee et al [128] showed circulating DNA levels in healthy blood donors were

higher in serum compared to plasma, and the level in serum actually increases after five days storage at 4°C. They also showed that the observed difference between the blood components was due to the clotting process; as the white blood cells lyse, nuclear fragments are released into the serum.

Buffy coat specimens contain higher concentrations of nuclear DNA in comparison to plasma, as made apparent by the earlier ct value observed. As the assay is to determine the presence of alleles, and is not reliant on any quantification, this study shows that buffy coat and plasma samples are both viable specimen types for determining IFNλ3 genotype. There are a number of advantages associated with the use of plasma or serum samples as an alternative to buffy coat for this assay. It eliminates the need to collect multiple blood tubes, as well as the extra steps required to separate buffy coat from the whole blood. It also allows for the one primary sample tube may be used for a range of diagnostic analyses, including antibody testing, HCV viral load, HCV genotype, and IFNλ3 genotype.

Overall, the results identified that the most efficient way to extract DNA for IFNλ3 genotype analysis was to use the Qiagen Minikit from plasma samples. It employs a widely used extraction method, meaning new techniques are not required beyond the scope of the normal laboratory operations. It does not necessarily require any specialized equipment, as it can utilize standard laboratory equipment.

### 2.4.2: Longevity Studies

DNA was extracted from twenty four plasma samples and ten buffy coat samples which had been frozen for two years at -20°C, and subsequently retested to determine IFN $\lambda$ 3 genotype. The allelic discrimination software was able to reliably determine the genotype for each of the samples. In order to assess any degradation of the sample, the ct values were also compared with the value from when the sample was originally tested. It was found that the cycle at which fluorescence was detected remained essentially the same (Table 2.3).

Biological samples stored for a long period of time have often been used for retrospective testing. Madisen et al [129] observed that whole blood stored at -70°C for up to eight weeks was able to yield high molecular weight DNA comparable to DNA extracted from fresh whole blood samples. The group had more extensive data showing that DNA extracted from fresh blood was able to be stored for 11 years without showing significant degradation. Plasma samples stored for nine years at either -70°C or -20°C and retested for HBV DNA, HCV RNA, and HIV-1 RNA showed only a slight reduction in viral load; the difference was similar to what is seen with intra-assay variability, and therefore not clinically significant [130]. In addition, buffy coat specimens have been shown to be stable for 9 years at -80°C, and remain suitable for genetic testing without major degradation [131]. This study has shown that the same applies to a qualitative assay assessing human genomic DNA, with negligible deviation between testing time-points, as assessed by ct values.

The observation that the genomic DNA required to test for IFN $\lambda$ 3 remains in the plasma samples over a two-year period allows for the retrospective testing of patient samples.

There are a number of instances in which retrospective testing may be useful. In a clinical setting, if a patient has failed treatment, retrospective testing from a previous bleed is an option. Another benefit is that if a sample has become compromised, in transit for example, there is the option to use a stored sample that has perhaps been received for other tests. It may eliminate the need to recall a patient for extra blood to be drawn.

IFN $\lambda$ 3 genotyping has been helpful in indicating why some patients in past clinical trials may have responded to treatment whereas others did not. It also allows researchers to view viral clearance from a transmission case in a new light. For example, there are two well publicized reports where women have been infected with HCV by contaminated anti-D blood products in Germany in 1978-79 [90] and in Ireland in 1977-78 [132]. In these two incidents, a proportion resolved their infection similar to previously reported clearance rates. Within each cohort, the individuals contracted the virus from a single source; the elimination of variance due to viral strain allowed researchers to observe disease progression at the time. With the advent of IFN $\lambda$ 3 SNP testing, the cohorts were able to be further scrutinized as to why some individuals cleared the virus and others remained chronically infected. Fresh blood was obtained following consent for the testing of women in these cohorts. The results presented here have identified IFN $\lambda$ 3 testing can be successfully carried out after storage of blood products for two years, however, results could not be extrapolated to storage for 30 years. Nonetheless, in circumstances where the patient has perhaps been infected due to an adverse event, it may cause them some distress to be recalled to give blood for further genetic testing. In examples like this, retrospective testing of IFN $\lambda$ 3 genotype on stored plasma samples (such as from the initial HCV diagnosis) would be advantageous.

A great benefit of this study is that as more host DNA markers are discovered in the future, we have identified that any stored plasma samples are able to be tested. This allows for linkage studies, as well as investigation why some phenotypes are observed without the need to collect further samples.

#### **2.4.3: Effect of Multiple Freeze-Thaw Cycles**

For many applications, plasma and serum samples are stored in either -70°C or -20°C freezers following sample arrival. The long term viability of IFN $\lambda$ 3 genotype relating to prolonged storage has been evaluated; the next step in the study was to determine the viability after multiple freeze-thaw cycles. Samples commonly undergo freeze and thaw cycles in normal laboratory testing, particularly in the setting of an institution that has long term storage of specimens for which multiple requests have been received. Multiple freeze-thaw cycles may be avoided by storing multiple aliquots of each specimen upon arrival. While this may be an option in serum and plasma banks in research facilities, it is largely impractical for many institutions due to the size of the facility and the equipment required.

Often, the characterization of a patient sample includes sequential requests, which includes freeze/thaw steps. A common request at initial diagnosis is to perform HCV antibody testing; if this returns a positive result, a PCR is performed. If this is in turn positive, HCV genotyping is requested, and then, if clinically relevant, an IFN $\lambda$ 3 genotype is requested. In instances such as this, thawing and refreezing specimens is largely unavoidable. This analysis has shown that multiple freeze-thaw cycles do not affect the stability of the genomic DNA

within a plasma sample. Three plasma samples, one of each IFN $\lambda$ 3 genotype, were each separated into ten 200 $\mu$ l aliquots, and underwent between 1-10 freeze-thaw cycles, with no significant effect.

Various studies have been performed on the effect of freeze-thaw cycles on a number of analytes. Serum and plasma from ten healthy volunteers was subjected to up to ten freeze-thaw cycles, and analysed for cholesterol, micronutrients, and sex-specific and sex-neutral hormones; no significant effects were identified [133]. This confirmed earlier studies using baboon sera that showed cholesterol is stable following multiple freeze-thaw cycles [134]. More recently, further analytes including ALT, AST, cholesterol and glucose have also been shown to be stable at multiple freeze-thaw cycles [135].

In conclusion, this study has identified that plasma samples are suitable specimen types for determination of IFN $\lambda$ 3 genotype, and is preferable to buffy coat samples in a diagnostic laboratory setting where plasma samples are routinely used for multiple assays. The longevity component and the multiple freeze-thaw analysis have demonstrated that the plasma samples are able to withstand a range of storage and handling procedures without a detrimental effect in identifying IFN $\lambda$ 3 genotype. This is a benefit in both routine diagnostic laboratory practices, as well as in retrospective studies.

## Chapter 3: Optimisation of the IFN $\lambda$ 3 Assay

### 3.1: Introduction

An in-house IFN $\lambda$ 3 rs12979860 genotype assay has been validated, using the TaqMan allelic discrimination protocol, with plasma as the preferred specimen type (see Chapter 2). Optimisation procedures for many laboratory-based tests often have to be implemented to ensure sensitivity, specificity, cost effectiveness, efficiency, and practicality. A number of different approaches may have to be undertaken, including modification of PCR cycling parameters, altering volumes of reagents or template, using products from different manufacturers, and evaluating commercial kits with in-house developed assays. Subsequent to the validation of the in-house custom designed single nucleotide polymorphism (SNP) analysis previously described, alternative reagents have been developed and have become available commercially. With the view of working in a narrow timeframe, and optimizing outcomes, critical parameters were modified to identify the best conditions. This included altering volumes of reagents and template, and assessing different reagents.

Additionally, a commercial assay, corresponding to the validated TaqMan in-house assay, has been released for detection of the IFN $\lambda$ 3 rs12979860 SNP. This kit, developed by TIB MolBiol (LightMix<sup>®</sup> Kit IFN $\lambda$ 3 rs12979860 kit; Roche, Cat. No. 40-0588-32) utilizes melt-curve analysis to determine the genotype. Commercial kits offer further reassurance that results are reliable, as well as offering an alternative to allow troubleshooting when problems do arise. Due to the extensive validation procedures and costs involved with providing ongoing



quality assurance, commercial kits are more expensive but usually provide greater reproducibility than in-house assays.

The commercially available TIB MolBiol kit uses melting curve analysis, an alternate technology to the validated TaqMan protocol, which is based on probe detection of amplification outcomes. Melting curve analysis is a technique that has been used for a number of years. First described in 1997 [136, 137], it allows for the detection and differentiation of products, primer dimers and non-specific binding according to the temperature at which double-stranded DNA dissociates (melting temperature or  $T_m$ ). The difference in  $T_m$  observed in different sequences is a result of the length of sequence and the base composition [136]; a C-G pair has three hydrogen bonds, while a T-A pair has two. This makes the C-G bond more stable, thus a higher C-G content means a higher temperature is required to separate the strands. A major advantage of melting curve analysis over conventional PCR and gel electrophoresis, is the ability to distinguish between products of the same size, but different sequences [136]. The starting temperature begins below the lowest expected  $T_m$  for a particular marker, and slowly ramps up to 90°C.

This study not only aimed to examine means of reducing costs by altering assay parameters, but also to assess the efficiency of the newly available commercial assay in relation to the methods already in place.

## **3.2: Methods**

### **3.2.1: DNA Extraction**

DNA extraction was performed using the Qiagen QIAamp DNA minikit as per manufacturer's instructions. Initial plasma input was 200 $\mu$ l, with a final elution volume of 100 $\mu$ l.

### **3.2.2: IFN $\lambda$ 3 Genotyping: ABI Allelic Discrimination Protocol**

IFN $\lambda$ 3 genotyping for both the reduced volume study and SensiFAST genotyping study was performed as described in chapter 2.2.3.

### **3.2.3: Reduced Volume PCR Reaction using In-House Assay**

Using the same DNA extract from the initial testing, twelve samples which had been previously tested were included as part of this study.

Two reduced PCR reaction volumes were evaluated:

a, 10 $\mu$ l TaqMan master mix, 1 $\mu$ l IFN $\lambda$ 3 primer/probe mix, and 9 $\mu$ l DNA template (total reaction volume of 20 $\mu$ l).

b, 11.9 $\mu$ l TaqMan master mix, 1.1 $\mu$ l IFN $\lambda$ 3 primer/probe mix, and 11 $\mu$ l DNA template (total reaction volume of 24 $\mu$ l).

Programming of the PCR on the ABI was altered to reflect the sample input volume.

### 3.2.4: Master mix Reagent: SensiFAST Genotyping Kit

Three previously tested samples of known IFN $\lambda$ 3 genotypes (one of each CC, CT and TT) were selected for this evaluation. SensiFAST Genotyping Kit (Bioline, Cat # BIO-36002—now discontinued) was used in a direct comparison to the TaqMan master mix.

Instead of the manufacturer's recommendation, alternate reagent volumes were employed; all calculated to stay within the suggested 20 $\mu$ l reaction volume (Table 3.1). As a direct comparison, cycling conditions on the ABI were as validated on the TaqMan master mix protocol.

	MM1	MM2	MM3
Master mix reagent	10 $\mu$ l	10 $\mu$ l	10 $\mu$ l
Probe	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l
Low ROX*	0.4 $\mu$ l	0.4 $\mu$ l	0.4 $\mu$ l
DNA	4.6 $\mu$ l	6 $\mu$ l	8.6 $\mu$ l
Nuclease Free Water	4 $\mu$ l	2.6 $\mu$ l	-
Total:	20 $\mu$ l	20 $\mu$ l	20 $\mu$ l

**Table 3.1: Master Mix volumes for SensiFAST Trial**

**\*Low ROX is a reaction-independent fluorophor (5-carboxy-X-rhodamine) which is recommended to use in the ABI 7500 instrument when running this assay.**

As the manufacturer's directions differed slightly, the suggested PCR parameters were followed (1 cycle of 3mins at 95°C, 40 cycles of 95° for 10 seconds and 60°C for 40 seconds), using the pre-designed pre- and post-read steps of 60°C for 1 minute as described in the ABI allelic discrimination guide.

The ct values were compared across different PCR conditions, master mix composition, and between TaqMan and SensiFAST reagents.

### 3.2.5: TIB MolBiol LightMix<sup>®</sup> Kit IFN $\lambda$ 3 rs12979860 Assay

Thirty-seven samples were tested (35 plasma and 2 buffy coat) using the TIB MolBiol LightMix<sup>®</sup> Kit IFN $\lambda$ 3 rs12979860 (Roche).

The 15 $\mu$ l reaction mix was as follows: 9.4 $\mu$ l nuclease free water, 1.6 $\mu$ l Mg<sup>2+</sup> 25mM, 2.0 $\mu$ l IFN $\lambda$ 3 reagent mix, 2.0 $\mu$ l Roche Master Mix (Cat.-No. 03 003 248 001). This was aliquoted into glass LightCycler capillaries, and 5 $\mu$ l sample or control DNA was added. PCR parameters were as displayed in Table 3.2.

	Denaturation	Cycling			Melting			Cooling
Cycles	1	45			1			1
Temperature (°C)	95	95	60	72	95	40	85	40
Hold mm.ss)	10.00	00.05	00.10	00.15	00.20	00.20	00.00	00.30
Ramp Rate (°C/s)	20	20	20	20	20	20	0.02	20
Acquisition Mode	n/a		Single				Continuous	n/a

**Table 3.2: PCR parameters for melt curve analysis using the TIB MolBiol LightMix<sup>®</sup> Kit IFN $\lambda$ 3 rs12979860 kit**

The assay was carried out using the Roche LightCycler (Version 1), with results analysed on version 4.1 of the software.

Results are assigned automatically via the software; one peak at 55°C represents homozygote TT; one peak at 59°C is homozygous CC; two peaks represent IFN $\lambda$ 3 CT, one peak at 51°C and one at 59°C. The results were compared to the results obtained using the ABI 7500 allelic discrimination protocol.

### **3.3: Results**

#### **3.3.1: Reduced volumes**

In order to utilize a step-pipette, which only allows aliquots of integers, the protocol for validated in-house was altered. A step-pipette allows the user to draw up a large volume of reagent, and dispense smaller aliquots into multiple tubes/wells. As multiple dispensing can be done without refilling, it reduces repetitive strain injury.

Twelve samples which had been successfully genotyped using the original protocol (25 $\mu$ l total reaction) were tested using a reduced 20 $\mu$ l volume, a reduction of 20%. In both the amplification and post-read analyses, three of the twelve did not produce a result, giving only a 75% efficiency rate. The remaining nine samples that were successfully genotyped were 100% concordant with the original results. When analyzing the amplification step and ct values between the two reaction volumes, the fluorescence was observed later in the reduced volume samples compared to the original volume, with an average ct value of 35.88 and 33.29 respectively (Table 3.3).

Subsequently, a second volume reduction was performed as an attempt to improve the results, while still allowing the use of a step-pipette. The reduction was 1 $\mu$ l in total (11.9 $\mu$ l

TaqMan master mix, 1.1 $\mu$ l IFN $\lambda$ 3 probe [13  $\mu$ l reaction mix], and 11 $\mu$ l DNA template; total volume 24 $\mu$ l). With this volume, genotyping of each sample was successful, and in 100% agreement with the original results. The ct was comparable between the original and reduced volumes, with an average of 30.99 and 30.89 respectively (Table 3.4). These results have allowed the use of the step-pipette, and therefore reduce multiple manual pipetting.

		REDUCED VOLUME (20 $\mu$ l)	ORIGINAL VOLUME (25 $\mu$ l)
Well	Detector	ct	ct
A1	VIC (C)	37.544	33.5629
	FAM (T)	36.2545	31.8085
A2	VIC (C)	36.2387	34.6162
	FAM (T)	Undetermined	Undetermined
A3	VIC (C)	Undetermined	33.5803
	FAM (T)	Undetermined	32.2124
A4	VIC (C)	35.1264	31.6678
	FAM (T)	Undetermined	Undetermined
A5	VIC (C)	32.6467	35.4481
	FAM (T)	31.1308	34.1234
A6	VIC (C)	Undetermined	Undetermined
	FAM (T)	36.9549	30.2068
A7	VIC (C)	Undetermined	36.7243
	FAM (T)	Undetermined	Undetermined
A8	VIC (C)	35.6282	34.5632
	FAM (T)	33.7007	33.5829
A9	VIC (C)	38.4529	30.9372
	FAM (T)	36.9732	29.5173
A10	VIC (C)	Undetermined	34.9123
	FAM (T)	Undetermined	Undetermined
A11	VIC (C)	38.4509	35.148
	FAM (T)	Undetermined	Undetermined
A12	VIC (C)	Undetermined	Undetermined
	FAM (T)	37.4029	33.3275

**Table 3.3: Ct values of initial reduced PCR reactions compared to original results. Reduced volumes are 10 $\mu$ l TaqMan master mix, 1 $\mu$ l IFN $\lambda$ 3 probe, and 9 $\mu$ l template (final volume of 20 $\mu$ l), a reduction of 20% from the validated protocol. Red text indicates variation from the original results.**

		24 $\mu$ L RXN	25 $\mu$ L RXN
Well	Detector	Ct	Ct
A1	VIC (C ALLELE)	29.4293	29.3208
	FAM (T ALLELE)	Undetermined	Undetermined
A2	VIC (C ALLELE)	34.4562	35.226
	FAM (T ALLELE)	32.7318	33.4771
A3	VIC (C ALLELE)	31.8398	32.0709
	FAM (T ALLELE)	30.6851	30.5072
A4	VIC (C ALLELE)	34.4733	34.9896
	FAM (T ALLELE)	33.7264	33.7236
A5	VIC (C ALLELE)	33.5866	33.3933
	FAM (T ALLELE)	Undetermined	Undetermined
A6	VIC (C ALLELE)	28.007	27.94
	FAM (T ALLELE)	26.4509	26.3361
A7	VIC (C ALLELE)	30.6292	30.4799
	FAM (T ALLELE)	29.228	29.0949
A8	VIC (C ALLELE)	29.1183	29.3498
	FAM (T ALLELE)	Undetermined	Undetermined
A9	VIC (C ALLELE)	Undetermined	Undetermined
	FAM (T ALLELE)	32.6047	32.7313
A10	VIC (C ALLELE)	29.6319	29.7567
	FAM (T ALLELE)	28.2665	28.3081
A11	VIC (C ALLELE)	31.3169	31.3818
	FAM (T ALLELE)	Undetermined	Undetermined
A12	VIC (C ALLELE)	30.015	29.8928
	FAM (T ALLELE)	Undetermined	Undetermined

**Table 3.4: Ct values of reduced PCR reactions (24 $\mu$ L RXN) compared to original results (25 $\mu$ L RXN). Reduced volumes are 11.9 $\mu$ L TaqMan master mix, 1.1 $\mu$ L IFN $\lambda$ 3 probe, and 11 $\mu$ L template.**

### 3.3.2: SensiFAST Genotyping Kit

With the view to substantially reduce the cost of the assay, the less expensive SensiFAST Genotyping Kit was assessed. Three samples of known genotypes (CC, CT and TT for IFN $\lambda$ 3

rs12979860) were tested using the SensiFAST genotyping master mix, using three different volumes of DNA and nuclease free water, adjusted to a total reaction volume of 20 $\mu$ l (Table 3.1). There was very little observed difference in cycle threshold between each of the three master mix preparations evaluated (Table 3.5).

When scrutinizing the amplification plots and data output, the results were not in agreement with the genotypes initially obtained in the in-house assay (sample names assigned by their initial result). In the CC sample, there was amplification in both the VIC (C allele), and the FAM (T allele) channels, which is indicative of a heterozygous genotype. Only amplification in the VIC channel is expected in a CC homozygote specimen. This was consistent across all three master mix compositions. The known CT and TT samples returned the expected results; the respective probes (alleles) were amplified as expected (Table 3.5).

Following the amplification step of the assay, the post-read component was carried out. The results from the allelic discrimination plot (Figure 3.1) were consistent with that obtained using the TaqMan master mix, but not with results obtained from the amplification data. The samples with the CC genotype were not aligned as tightly along the X-axis as was shown using the TaqMan mastermix, although the “C allele” was still determined by the software, and were appropriately represented by the red circle (Figure 3.1 A).

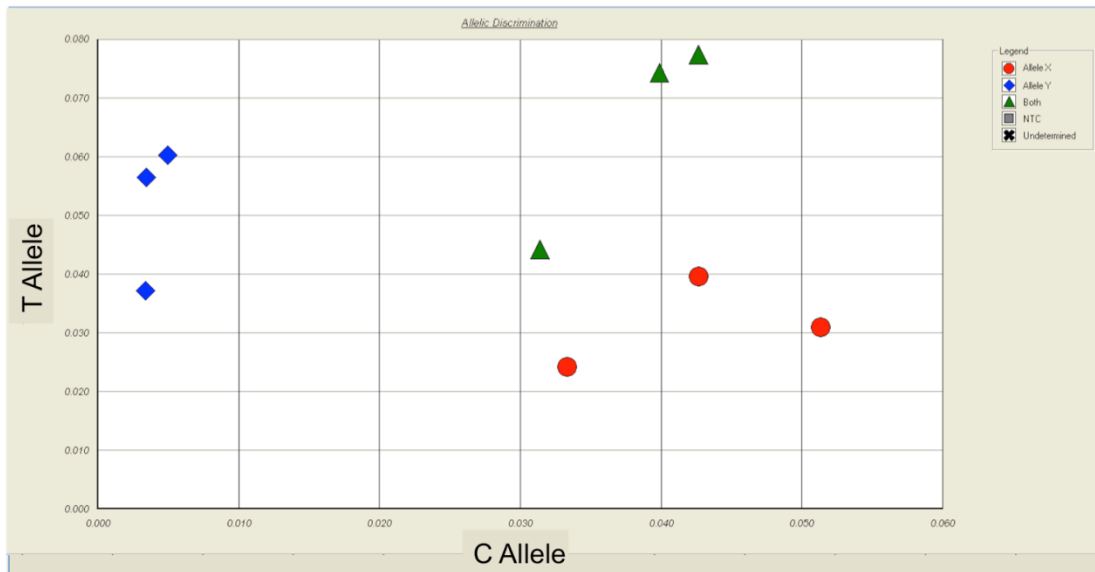


			SENSIFAST MASTER MIX		
			MM 1 (ct)	MM 2 (ct)	MM 3 (ct)
SAMPLE	CC	VIC (C)	35.75	35.19	31.19
		FAM (T)	36.63	35.56	35.67
	CT	VIC (C)	34.53	34.38	34.01
		FAM (T)	33.81	28.56	27.83
	TT	VIC (C)	UNDETERMINED	UNDETERMINED	UNDETERMINED
		FAM (T)	34.71	33.41	32.82

**Table 3.5: The ct values obtained from the amplification plot:**

The sample name denotes the results previously obtained using the TaqMan master mix; one sample from each possible genotype, using three different master mix compositions (MM1, MM2, MM3). The red text indicates the anomalies found in the results using the SensiFAST Genotyping Kit.

To determine if the amplification of the T allele in the CC samples was unique to the one sample, an additional four known CC samples were tested with the SensiFAST master mix. In each of these four samples, there was amplification of both the C and the T alleles as there was in the initial CC test specimen (Table 3.6). Unlike in the previous comparison, the post-read was not able to give a definitive genotype, and as a result each was described as ‘undetermined’ (Figure 3.2). This illustrated the product was inadequate for the purpose of the assay, and was not introduced as a replacement reagent. Sale of this product has since been discontinued.

**A****B**

Well	Sample Name	Genotype Call	Quality (%)
A1	CC (MM1)	C ALLELE (CC)	99.52
A2	CT (MM1)	Both (CT)	98.64
A3	TT (MM1)	T ALLELE (TT)	99.37
B1	CC (MM2)	C ALLELE (CC)	99.62
B2	CT (MM2)	Both (CT)	99.99
B3	TT (MM2)	T ALLELE (TT)	99.8
C1	CC (MM3)	C ALLELE (CC)	99.76
C2	CT (MM3)	Both (CT)	99.98
C3	TT (MM3)	T ALLELE (TT)	99.97

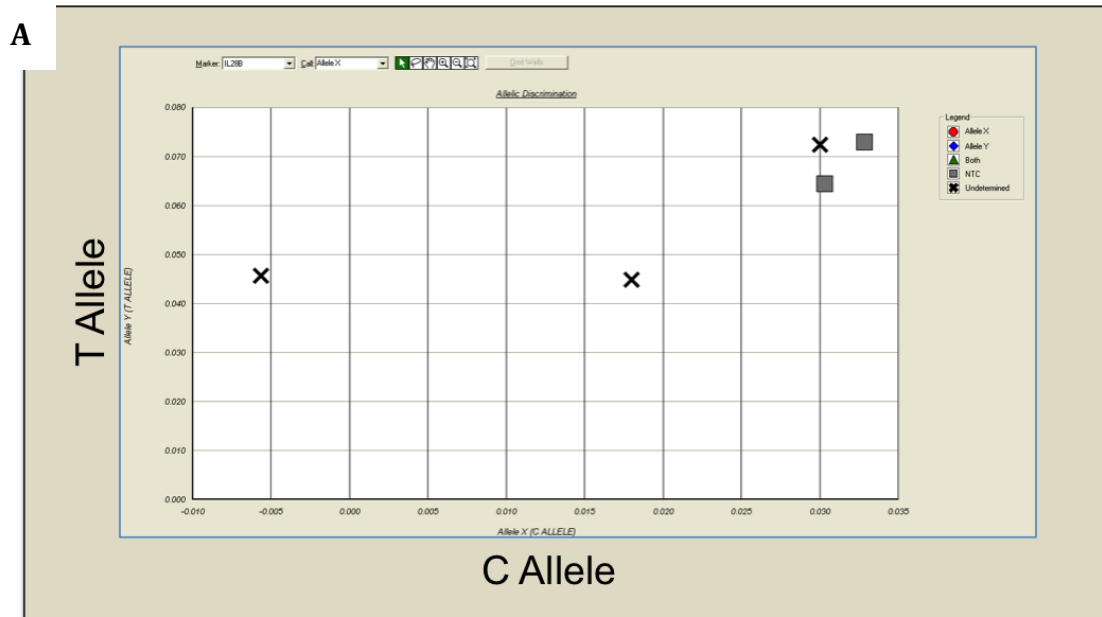
**Figure 3.1: Results of post-read analysis following use of SensiFAST master mix.**

**A.** Allelic discrimination plot; blue diamond = TT, green triangle = CT, red circle = CC.

**B.** Results output following post-read analysis; all results fall well within the quality threshold range.

Well	Sample Name	Detector	ct
A1	CC 1	VIC (C)	32.1666
A1	CC 1	FAM (T)	33.868
A2	CC 2	VIC (C)	27.0574
A2	CC 2	FAM (T)	28.4395
A3	CC 3	VIC (C)	32.6265
A3	CC 3	FAM (T)	33.1824
A4	CC 4	VIC (C)	33.6622
A4	CC 4	FAM (T)	35.205
A5	ntc	VIC (C)	Undetermined
A5	ntc	FAM (T)	Undetermined
A6	ntc	VIC (C)	Undetermined
A6	ntc	FAM (T)	Undetermined

**Table 3.6: A further four IFN $\lambda$ 3 genotype CC specimens were assayed using the SensiFAST master mix; Red text indicates the channel in which amplification should not have occurred. Two 'No Template Controls' (NTC) were included; expected result of 'undetermined' was achieved in both channels of each sample**



**B**

Well	Sample Name	Result	Quality (%)
A1	CC 1	Undetermined	0
A2	CC 2	Undetermined	0
A3	CC 3	Undetermined	0
A4	CC 4	Undetermined	0

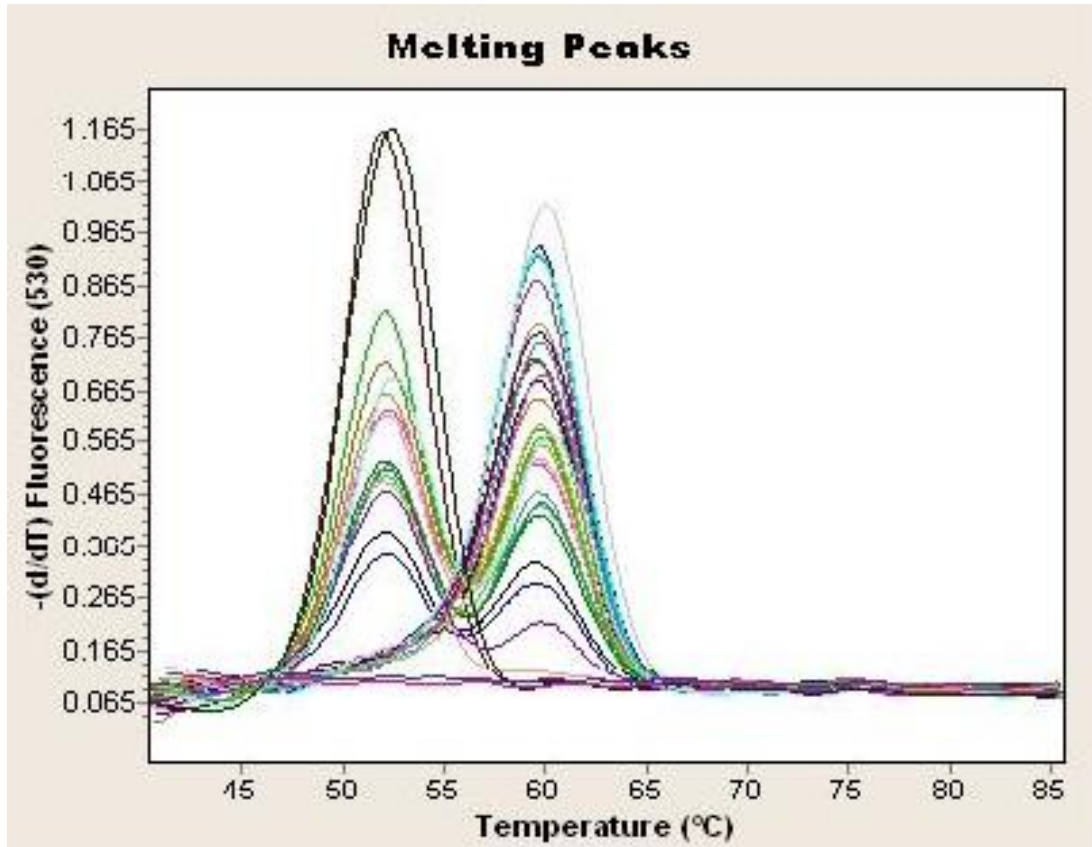
**Figure 3.2: Results of post-read analysis following use of SensiFAST master mix on known IFN $\lambda$ 3 CC samples, shown in Table 3.6. (A) Allelic discrimination plot; 'X' is representative of samples for which genotypes cannot be determined; boxes represent NTC wells. (B) Results output following post-read analysis.**

### 3.3.3: TIB MolBiol LightMix Kit

A total of 37 samples were compared using the TaqMan allelic discrimination assay and the TIB MolBiol LightMix kit. Genotypes obtained from the Roche LightCycler using the LightMix assay acquire readings using a melt curve analysis; the C allele has a melting point of 59°C,

and the T allele melts at 51°C. If the melting curve shows one peak, it is homozygote for that allele, and if it has two peaks, it is a heterozygote (Figure 3.3).

The majority of samples (95%) gave concordant results for the TaqMan and TIB MolBiol assays. Two samples were undetermined in the ABI but were assigned a genotype in the LightCycler. Five were initially called as 'unknown' in the LightCycler but once the quality threshold was reduced (down to a minimum of 90%), three of these five correlate with the ABI result (Table 3.7). For the remaining two samples, the buffy coats were extracted and run in the LightCycler; one (sample 4) matched the LightCycler plasma result, and one (sample 5) matched the ABI plasma result (summarized in Table 3.7).



**Figure 3.3: Melt-curve analysis from LightCycler software using TIB MolBiol IFN $\lambda$ 3 rs12979860 kit. A solitary peak at 51°C (A) specifies a 'T' allele; a solitary peak at 59°C (B) specifies a 'C' allele. If a peak is present at both 51°C and 59°C, the sample is IFN $\lambda$ 3 genotype CT.**

	<b>TaqMan</b>	<b>LightMix</b>
1	CT	CT*
2	CT	CT
3	CC	CC
4	TT	CT*
5	CT	CT
6	CC	CC
7	CC	CC
8	CC	CC
9	CT	TT*
10	CC	CC
11	CC	CC
12	CT	CT
13	CC	CC
14	CT	CT
15	CT	CT
16	UNDETERMINED	CT*
17	UNDETERMINED	CC*
18	TT	TT
19	CC	CC
20	CT	CT
21	CC	CC
22	CC	CC
23	CT	CT
24	CC	CC
25 <sup>#</sup>	CT	CT*
26 <sup>#</sup>	CT	CT
27	CC	CC*
28	CT	CT
29	CT	CT
30	CC	CC
31	CC	CC
32	CC	CC
33	CT	CT
34	TT	TT
35	CT	CT
36	TT	TT
37	TT	TT

**Table 3.7: Comparison of 37 samples tested using both TaqMan allelic discrimination and LightMix protocols.**

Results highlighted in yellow represent discrepant results between the two assays; DNA extracted from buffy coat was required to determine true genotype status.

Results with (\*) indicate LightCycler quality call needed to be reduced to 90% to obtain a result.

Samples 25 and 26 (#) were paired plasma and buffy coat specimens respectively.

### 3.4: Discussion

#### 3.4.1: Reduced Volumes

It was hypothesized that reducing the volume of the PCR would not detract from the efficiency in determining IFN $\lambda$ 3 genotype, while achieving a cost saving. Furthermore, changing PCR reaction volumes can improve the health and reduce injury to laboratory workers by allowing the use of a step-pipette. Some studies have shown that by reducing the volume of the PCR mix, while maintaining the template volume, the efficiency is increased - presumably due to the DNA template having more access to the primers in the mix [138]. Various studies have shown that PCR protocols are able to be scaled back by up to 50% without losing any efficiency [139-142]. In this study however, the initial reduction of 20% was unsuccessful, yielding only 75% genotyping efficiency. Of the samples that were successfully genotyped, the amplification was less efficient in the trial protocol, as reflected in the higher ct values. With such a large number of samples not being able to be accurately genotyped, it is presumed that the manufacturer has performed extensive optimization studies. Due to these comparisons, the failure rate in this trial was much higher than anticipated.

A common complaint of scientific staff is repetitive strain injury due to frequent pipetting [143, 144], with an increase in hand complaints compared to the general population (44% compared to 24% respectively) [144]. The volume of master mix used in the IFN $\lambda$ 3 genotyping assay, 13.75 $\mu$ l, does not allow for the use of a standard step-pipette which is limited to increments of 1 $\mu$ l. Hence, to avoid injury and strain, a lesser reduction in volume was evaluated. If unsuccessful, an increase could be assessed to bring the final master mix



volume up to 14 $\mu$ l, which would also require an increase in DNA template. While this would fractionally increase the cost of the assay, it would not be significant, but would allow shorter working times due to the use of optimal laboratory tools.

The overall reduction in reaction mix was scaled back to a minor reduction of just 1 $\mu$ l. By reducing the mix from 25 $\mu$ l to 24 $\mu$ l, the efficiency did not appear to be compromised. This reduction was evaluated using twelve samples, with results showing 100% agreement with the original results using the larger volume. In addition to this, the difference in ct between the two was a mere 0.1 cycles and likely not statistically significant.

While this represents a miniscule cost saving in reagents, it does improve hands-on time in setting up the PCR plate. As the master mix and probe component is reduced to a whole number (13 $\mu$ l), it allows for the use of a step-pipette, saving not only time, but also reducing physical stress on the scientist's hands. As this reduction was successful, it remains to be the protocol used for the determination of IFN $\lambda$ 3 genotype.

#### **3.4.2: SensiFAST Genotyping Kit**

With the aim of reducing costs, the SensiFAST Genotyping master mix was evaluated against the TaqMan master mix, with a comparative cost saving of approximately 50%. The results seen with the SensiFAST master mix were unexpected, as they were not consistent with the TaqMan product. Using the SensiFAST reagent, there was amplification in both the VIC (C allele) and FAM (T allele) channels in samples previously tested with the TaqMan master mix and genotyped as IFN $\lambda$ 3 CC (Figure 3.1). This result indicates the patient had IFN $\lambda$ 3 CT

genotype due to both probes fluorescing. However, once the post-read had been performed, the allelic discrimination software correctly assigned a 'CC' genotype to these samples. The other two samples in the initial test (known CT and TT specimens) gave expected results in both the amplification and the allelic discrimination plots. These results were seen across all three master mix compositions tested.

To determine if this result was sample-specific, four additional IFN $\lambda$ 3 CC specimens were selected and re-tested using the SensiFAST master mix. As the same results were identified on the amplification plots on each of these samples (Table 3.6), it was confirmed that it was unrelated to the sample SNP, and was in fact due to the alternate master mix product itself, and not a specific sample or master mix volume composition. The fact that the representative points on the allelic discrimination plot were not as tightly aligned along the X-axis (Figure 3.1 A) as with the TaqMan master mix may be a reflection of the results seen in the amplification plot, whereby both the C and T allele were detected (Table 3.6).

By definition, there is only a single nucleotide difference between each genotype; this allows the probes to bind with lower efficiency to the alternative allele. Due to the nature of the allelic discrimination assay, the amplification plots are routinely only used for troubleshooting. However, as there was no distinction between each allele, it would not have been possible to use the data to troubleshoot. If the results were validated using the amplification data alone, they would be incorrectly released as CT heterozygote, and have the potential of altering treatment regimens for the patient.

In addition to the potentially incorrect results in a diagnostic setting, if the SensiFAST master mix was utilized in further predictive studies, the results may not be accurate. There is the potential that genetic associations to certain phenotypes are not observed, simply from using a product not optimized for the assay.

Subsequent to this comparative analysis, this particular SensiFAST genotyping master mix has been discontinued, and alternative reagents from the same manufacturer (also named SensiFAST, but with different catalogue numbers) have superseded this product. This study did not seek to include these in the analysis.

#### **3.4.3: TIB MolBiol LightMix Kit**

The use of the TIB MOLBIOL LightMix kit utilizes melting curve analysis technology; the  $T_m$  for the C allele is 59°C, and the T allele is 51°C which is easily distinguishable in the data output. The study yielded good results, with a concordance rate of 95% with the TaqMan allelic discrimination assay. The software automatically assigns genotype assuming a certainty of 95%. If the samples fall outside this range, the certainty threshold may be reduced to 90%. If this adjustment is performed, the user must carefully analyse each sample within the altered range to ensure accuracy.

Many studies have taken advantage of this reliable assay to perform their IFN $\lambda$ 3 genotyping [111, 145-147]. It is a user-friendly assay, which yields accurate results in a short amount of time. The disadvantage of using this kit is the use of specialized equipment; the Roche LightCycler, and its accompanying centrifuge and dedicated consumables. Whilst

laboratories are able to use the instrument for a variety of tests, it is not as versatile as a standard real-time machine which uses a 96-well plate format such as the ABI 7500. The Roche LightCycler uses glass capillaries that must be handled multiple times during both the set-up and in the post-run clean up, with the potential to break at various times. It also has a total capacity of 32 samples; once controls (three IFN $\lambda$ 3 controls, and one negative control) are included, there is the capacity for running 28 clinical or study samples. This is compared to a 96-well plate format, which has the potential to run 91 clinical samples, with the addition of the recommended controls (three IFN $\lambda$ 3 controls, and two negative controls).

The cost of running a commercial assay rather than an in-house assay must also be considered. Based on the cost of the assay, from DNA extraction through to genotype analysis, the cost of the TaqMan allelic discrimination assay on the ABI is 60% of the cost of the LightMix kit. If doing a direct comparison of the LightMix kit vs TaqMan probe and TaqMan master mix, the TaqMan (ABI method) is most affordable at less than 20% of the cost.

The advantages of the commercial assay mean there is the extra validation and quality assurance performed by the manufacturer. However, a laboratory can readily perform its own validation, and our laboratory has decided to continue with the TaqMan allelic discrimination assay and use the LightMix kit as a reference point [145].

### 3.5: Conclusion

Decreasing the volume of the PCR by just 1 $\mu$ l has proven to be efficient – not significant for cost saving, but in labour time and effort. By being able to utilize as step-pipette, which does not aliquot in fractions of microliters, the repetitive nature of aliquoting master mix into a 96-well plate has been significantly reduced.

Due to the non-specific amplification when using the SensiFAST master mix, this product will not be considered for use within the laboratory. The lack of troubleshooting ability whilst using this product renders it unsuitable for this application. The TaqMan master mix continues to be the reagent of choice, as it can be used to reliably determine the IFN $\lambda$ 3 genotype of clinical samples.

The TIB MOLBIOL LightMix serves as a suitable alternative to an in-house assay, and will be retained as a backup and may be used for confirmation of results if necessary. Though it is a reliable assay, the higher cost involved in running the test can be easily avoided by continuing with the TaqMan allelic discrimination protocol.

## **Chapter 4: Association of IFN $\lambda$ 3 SNPs with Different HCV Genotype 1 Subtypes**

### **4.1: Introduction**

In Australia, the most prevalent HCV genotypes are HCV-1 (including subtypes 1a and 1b) (52%), and HCV-3 (31%), followed by HCV-2 and HCV-4 [148, 149]. Patients infected with HCV genotype 2 or 3 have a greater chance of achieving a sustained viral response (SVR) whilst on pegIFN/RBV therapy (80%) than those infected with genotype 1 or 4 (40-45%) [150]. The duration of therapy varies depending on the HCV genotype, with patients infected with HCV genotypes 2 and 3 having treatment for 24 weeks and non-genotype 2 and 3 (predominantly genotype 1) having to undergo 48 weeks of treatment. Other factors positively influencing viral clearance include younger age [151], female gender [152], ethnicity [69, 153, 154], lower body mass index [155], and IFN $\lambda$ 3 genotype [41, 43, 44, 53].

When the association between IFN $\lambda$ 3 genotype and response to pegIFN/RBV therapy was identified, specific genomic sequences for alleles (CC at rs12979860, or TT at rs8099917) were identified as being advantageous for patients with HCV genotype 1 to achieve better treatment outcome [41, 43, 44]. It was subsequently shown that there is also a positive association between favourable IFN $\lambda$ 3 alleles and an SVR in patients infected with HCV genotype 2 and 3 but who do not achieve a rapid virological response (HCV RNA negativity at week 4 of treatment) [156].

There is scant data on whether there is any difference in response rates for patients infected with HCV genotype 1 subtypes 1a and 1b, respectively, whilst on pegIFN/RBV therapy, primarily because treatment remains the same regardless of subtype. In addition, different effects due to IFN $\lambda$ 3 genotype between HCV subtypes 1a and 1b in HCV mono-infected patients has not previously been investigated. There have been studies in the setting of HIV/HCV co-infection [157]; however this data is also limited. It is clear that there is gap in the knowledge surrounding the interplay between HCV genotype 1 subtypes and IFN $\lambda$ 3 genotypes, in spite of the fact that the initial association was recognized in patients infected with HCV genotype 1.

As a state reference laboratory providing diagnostic testing to a vast number of pathology services across not only the Victoria, but also nationally, VIDRL has an extensive database. The diagnostic services in regard to HCV include antibody testing, qualitative PCR, quantitative PCR (viral load), HCV genotyping, HCV core sequencing, and IFN $\lambda$ 3 genotyping. With all these tests combined, along with clinical notes provided from the healthcare professionals, the database contains a significant amount of information. This potentially allows for the information to be accessed and analysed in a number of ways.

By assessing this data, there is the potential to further characterize the IFN $\lambda$ 3 distribution according to HCV genotype 1 subtypes, as well as extrapolate any further differences in treatment response between subtypes.

## **4.2: Methods**

### **4.2.1: Patients**

This was a retrospective analysis of 4,446 individuals tested for IFN $\lambda$ 3 genotype up until the end of December 2012. The database was scrutinized to identify those patients for whom additional information such as HCV genotype and HCV RNA status was available whilst on therapy. Patients with at least four blood draws, including baseline, and 6 months post-therapy, were included in the analysis.

### **4.2.2: IFN $\lambda$ 3 Genotyping**

IFN $\lambda$ 3 genotyping (rs12979860) was performed as described in Chapter 2, using the Qiagen QIAamp DNA Minikit to extract genomic DNA from plasma samples.

### **4.2.3: HCV Genotyping**

HCV genotyping was performed using Versant Hepatitis C Virus Genotype Assay (Line Probe Assay - LiPA) version 2.0. Regions from both the 5' untranslated region (5'UTR) and the core region were amplified, prior to a reverse-phase hybridization assay. Genotypes are diagnosed via a banding pattern and cross-referenced with an interpretation chart.

### **4.2.4: HCV Detection**

Detection of HCV RNA is routinely performed by reverse transcription-PCR in either qualitative or a quantitative assays; for this analysis data was included from both.



#### **4.2.4.1: Qualitative Assay**

HCV RNA status was performed using the Roche COBAS Ampliprep/COBAS Amplicor HCV test, version 2.0, with a lower limit of detection of 50 IU/ml in plasma or serum. The assay has three major steps; RNA extraction, reverse transcription to generate cDNA and PCR amplification with simultaneous detection using TaqMan probes.

#### **4.2.4.2: Quantitative Assays**

Prior to December 2011, all HCV quantitative tests were performed using Siemens Versant HCV RNA 3.0 (bDNA) assay; the dynamic range of 615 to 7,690,000 IU/mL. From December 2011, HCV RNA levels were detected using the Abbott RealTime HCV test with the m2000sp and m2000rt instruments; the test has a dynamic range of 30 to 100,000,000 IU/mL when 0.2mL sample volumes are used. This quantitative assay follows a similar principle as the qualitative assay but has the addition of calibration standards to generate a calibration curve which is used for interpolation of values.

#### **4.2.5: Statistical Analysis**

Statistical analysis performed in this chapter was by student t-test, using graph-pad software. Significance was claimed when values were <0.05.

### **4.3: Results**

A total of 4446 patient samples had been received for IFN $\lambda$ 3 genotype testing to the end of December 2012. Of these, 1,699 also had HCV genotype testing performed at VIDRL. The breakdown of HCV genotype is displayed in table 4.1. The majority (69%) of patients were infected with HCV genotype 1, with the next most prevalent being genotype 3 (24%). HCV genotypes 2 (2%), 4 (2%), and 6 (3%) were in the minority. Overall, IFN $\lambda$ 3 distribution was:

CC (40%), CT (47%), and TT (12%). Two groups had higher percentages of IFN $\lambda$ 3 CC patients compared to CT; HCV genotype 2 (60%) and genotype 6 (79%).

HCV Genotype	IFN $\lambda$ 3 Genotype			
	CC	CT	TT	non-CC
	n (%)	n (%)	n (%)	n (%)
1 (n=1172)	433 (37%)	575 (49%)	164 (14%)	739 (63%)
2 (n=35)	21 (60%)	12 (34%)	2 (6%)	14 (40%)
3 (n=412)	190 (46%)	188 (46%)	34 (8%)	222 (54%)
4 (n=33)	7 (21%)	20 (61%)	6 (18%)	26 (79%)
6 (n=47)	37 (79%)	9 (19%)	1 (2%)	10 (21%)

**Table 4.1: IFN $\lambda$ 3 and HCV Genotype Distribution. A total of 1,699 patients had been tested for both HCV and IFN $\lambda$ 3 genotyping.**

A total of 1,172 patient samples were HCV genotype 1; these were further sorted according to subtype. Seventy of these HCV-1 patients could not be subtyped, and an additional 30 were genotyped as 1a/b; these were excluded from the analysis. Of those for which subtypes were available, 788 were HCV 1a, and 284 were HCV 1b. Of those patient samples with HCV 1b, 57 were excluded from analysis as genotyping was performed using the first generation LiPA; these patient samples may have been mistyped due to deficiencies in the original test.

By selecting patients for whom there were results for defined time points (minimum of four bleeds, including baseline and 6 months post-therapy) the total patient count was reduced

from 1,015 down to 289. Of these, 204 samples (71%) were HCV 1a, and 85 (29%) were HCV 1b. The breakdown of IFN $\lambda$ 3 genotype according to HCV genotype is outlined in table 4.2. Patient data was further analysed to determine frequencies of SVR according to IFN $\lambda$ 3 genotype, within the HCV genotype 1 subtypes (Table 4.3). The IFN $\lambda$ 3 distribution is marginally different in the HCV genotype 1 patients compared to the overall population; TT is the same (12%), but the favourable CC genotype is lower in the HCV-1 patients than the overall data set (38% compared to 40%). The heterozygote IFN $\lambda$ 3 CT is similar in both HCV 1 subtypes analysed, and a non-significant trend shows more IFN $\lambda$ 3 TT patients are infected with HCV 1a than HCV 1b.

		IFN $\lambda$ 3 Genotype			
		CC	CT	TT	non-CC
HCV Genotype	Gender	n (%)	n (%)	n (%)	n(%)
1a	Male (n=148)	55 (37%)	71 (48%)	22 (15%)	93 (63%)
	Female (n=56)	19 (34%)	32 (57%)	5 (9%)	37 (66%)
	Total (n=204)	74 (36%)	103 (50%)	27 (13%)	130 (64%)
1b	Male (n=45)	21 (47%)	20 (44%)	4 (9%)	24 (53%)
	Female (n=40)	14 (35%)	22 (55%)	4 (10%)	26 (65%)
	Total (n=85)	35 (41%)	42 (49%)	8 (9%)	50 (59%)
Total	Male (n=193)	76 (39%)	91 (47%)	26 (13%)	117 (61%)
	Female (n=96)	33 (34%)	54 (56%)	9 (9.5%)	63 (66%)
	Total (n=289)	109 (38%)	145 (50%)	35 (12%)	180 (62%)

**Table 4.2: Distribution of IFN $\lambda$ 3 genotypes within HCV 1 subtypes.**

Females were more likely to clear the virus than males (49% and 23% respectively,  $p=0.0002$ ). This difference was amplified in HCV 1b patients, with females exhibiting a clearance rate of 70%, compared to the rate in males of 24% ( $p<0.0001$ ). The only subset in which this did not hold true was in the HCV 1a, IFN $\lambda$ 3 CC patients (Table 4.3).

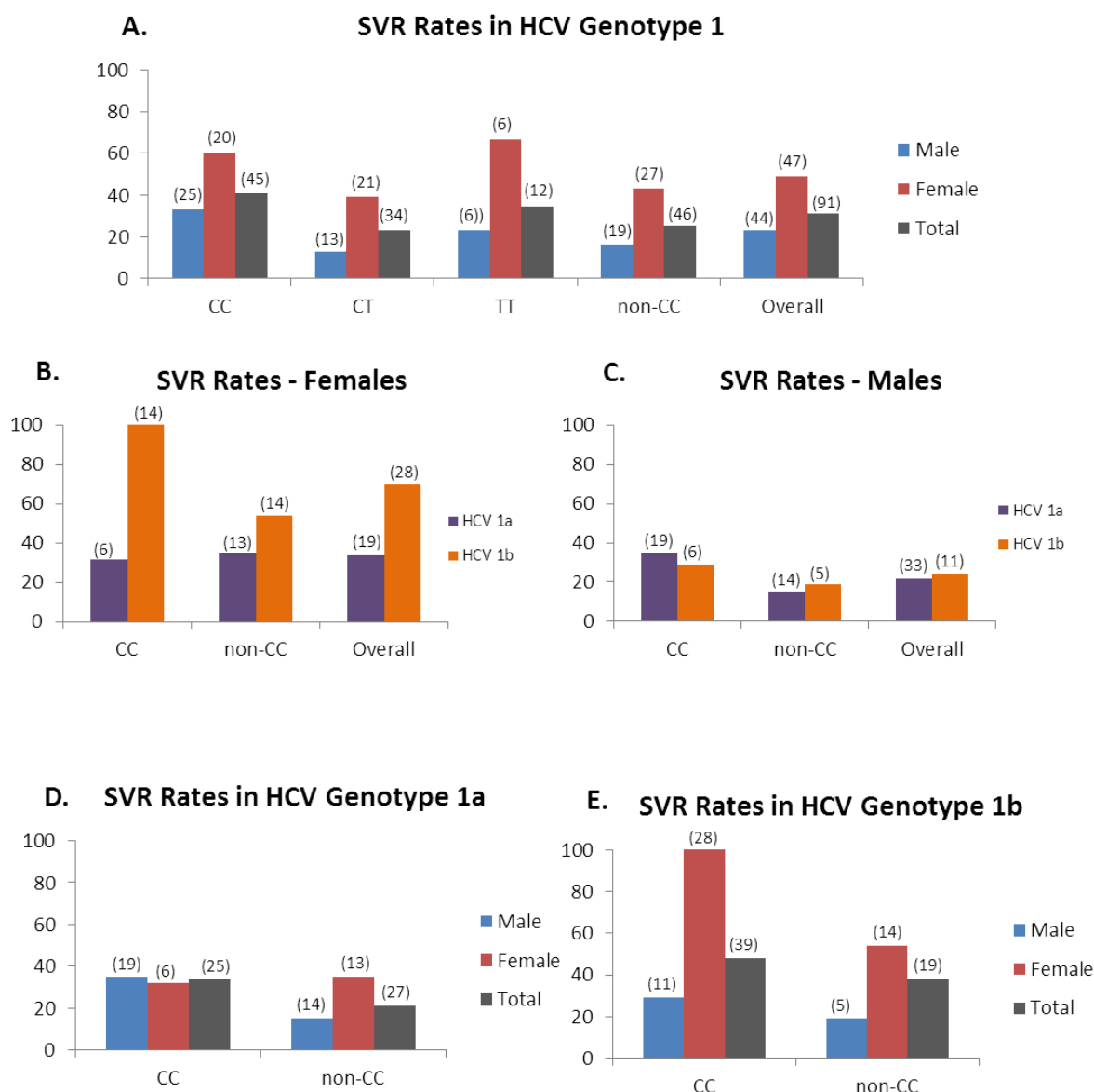
HCV Genotype	Gender	SVR (%)				
		IFN $\lambda$ 3 Genotype				
		Total	CC	CT	TT	non-CC
1a	Male (n=148)	33 (22%)	19 (35%)	9 (13%)	5 (23%)	14 (15%)
	Female (n=56)	19 (34%)	6 (32%)	9 (28%)	4 (80%)	13 (35%)
	Total (n=204)	52 (25%)	25(34%)	18 (17%)	9 (33%)	27 (21%)
1b	Male (n=45)	11 (24%)	6 (29%)	4 (20%)	1 (25%)	5 (19%)
	Female (n=40)	28 (70%)	14 (100%)	12 (55%)	2 (50%)	14 (54%)
	Total (n=85)	39 (46%)	20 (48%)	16 (38%)	3 (38%)	19 (38%)
1a and 1b	Male (n=193)	44 (23%)	25 (33%)	13 (14%)	6 (23%)	19 (16%)
	Female (n=96)	47 (49%)	20 (60%)	21 (39%)	6 (67%)	27 (43%)
	Total (n=289)	91 (31%)	45 (41%)	34 (23%)	12 (34%)	46 (26%)

**Table 4.3: Association of HCV Genotype and IFN $\lambda$ 3 genotype with Gender.**

When grouped by IFN $\lambda$ 3 genotype, the data shows patients with IFN $\lambda$ 3 TT respond better to pegIFN/RBV therapy than IFN $\lambda$ 3 CT (34% vs 23%) (Figure 4.1 A). However, the data for IFN $\lambda$ 3 TT patients is low in numbers, and may not be reliable. When sorting by IFN $\lambda$ 3 CC vs non-CC, the data shows that IFN $\lambda$ 3 CC patients respond better to treatment than non-CC patients, regardless of HCV 1 subtype (fig 4.1 A). A greater response was observed in IFN $\lambda$ 3 CC

females infected with HCV 1b than HCV 1a, though a larger sample size is required to deem any significance (Figure 4.1 B and 4.1 C).

HCV genotype 1b had higher overall SVR rates than HCV 1a (46% and 25% respectively;  $p=0.0030$ ) (Figure 4.1 D and E); this was consistent across all HCV/IFN $\lambda$ 3 genotype combinations for both males and females, with the exception of males carrying the IFN $\lambda$ 3 CC genotype in which HCV 1a patients had a similar SVR to HCV 1b (35% and 29%,  $p=0.4486$ ; not significant).



**Fig 4.1: Sustained virological response according to HCV-1 subtypes and IFN $\lambda$ 3 genotypes.** The effect of gender and IFN $\lambda$ 3 genotype in clearance of HCV genotype 1 (combined) is illustrated in (A). (B) and (C) illustrate the effect that HCV 1 subtypes and IFN $\lambda$ 3 genotypes have on females and males respectively. HCV genotype 1 subtypes are illustrated separately (HCV 1a (D) and HCV 1b (E)).

#### 4.4: Discussion

From January 2011 to December 2012, IFN $\lambda$ 3 genotype analysis had been requested for 4,446 individuals. HCV genotype data was available for 1,699 of these patients. HCV-1 was

the most common genotype (69%), followed by HCV-3 (24%), HCV-6 (3%), HCV-2 (2%), and HCV-4 (2%). These figures are in line with other studies conducted, where HCV-1 and 3 are the more predominant genotypes in Australia [16, 148, 149, 158].

Patients were sorted based on IFN $\lambda$ 3 genotype; CC vs non-CC. The distribution between each HCV genotype reflected the response rate seen in pegIFN/RBV treatment. The favourable IFN $\lambda$ 3 CC genotype is poorly represented in the more difficult to treat HCV genotypes 1 and 4, with 37% and 21%, respectively. In contrast, the IFN $\lambda$ 3 CC distribution in HCV genotypes 2, 3, and 6, are 60%, 46%, and 79% respectively. Although in this study, patients were not studied according to ethnicity, the distribution between IFN $\lambda$ 3 and HCV genotypes display an expected pattern that is in agreement with both HCV genotype and IFN $\lambda$ 3 genotype distributions globally. It has been recognized that people of Asian descent respond well to pegIFN/RBV therapy [153, 159, 160], in which there is a higher frequency of the favourable IFN $\lambda$ 3 CC [53, 161, 162]. In addition, the predominant genotypes in Asia are HCV 1 and 6 [160, 163, 164]; the effect of IFN $\lambda$ 3 on these genotypes is more pronounced than in other HCV genotypes. In contrast, African populations are difficult to treat [68-70], being predominantly IFN $\lambda$ 3 non-CC [41, 53, 165], as well as having difficult to treat HCV 1 or 4 [72, 148, 164]. If ethnicity had been included in this analysis, it is likely to have explained some of the SVR rates observed. IFN $\lambda$ 3 CC and CT are common throughout European and Caucasian populations [41, 53]. In these populations, HCV genotype 3 has taken over from genotype 1 as the most common variant for new incident infections; older patients whom acquired HCV via blood transfusions tend to be infected with HCV-1, whilst younger injecting drug users carry HCV-3 [166-169].

When selecting the HCV-1 patients and grouping according to subtype, 289 of the 1,172 were suitable for this analysis. Others were excluded for reasons such as lack of clinical notes, insufficient HCV RNA data, or not having subtype data. Patients genotyped as HCV 1b prior to the introduction of the second generation LiPA kit (LiPA 2) were excluded from analysis. The first generation LiPA based its genotyping on nucleotide differences in the conserved 5' untranslated region (5' UTR) which was unable to discriminate HCV genotypes 6c-I from 1b. The second generation LiPA assay is able to distinguish between HCV 1b and HCV 6c-I by utilizing the core region of the HCV genome. Many of those with a possible mistype in the first generation LiPA include South-East Asian HCV patients; even those who were infected with genuine HCV genotype 1b (as determined by sequence analysis) responded better to treatment than Caucasian patients infected with HCV 1b [159]. As we now know, at least part of this difference would be attributable to the higher frequency of the favourable CC IFN $\lambda$ 3 genotype in the South-East Asian population.

Once the patients were grouped into HCV-1 subtypes, it became apparent there were distribution differences between the two groups. Overall, there were more males in the study (67%), which is consistent with other studies [149, 170, 171]. The lower number of females in this study, and many others, may be due to lower rates of infection [172, 173], or higher rates of spontaneous viral clearance seen in women [88, 174-176]. Interestingly, there was a difference between gender distribution between HCV-1 subtypes; in HCV-1a, females constituted 27% of patients, compared to 47% in HCV-1b.

There was a greater proportion of patients infected with HCV 1b achieving SVR than HCV 1a (46% vs 25%) (Figure 4.1 D and E), a trend which was also seen when separated by gender



and IFN $\lambda$ 3 genotype. The only subset for which this did not hold true was in IFN $\lambda$ 3 CC males; SVR rates in HCV 1a were higher than HCV 1b (35% and 29%) (Figure 4.1 D and E). There is little in the literature regarding differences in response to pegIFN/RBV therapy between the HCV genotype 1 subtypes, as most studies report on differences between genotypes rather than subtypes. Interestingly, the studies that have focused on response at the subtype level found contradicting results. Jensen et al [177] found similar results to that presented here; that HCV 1b was more responsive to treatment. The opposite was observed by Pellicelli et al, who found that dual therapy is more effective on patients infected with HCV 1a than HCV 1b [178]. Due to the contradictions and low number of studies in this area, it shows that our knowledge is limited and has the potential to be expanded.

In recent years, there has been more focus on the HCV genotype subtypes 1a and 1b due to the potential differences in response to new treatment options. Improved SVR rate is seen in HCV 1b compared to 1a in difficult to treat patients on pegIFN/RBV who have then gone on to a course of boceprevir, with an SVR of 47% and 25%, respectively [179]. This difference is partly due to a higher resistance rate, and double mutations, in HCV 1a virus compared to HCV 1b [180]. The difference in SVR between HCV 1a and 1b is greater in IFN $\lambda$ 3 CT and TT patients [179], which is likely due to the already high SVR in CC patients on dual therapy. There is some evidence that there are two distinct clades within the HCV 1a subtype as determined by phylogenetic analysis, with sequence changes throughout the HCV genome [181]. The isolates in this study were from the United States and Europe, and showed no specific geographic distribution. Although this is well outside the scope of the present study, it would be of interest to determine if this confers any difference to the IFN $\lambda$ 3 response.

The difference in response rates between HCV genotype 1a and 1b is also seen in interferon-free treatment regimens [182]. Distributing differing doses of faldaprevir (a protease inhibitor) and deleobuvir (a non-nucleoside polymerase inhibitor) to treatment naïve individuals, HCV 1a patients exhibit a clearance rate of 11-47%, compared to 56-85% for HCV 1b patients. IFN $\lambda$ 3 also played a role in viral clearance under this treatment regimen, with an SVR rate of 58-84% for IFN $\lambda$ 3 CC patients, and 33-64% for non-CC patients infected with HCV genotype 1 [182].

The distribution of IFN $\lambda$ 3 genotype within the HCV genotype 1 patients (38% CC, 50% CT, 12% TT) was similar to the overall population (40.5% CC, 47.5% CT, 12% TT), with IFN $\lambda$ 3 CC being slightly less common in the HCV 1 subset. There was also an observed gender difference in IFN $\lambda$ 3 distribution within the HCV 1 patients. The heterozygote CT was highly represented in females (56%, compared with 47% in both HCV 1 males, and the overall population), and the non-responsive TT was lower at just 9.5% (13.5% in males, 12% in overall population) (Table 4.2). This may also be a reflection of higher spontaneous clearance of HCV seen in females and in IFN $\lambda$ 3 CC patients. Whether these observations are a reflection of overall population or due to the inclusion criteria is unclear.

Low numbers of TT patients in this study is a reflection of the inclusion criteria; patients in this study had HCV testing performed at specific time points, including a post-therapy bleed. Many TT patients would not have met the criteria as treatment is discontinued if there is no viral response after 12 weeks of treatment [183, 184].. It was for this reason that further analysis was carried out as IFN $\lambda$ 3 CC vs non-CC, a plan that many other studies have

followed [94, 182, 185-190]. It would be of interest to explore what proportion of TT patients had their treatment ceased, compared to those whose ongoing testing was performed at different laboratories. The ability to follow these patients longitudinally may have added value to these studies.

When looking at the HCV 1 patients as a whole, SVR was observed in 41% of IFN $\lambda$ 3 CC patients, compared to 26% of non-CC patients. This is consistent with the 2-3 fold increase in SVR for IFN $\lambda$ 3 CC patients observed in previous studies [41, 191]. Once these patients were stratified according to HCV 1 subtype, it was observed that the benefit afforded by IFN $\lambda$ 3 CC was greater in HCV 1b patients than HCV 1a patients.

Although there is little data in HCV mono-infected patients, there have been studies investigating the interaction between HCV genotype 1 subtypes and IFN $\lambda$ 3 genotypes in HIV/HCV co-infected patients. These studies found the opposite effect of IFN $\lambda$ 3 effect within HCV 1 subtypes as observed here; the greatest influence of IFN $\lambda$ 3 genotypes was seen in HCV 1a [192, 193]. Again however, these studies are limited, and whether the dynamics of being co-infected with HIV has some interplay needs to be determined.

#### **4.5: Summary**

A major disadvantage to this study is lack of follow up data for those who discontinued treatment. As a laboratory providing diagnostic services for a number of pathology centres and hospitals across Australia, there no access to discontinuation circumstances. Once treatment has stopped, monitoring also ceases. As the inclusion criteria for this

retrospective analysis included HCV RNA status analysis at various time points, there is more than likely a greater number of patients who have failed to respond to treatment than is represented here.

From the data provided here, it appears that the beneficial effect of IFN $\lambda$ 3 CC is greater in HCV 1b than 1a in females, although this may be limited by small sample numbers. Collaborations with hepatitis clinics in the future may provide further insight into the difference in effect IFN $\lambda$ 3 genotype has on HCV subtypes 1a and 1b.

## Chapter 5: IFN $\lambda$ 3 Genotype Distribution Throughout Australia

### 5.1: Introduction

Following on from the previous Chapter, in which there was an observed distribution pattern between IFN $\lambda$ 3 genotype and HCV genotype, it became feasible that immigration patterns throughout Australia could be extrapolated by analyzing the frequencies of the three IFN $\lambda$ 3 genotypes.

In Australia, the population is dynamic, and contains a distinct mixture of ethnicities compared to other western countries. Due to the combination of geographical isolation and colonial British early settlement, the Australian population consists of a largely Caucasian and European majority, but with an increasing number of Asian and African immigrants. Data from the Australian Bureau of Statistics shows that as of 2013, 27.7% of Australian residents were born overseas, up from 23.6% in 2003 [194]. Immigration patterns have changed over time, and often occur in waves. The Aboriginal Australians lived in isolation until the British claimed Australia as a territory in 1788 when an influx of Britons, including convicts and farmers, laid claim to the land. Following this, a large number of Chinese settlers arrived in Australia during the gold rush in the 1850s. There was a large influx of western Europeans following the Second World War, and more recently South East Asians, and most recently we have seen migration from India and Africa. As of 2006, the top five countries of birth for immigrants to Australia were the UK, New Zealand, China, Italy, and Vietnam [195].

As a largely multicultural country, the findings from overseas studies involving treatment response, as well as genome wide association studies (GWAS) in isolated populations, often cannot be extrapolated to the Australian population. There is however, the advantage of being able to consolidate data from a number of different countries and populations, and seeing how the combined information relates to a diverse population.

GWAS tend to be performed in affluent countries, where they have the resources and the skilled workforce to undertake large studies. In the USA, a number of GWAS have been undertaken, leading to the discovery of genes associated with a number of phenotypes, including toxicity to cancer treatments [196], late-onset Alzheimer's disease [197], and response to HCV treatment [41]. The cohorts in these American studies often consist of three defined population groups; whites (Caucasian), blacks (African Americans), and Hispanics [41, 165, 198].

Australia has approximately 230,000 people infected with HCV, of which around 58,000 have moderate to severe liver disease [199]. Migrants to Australia from high-prevalence countries constitute approximately 7.2% of the total HCV cases [200]. Of immigrants with HCV antibodies, approximately 10.9% had already been exposed prior to migrating from countries of high prevalence [201]. High prevalence countries include Egypt (22%) [17], Vietnam (6.1% ) [160] and China (3.2%) [160].

IFN $\lambda$ 3 genotypes and allele frequencies vary among the global population. The protective allele (CC in rs12979860) is almost exclusive in populations in areas such as South-East Asia [53, 93], while the major allele in populations in Africa, and those of African descent

(rs12979860 TT), offers poor-response to HCV treatment [41, 53]. Caucasian and European populations are intermediate; often carrying at least one copy of the good response allele (ie, CC or CT for rs12979860) [32, 156, 191].

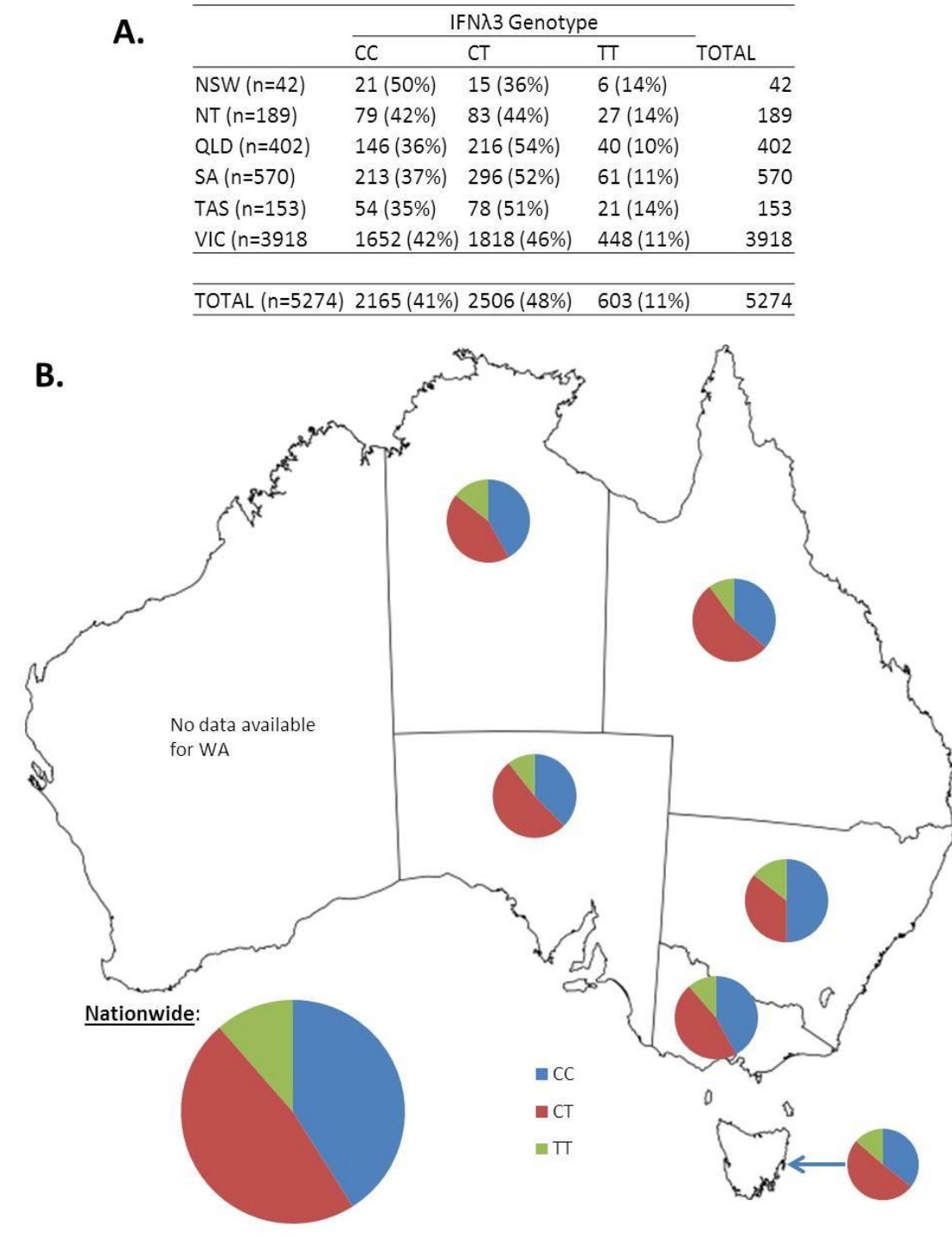
The aim of this Chapter was to determine if the frequency of IFN $\lambda$ 3 genotypes among different Australian States and Territories varies, and if there is any significance to it in terms of ethnicity and migration.

## **5.2: Methods**

The VIDRL database was screened for patient location according to State and Territory, and analysed according to IFN $\lambda$ 3 genotype. Ethnicity was crudely assigned to each patient based on surname. The three distinct groups analyzed were 'Asian', 'Non-Asian', and 'Unknown'. 'Unknown' accounts for the specimens which have been received but coded for anonymity.

## **5.3: Results**

IFN $\lambda$ 3 (rs12979860) genotyping had been performed on a total of 5274 patients between January 2010 and December 2012. Being a State Reference Laboratory for Victoria, the majority (74%) of the samples were from Victorian residents. NSW was the only State or Territory which had patients with CC as the predominant genotype. In each of the other States and Territories the main patient genotype was CT (Fig. 5.1).



**Figure 5.1: rs12979860 Genotype distribution patterns according to States and Territories in Australia. (A) This shows the number of IFN $\lambda$ 3 genotypes for each State and Territory, which are then represented graphically in (B). No data was available for WA. All data from ACT is included in NSW dataset.**



The total nationwide genotype distribution (48% CC, 41% CT, 11% TT) represents an allelic frequency of 69% C allele, and 31% T allele. This data has been included on the global map from previously published data (Fig. 5.2).



**Figure 5.2: Global allele frequency of rs12979860. Pie charts show frequencies of C (green) and T (blue) alleles for regions assessed. Data acquired from the present study has allowed for IFN $\lambda$ 3 genotype to be applied to the global map. Modified from Thomas *et al*, 2009 [53]**

Of the 5274 patients tested for IFN $\lambda$ 3 genotype, 9% were Asian, according to the ethnicity criteria (Table 5.1 A). The 'unknown' proportion was minor, constituting just 3.2% of samples which had tests performed. The genotype distribution in the Asian vs Non-Asian is marked, with a higher number of Asians carrying the homozygous CC genotype (74%, Table 5.1), while the non-Asians are predominantly CT (50%, Table 5.1). The presence of the CC genotype in the Asian population is substantially higher than in the general population (41%, Figure 5.1). The ethnicity data and the State data were combined to provide the distribution

of ethnicities in each State (Table 5.1). Victoria had the highest number of Asian patients included in this study (10.13%), and Tasmania had the lowest (1.31%).

**A.**

	IFN $\lambda$ 3 Genotype			
	TOTAL	CC	CT	TT
Asian	474 (9%)	351 (74%)	114 (24%)	9 (2%)
Non-Asian	4633 (88%)	1746 (38%)	2319 (50%)	568 (12%)
Unknown	167 (3%)	68 (40%)	73 (44%)	26 (16%)
TOTAL	5274	2165 (41%)	2506 (48%)	603 (11%)

**B.**

	Asian	Non-Asian	Unknown
NSW (42)	3 (7.10%)	39 (92.8%)	0 (0%)
NT (189)	10 (5.29%)	113 (59.79%)	66 (34.92%)
VIC (3918)	397 (10.13%)	3462 (88.36%)	59 (1.51%)
QLD (402)	13 (3.23%)	361 (89.80%)	28 (6.97%)
SA (570)	49 (8.60%)	513 (90.00%)	8 (1.40%)
TAS (153)	2 (1.31%)	145 (94.77%)	6 (3.92%)
TOTAL	474 (9%)	4633 (88%)	167 (3%)

**Table 5.1: Breakdown of ethnicity according to States. Determination of the distribution of the IFN $\lambda$ 3 genotype in relation to ethnicities (based on surnames) nationwide (A) and in States and Territories (B).**

#### 5.4: Discussion:

Once the 5274 IFN $\lambda$ 3 genotypes had been analysed, it was found that the Australian distribution is similar to the proportions found in other western societies (CC=41%, CT=47.5%, TT=11.5%; Figure 5.1A) [53]. As a country that had many waves of migration for the past 50 years, the Australian distribution was not unexpected. This study was investigating on the basis of Asian vs non-Asian ethnicity; 9% of the study population was Asian, 88% were non-Asian, and for 3% the ethnicity could not be established.

Surprisingly, the presence of the IFN $\lambda$ 3 CC genotype seems to be low. The majority of Asian patients included in the analysis were from South-East Asia, including Vietnam, where IFN $\lambda$ 3 CC is almost exclusive [53]. This study however has identified only 74% of Asians carrying the protective CC genotype, with 24% and 2% having CT and TT genotypes, respectively. An explanation for this, which reveals a major limitation of this study is the crude assignment of ethnicity according to surname. By assigning ethnicity in such a fashion, it disregards the fact surnames can change due to events such as marriage. In a multicultural population such as Australia, it is not uncommon for people to have mixed heritage. Once expanded to second and third generation Australians, an Asian surname (for example) may be retained, but the Asian heritage may only represent a minor percentage of that person's genetic makeup. This helps to explain the higher than expected number of Asian patients (based on the study's criteria) with the CT genotype. If ethnicity had been self-reported, the figures would perhaps reveal a different result.

This study was unable to determine a link between the relative distribution of IFN $\lambda$ 3 genotype and Asian/non-Asian ethnicity within Australia. Although 74% of Asian patients were IFN $\lambda$ 3 CC, they constituted 9% of the overall population (Table 5.1A). In New South Wales, the State with the highest percentage of IFN $\lambda$ 3 CC (50%), had only the 3<sup>rd</sup> highest Asian population at 7%. Victoria had the highest proportion of Asians in the study (10%), and IFN $\lambda$ 3 CC constituted 42% of the population.

In Queensland, South Australia, and Tasmania there was a different distribution compared to New South Wales and Victoria. These States all had the heterozygote IFN $\lambda$ 3 CT allele as

the majority (53%, 52%, and 51% respectively; Figure 5.1A). They also had low numbers of Asian patients (3%, 9%, and 1% respectively). These States not only showed a shift away from the IFN $\lambda$ 3 CC genotype, but also from the IFN $\lambda$ 3 TT genotype. Tasmania is an exception to this, with a higher number of IFN $\lambda$ 3 TT patients than the national average (13% and 11% respectively; not significant). These figures are consistent with studies looking at Caucasian patients as a subset of the population [165, 202]. Ethnicity was unable to be determined for the majority of specimens from Northern Territory, so little information was gained from that region. This is due to specimens being coded prior to being dispatched to VIDRL. While there is no direct link between Asian migration and IFN $\lambda$ 3 distribution in Australia, there may be an association with immigration in a broader sense. There was insufficient information in the VIDRL database to draw any insight regarding IFN $\lambda$ 3 distribution in Indigenous Australian individuals, though other studies have shown varied results. One study showed that 33% of Indigenous Australians have the CC genotype (n=33), though it is unclear if the ethnicity is self-reported, or if it includes individuals of mixed ethnicity [203]. Another small study showed that of five Indigenous Australians tested in the NT, all of them were genotype CC [204]. The small number of samples examined in these studies, and the differing IFN $\lambda$ 3 genotype prevalence, illustrate that more studies are required in this area.

Since the introduction of HCV antibody screening of blood products in 1990, newly diagnosed individuals with HCV tend to be from marginalized populations. Because of this, there is little in the literature extrapolating a link between HCV exposure and immigration patterns into Australia. There are excellent datasets of HCV infected individuals in Australian populations, but due to the nature of the populations, the patient details remain confidential; these datasets include those of people who inject drugs [24, 205] and prison

studies [206, 207]. Concealed information includes any inference of ethnicity, although other information such as date of birth and gender is included in test request forms.

One obvious limitation to linking the data to the wider Australian population, is that the majority of those tested for IFN $\lambda$ 3 genotype are HCV RNA positive; the testing is performed either prior to initiating to treatment, or assessing reasons of treatment failure. If the general population were to be tested, perhaps more information would be able to be extrapolated. To infer anything of significance from this study, further information would need to be collected, such as (1) proportion of Asian/non-Asian injecting drug users compared to overall population, and (2) proportion of Asian/non-Asian prisoners compared to overall population. This information would then need to be analysed State-by-State, and back-referenced to the IFN $\lambda$ 3 genotypes.

By sorting the population into Asian vs non-Asian, this study was not able to deduce any ethnic distribution or migration patterns according to IFN $\lambda$ 3 genotype, due to the high number of CC patients in the non-Asian population. However, there is a distribution pattern between States; the more ethnically diverse states (New South Wales and Victoria) have a wider spread of IFN $\lambda$ 3 genotypes, whilst the more ethnically static states (Queensland, Tasmania, and South Australia) have a higher proportion of IFN $\lambda$ 3 CT (Figure 5.1A), which is common in Caucasian populations. This study would have greater power with a larger scope of investigation, including self-reported ethnicity.

## Chapter 6: Gene Duplication, Phylogeny and Function of the Interferon $\lambda$ Family Members

### 6.1 Background

As discussed in Chapter 1, members of the interferon  $\lambda$  (IFN $\lambda$ ) family are increasingly being acknowledged as critical mediators of immune and inflammatory responses, notably the antiviral response to RNA viruses and sterile inflammatory responses, including asthma and other allergic reactions. Recently, their antineoplastic potential has also been investigated [208].

A large body of information relating to the IFN $\lambda$ s has accumulated since 2009, when IFN $\lambda$ 3 first began attracting attention from researchers and clinicians as a result of genome-wide association studies (GWAS) that revealed the association between SNPs upstream of the *IFN $\lambda$ 3* promoter and spontaneous and treatment-induced clearance of HCV infection (details in Chapter 1). Despite their recognized importance, relatively little is currently known about the molecular mechanisms that regulate the expression and biological activities of the IFN $\lambda$  family, especially its most recently identified member, IFN $\lambda$ 4. This Chapter reports the results of a preliminary exploration of this topic, concentrating on similarities and differences between the individual genes and their products. It is divided into two sections - theoretical and experimental – and, due to the paucity of available information about IFN $\lambda$ 4, is restricted mainly to the canonical members IFN $\lambda$ s 1 to 3.

### 6.2: Part 1: Theoretical

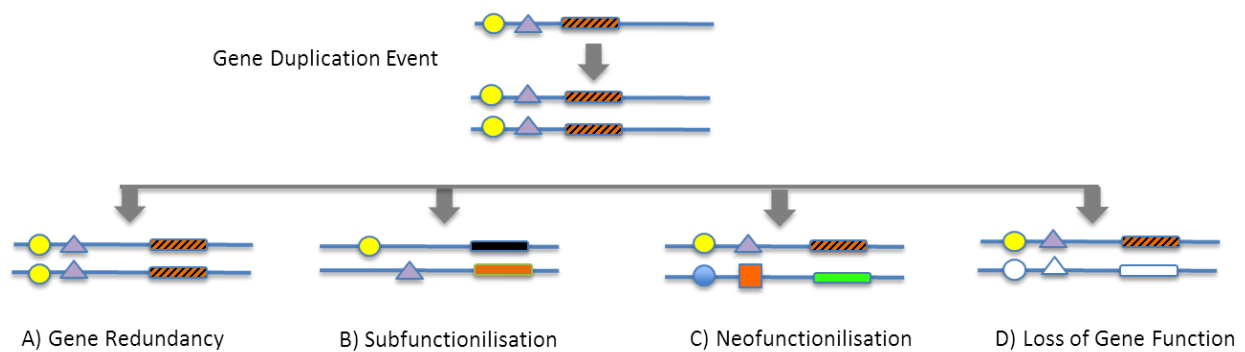
#### 6.2.1 Aims and Methods

The first part of this study aimed to extend current understanding of the evolution and function of the IFN $\lambda$  family by collecting, collating and analysing data mined from, or generated with the aid of, a variety of resources that are freely accessible available *via* the internet. Individual sources are identified in the text and figure captions.

## 6.2.2 Results and Discussion

### 6.2.2.1 Gene duplication

Gene duplication is a ubiquitous phenomenon in all life kingdoms and arguably the main source of genetic innovation and robustness [209, 210]. Following a gene duplication event, there are several possible fates for the products generated. There is an energy cost to the organism to maintain a redundant gene, which may eventually result in negative selection. Alternatively, a gene may evolve to perform a highly specific function, become redundant when that function is no longer required, and, because it is no longer under positive selection pressure, accumulate mutations that eventually either endow it with a new function or transform it into a non-functional pseudogene (Figure 6.1) [210].



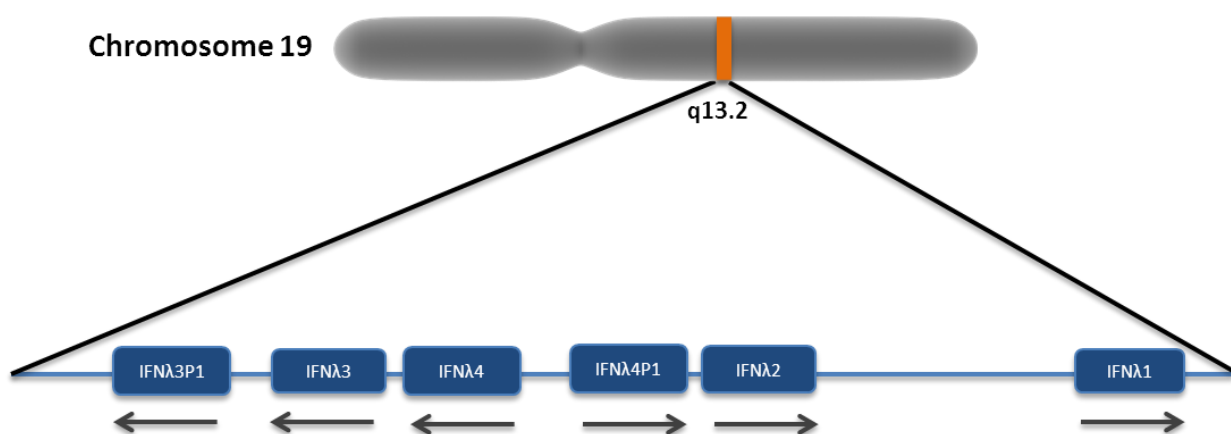
**Figure 6.1: Alternative outcomes of gene duplication.** There are at least four possible outcomes of gene duplication, which are not mutually exclusive. **A)** Duplicated genes both remain functional, which increases the capacity to deal with a specific need and is generally a short-term outcome. **B)** Subfunctionalisation describes the situation where both copies retaining partial function and both are required to achieve full function. **C)** Neofunctionalisation may occur when only one product is sufficient for function, freeing the duplicate to evolve a new function, which can involve mutations that affect both regulatory elements and coding regions. **D)** A redundant copy accumulates mutations that eventually silence its expression, although the gene may be retained.

The IFNs, as key regulators of inflammation and immunity, provide some intriguing examples of gene duplication and its *sequelae*. The fact that the genomes of all vertebrates contain genes for three families of IFNs with numerous subtypes that perform seemingly redundant or overlapping functions should not be surprising, considering that the immune system needs to maintain a large reserve capacity so as to be prepared to respond immediately to a wide variety of “expected” challenges from organisms that already present as part of the bodily ecosystem (recently christened the “holobiome”) [211]. In addition, it

needs to be capable of responding quickly to both “unexpected” challenges posed by new, rapidly evolving pathogens and large environmental changes. Type I and III IFNs, which are synthesised and secreted mainly in response to viral infection and other stressful stimuli, communicate stress signals *via* the JAK-STAT pathway, which is also used by other cytokines [212]. The apparent functional redundancy of the cytokines that share this signalling pathway is explicable in terms of location and timing, which are determined respectively by the tissue distribution of specific cytokine receptors and the characteristics of individual cytokines.

#### 6.2.2.2 *IFN $\lambda$ locus topography, phylogeny and evolution*

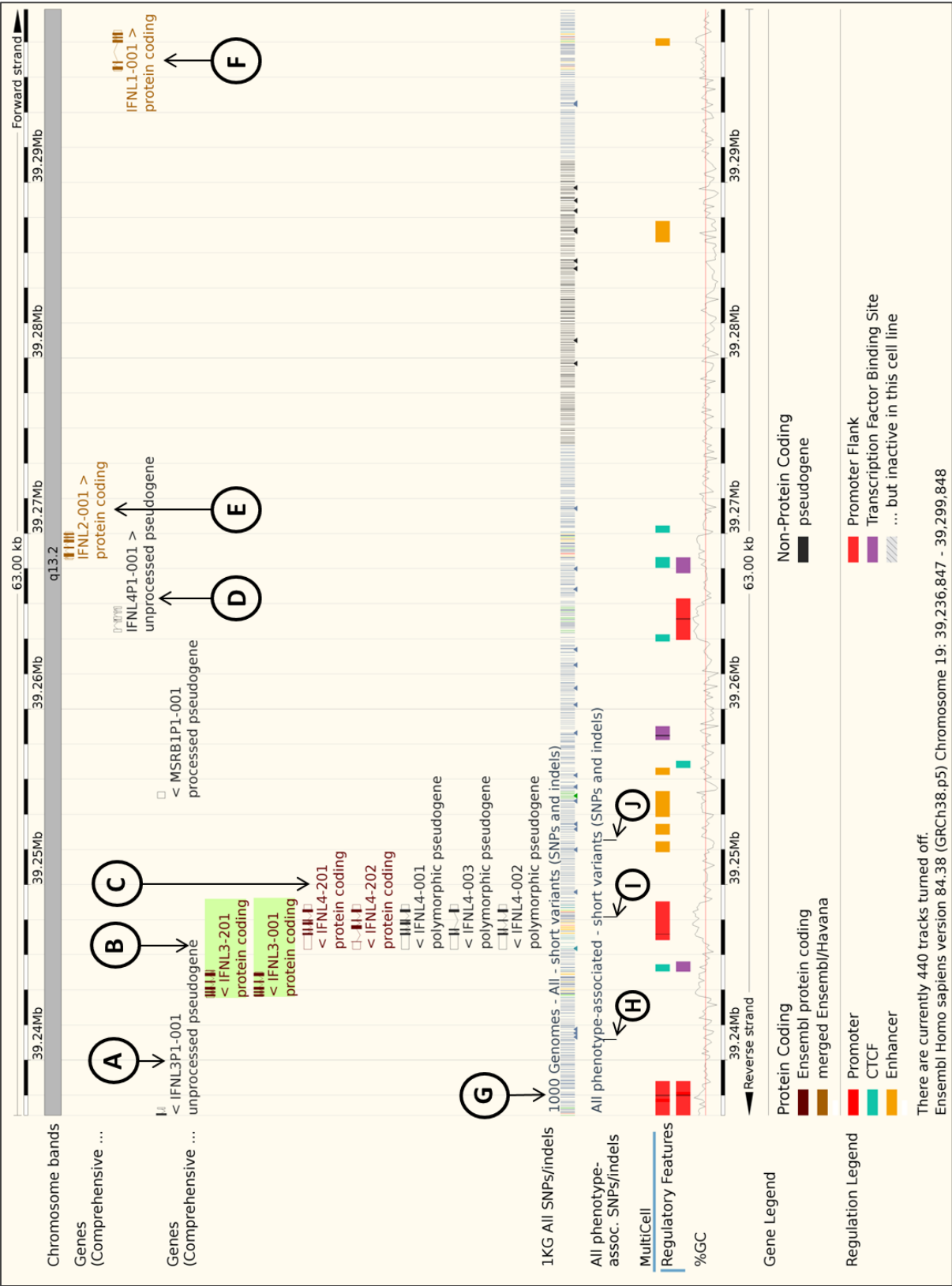
In the case of the type III IFN genes, two or more duplication events have occurred, resulting in three similar functional genes in the majority of human populations (four in people of African origin). The *IFN $\lambda$ 1*, *IFN $\lambda$ 2*, *IFN $\lambda$ 3*, and *IFN $\lambda$ 4* genes are clustered together with two pseudogenes - *IFN $\lambda$ 3P* and *IFN $\lambda$ 4P* in a ~65kb region that has been mapped to band q13.2 on the long arm of human chromosome 19 (Figures 6.2 and 6.3). Synteny has been maintained during evolution, so the relative orientation of the *IFN $\lambda$*  genes is conserved amongst mammals, with *IFN $\lambda$ 1* and *IFN $\lambda$ 2* transcribed in the opposite direction relative to *IFN $\lambda$ 3* and *IFN $\lambda$ 4* [213]. *IFN $\lambda$ 1* and *IFN $\lambda$ 2* are on the forward (positive) strand, with *IFN $\lambda$ 3* and *IFN $\lambda$ 4* on the reverse (negative) strand.



**Figure 6.2: Orientation of *IFN $\lambda$*  genes (and pseudogenes) on human chromosome 19**



Figure 6.3 (following page) provides a more detailed map of the *IFN $\lambda$*  locus imported from the Ensembl Gene Browser (<http://www.ensembl.org/index.html>) [214]. The map encompasses the relative positions of IFN $\lambda$ 3P1 (A), IFN $\lambda$ 3 (B), IFN $\lambda$ 4 (C), IFN $\lambda$ 4P1 (D), IFN $\lambda$ 2 (E), and IFN $\lambda$ 1 (F). It shows there are a multitude of SNPs in the region (G), though only three have been associated with a phenotypic effect; rs12980275 (H), rs12979860 (I), and rs8099917 (J).



A high degree of mutability is predictable and essential for loci that must evolve rapidly in response to novel pathogens and environmental change. The map also shows the locations of some important regulatory elements, notably CTCF-binding and enhancer sites, which act in concert to regulate chromosome structure [215]. From the relative positions of the *IFN $\lambda$*  genes, it seems improbable that *IFN $\lambda$ 1* and *IFN $\lambda$ 4* were generated by direct duplication, as early studies suggested [216, 217]. Instead, the orientation and location of the *IFN $\lambda$ 4P1* pseudogene immediately upstream (5') of *IFN $\lambda$ 2*, suggests that the *IFN $\lambda$ 4P1*>*IFN $\lambda$ 2* pair and the oppositely oriented *IFN $\lambda$ 3*<*IFN $\lambda$ 4* pair are products of a duplication and retro-transposition event, which is supported by close sequence homology.

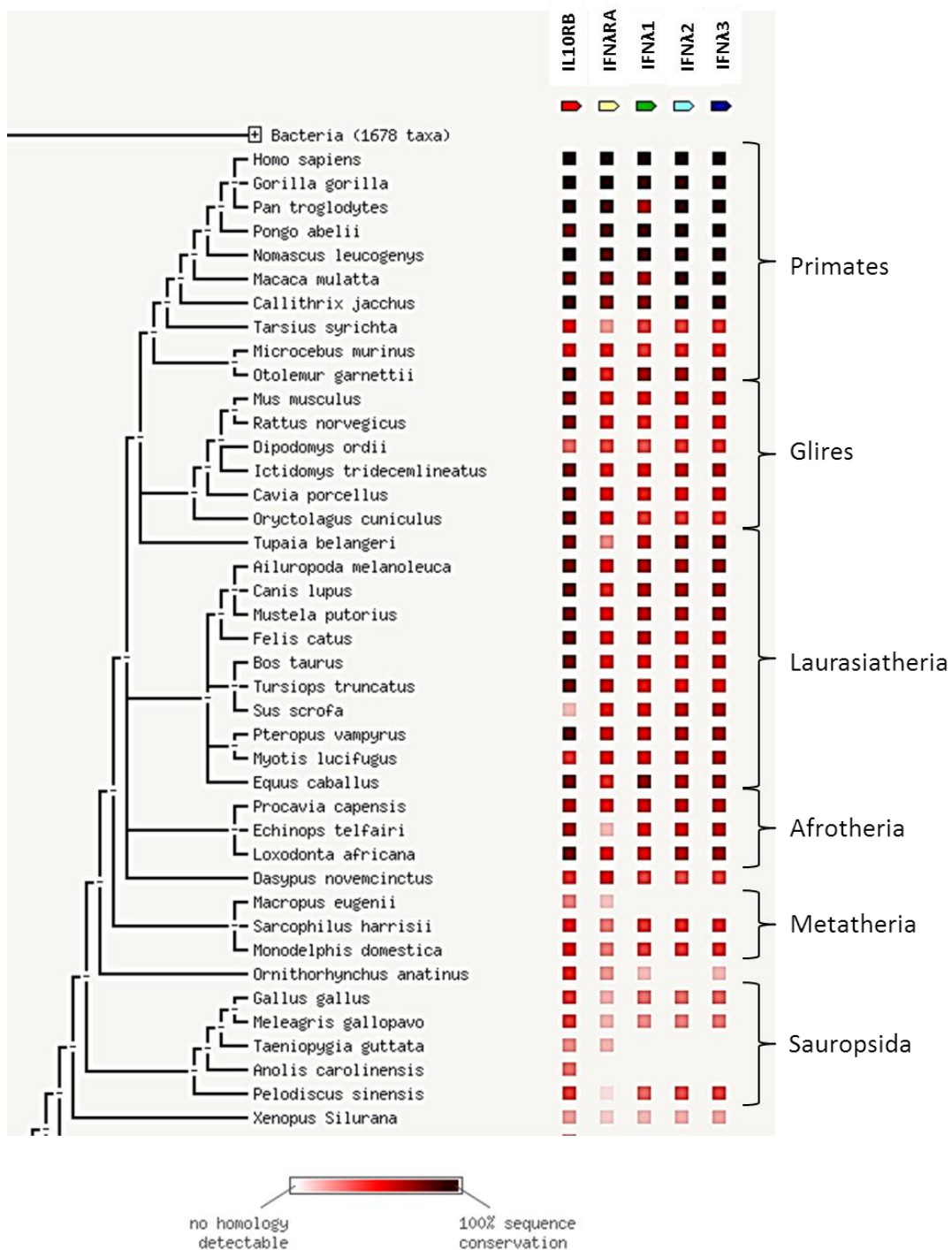
Besides suggesting that they originated by duplication, the location of the *IFN $\lambda$*  genes in two groups on opposite DNA strands has an interesting implication: one group may evolve more rapidly because, during replication, cytosine bases in the lagging strand are more prone to deamination (and therefore subsequent error-prone repair) than those in the leading strand [218]. It is also noteworthy that, of the at least six copies of the primordial human *IFN $\lambda$*  gene that have been identified, half (*IFN $\lambda$ s 1-3*) have remained functional during the >300 million years of evolution that separates primates from amphibians (Figure 6.4). Of the remaining three, two (*IFN $\lambda$ s 3P and 4P*) have become pseudogenes and one (*IFN $\lambda$ 4*) has transformed only very recently - at least in evolutionary terms - into a transcribed but non-coding pseudogene in most of the human population and other than ethnic groups of African ancestry [213].

### 6.2.3 *IFN $\lambda$* Phylogeny: Homologs of *IFN $\lambda$ s* and their Receptors in other Vertebrates

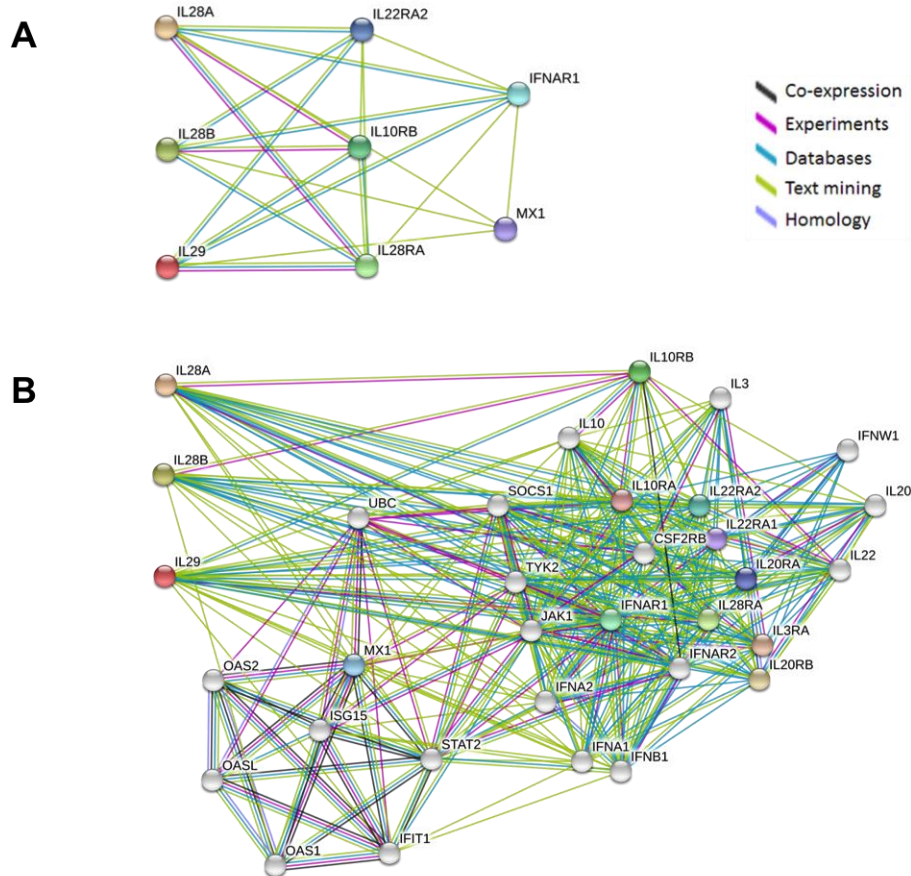
#### 6.2.3.1 *Interactions of IFN $\lambda$ s with their receptors and other proteins*

“String 10” is a free online database, which compiles and stores a large range of information about interactions between proteins based on experimental and other evidence [219]. The webserver (<http://string-db.org>) still uses the original nomenclature for *IFN $\lambda$ s*; IL28A (*IFN $\lambda$ 2*), IL28B (*IFN $\lambda$ 3*), and IL29 (*IFN $\lambda$ 1*) and does not yet include data relating to *IFN $\lambda$ 4*. Among its other functions, String10 can generate graphical representations of protein interactions as well as charts that depict protein co-occurrence and co-evolution. Figures 6.4 and 6.5 were produced using these functions.

Interacting partners of the three canonical IFN $\lambda$  proteins are depicted below (Figure 6.5A). Greater confidence may be placed in the associations determined by experiments (magenta) and databases (turquoise); other illustrated linkages may be included in the ‘text mining’ category (green). Most of the direct interactions between IFN $\lambda$ s and other proteins predicted by String 10 involve cytokine receptors and associated signaling proteins, which suggests that competition for receptors is likely to occur at high ligand concentrations.



**Figure 6.4.** Evolutionary succession and co-occurrence tree showing evidence for the existence of homologs of human IFNλs 1, 2 and 3 and their receptors, during ~ 300 million years of evolution from amphibians (Xenopus) to Primates. Receptor homologs, but not IFNλ homologs, have been identified in a few instances in which it is probable that other more primitive cytokines use the receptors and/or that IFNλ homologues exist and await discovery and characterisation.

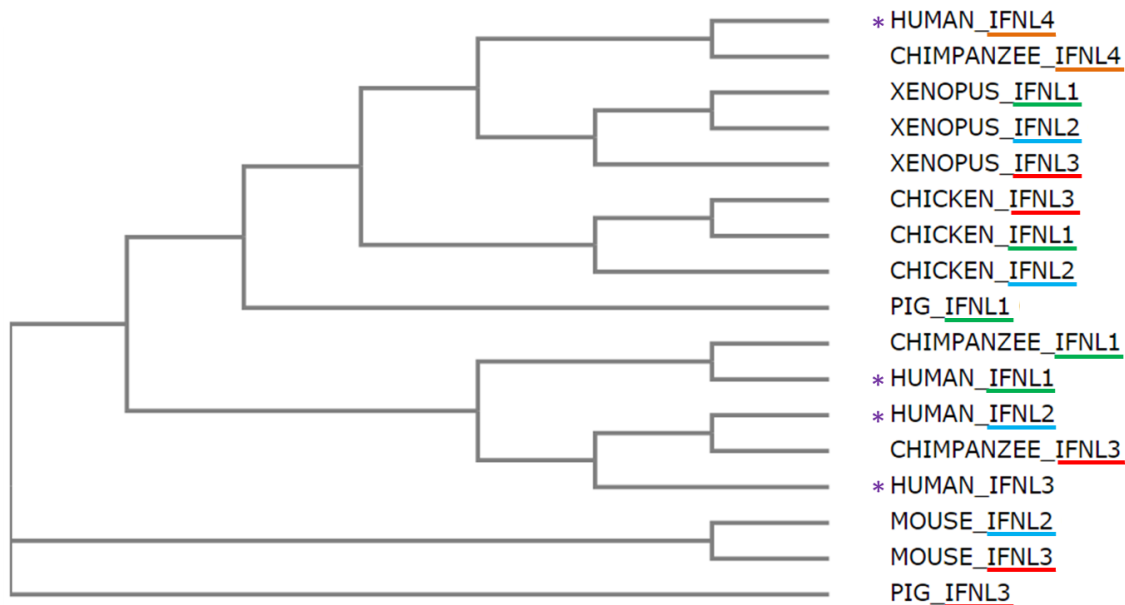


**Figure 6.5: Predicted interaction networks constructed using the String 10 server. A)** The input proteins (IL29 [IFN $\lambda$ 1], IL28A [IFN $\lambda$ 2], and IL28B [IFN $\lambda$ 3]), which are grouped at the left hand side of the diagram have many common interactions, apparently without direct interactions with each other (implying that they do not form hetero-aggregates). There are many interactions with a variety of cytokine receptors (identified by 'R' in the abbreviation), which implies that IFN $\lambda$ s may compete for receptor occupancy with each other and with similarly structured cytokines. **B)** Extending the network by increasing the number of connections reveals proteins known to be induced by IFNs (such as the OAS-group), IFN regulators (eg. IRF- group), and those involved in JAK-STAT signaling.

Currently it is uncertain which “modern” IFN $\lambda$  represents the ancestral gene with different research groups having reached different conclusions. Lasfar *et al.* [216], for example, deduced that IFN $\lambda$ 3 and IFN $\lambda$ 4 were created by duplication of IFN $\lambda$ 2 and IFN $\lambda$ 1, respectively. When their findings were published, it was widely assumed that IFN $\lambda$ 4 was invariably an unexpressible pseudogene, and they proposed that IFN $\lambda$ 4 was a highly mutated and non-functional duplicate of IFN $\lambda$ 1. A different conclusion was reached by Fox *et al.* [217], who analysed nucleotide sequence homology, and reached the opposite conclusion -, that IFN $\lambda$ 4 was the precursor of IFN $\lambda$ 1, with further duplications producing IFN $\lambda$ s 2 and 3. In their phylogenetic tree, all human IFN $\lambda$  genes reside on different branches to other species

(mouse, dog, guinea pig). Another study by Wack *et al.* [220] compared protein sequence, and grouped the *IFN $\lambda$ 4* proteins from five mammalian species (human, ferret, hedgehog, pig, and opossum) together on a separate branch from *IFN $\lambda$ s* 1 to 3, thereafter grouped according to species.

Figure 6.6 shows a phylogeny produced using Clustal Omega [221] (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) by inputting representative *IFN $\lambda$*  nucleotide reference sequences from human, chimpanzee, xenopus, chicken, pig, and mouse, which were downloaded from the Ensembl Genomes database [214]. *IFN $\lambda$ 4* sequences were only available for human and chimpanzee. All *IFN $\lambda$*  genes from individual species are clustered together, separate from the two *IFN $\lambda$ 4* sequences. In most instances the *IFN $\lambda$ 2* and *IFN $\lambda$ 3* genes from individual species cluster together.



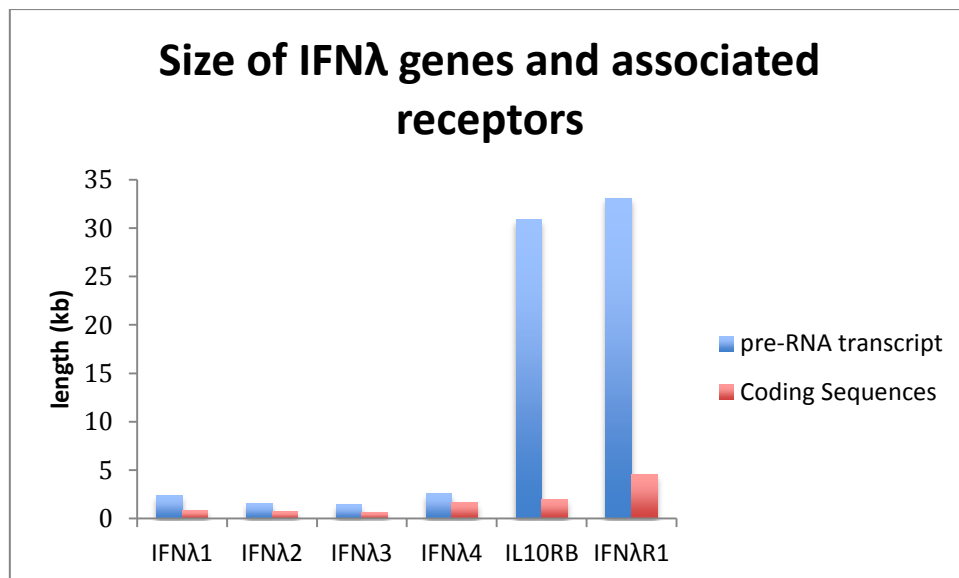
**Figure 6.6: Phylogeny of representative *IFN $\lambda$*  genes.** For visual ease, human branches are denoted with a purple asterisk (\*), and each gene homologue is underlined with a separate colour (*IFN $\lambda$ 1* is green, *IFN $\lambda$ 2* is blue, *IFN $\lambda$ 3* is red, and *IFN $\lambda$ 4* is orange).

#### 6.2.4 Structures of the *IFN $\lambda$* genes and the genes that encode their receptors

IFN $\lambda$ s share a common cell surface receptor, consisting of two subunits, IFN $\lambda$ R1 and IL10RB [222]. Compared to the genes that encode the IFN $\lambda$  ligands, the cytokine receptor genes are much larger and more complex, with a large intron:exon ratio and complex non-coding 3' and 5' control regions (Table 6.1 and Figure 6.8).

Gene	pre-RNA transcript length (kb)	Coding Sequences (kb)	translated product size (aa)	exons/ introns	ratio coding: non-coding
<i>IFN<math>\lambda</math>1</i>	2.35	0.857	200	5/5	0.36
<i>IFN<math>\lambda</math>2</i>	1.58	0.737	200	6/6	0.47
<i>IFN<math>\lambda</math>3</i>	1.40	0.656	196	6/6	0.47
<i>IFN<math>\lambda</math>4</i>	2.54	1.646	179	6/6	0.65
<i>IL10RB</i>	30.88	1.963	325	7/6	0.06
<i>IFN<math>\lambda</math>R1</i>	33.1	4.55	520	7/6	0.14

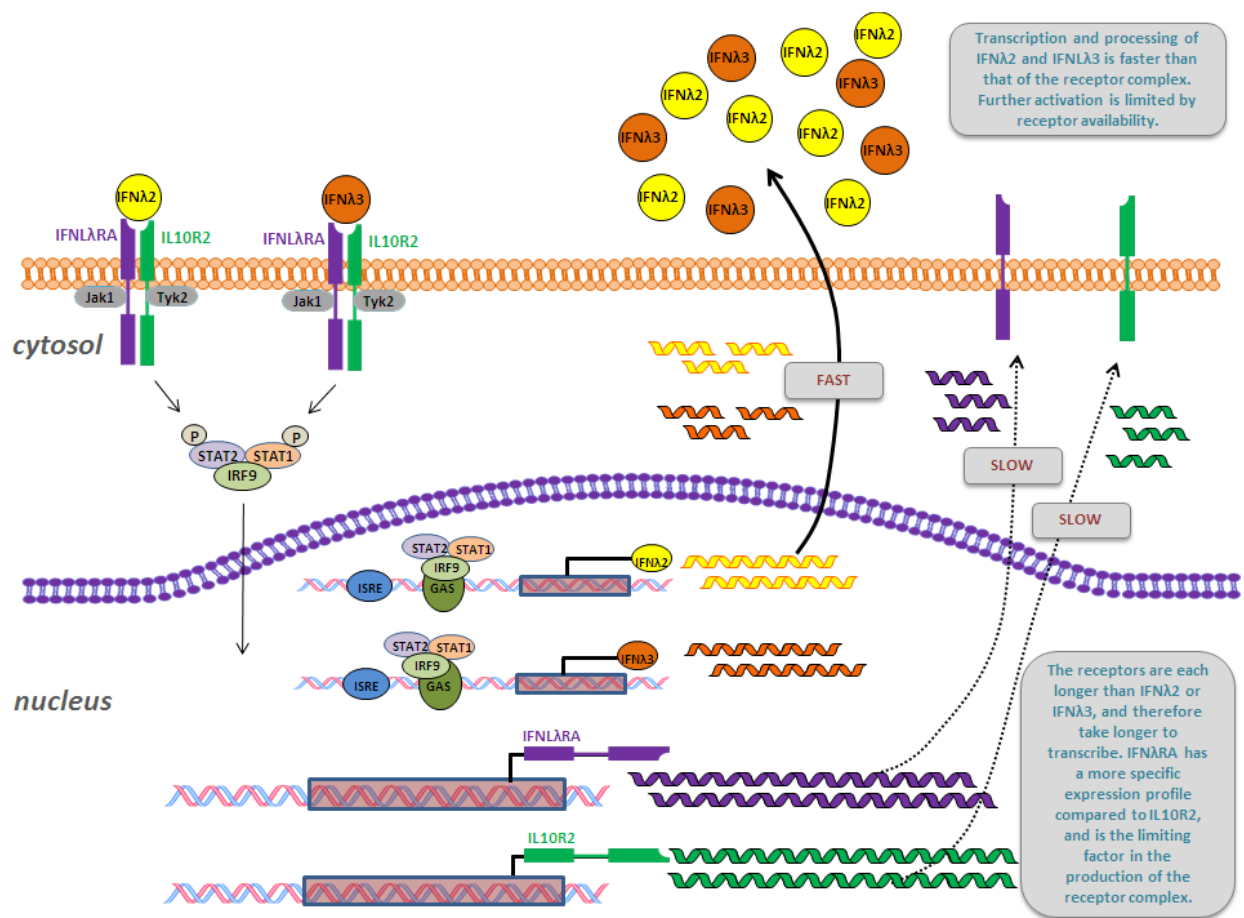
**Table 6.1: Comparison of gene product sizes of *IFN $\lambda$*  genes, and their receptors.**



**Figure 6.8: Size comparison of *IFN $\lambda$*  genes (*IFN $\lambda$ 1-4*), and the associated receptors, IL10RB and IFN $\lambda$ R1. The receptors not only have a much larger pre-RNA transcript than the IFN $\lambda$ s, but the ratio of coding to non-coding is also much lower, showing the immense control mechanisms in place in the receptors.**



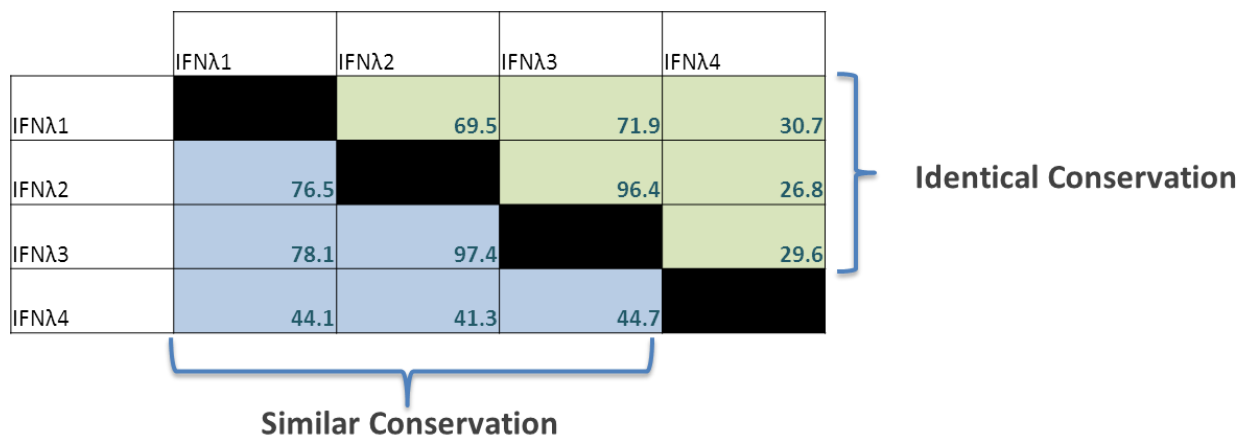
The size difference implies that the small cytokine molecules could be synthesised *de novo* in a small fraction ( $<0.1$ ) of the time needed for *de novo* synthesis of their larger cognate receptors. Consequently, during chronic infection or inflammation, depletion of the receptor population may eventually become a factor that limits response (Figure 6.9).



**Figure 6.9:** Long-term activation of IFN $\lambda$ 2 and IFN $\lambda$ 3 may be limited by receptor (IFN $\lambda$ RA and IL10R2) availability. Due to the relative size difference, the receptors take a much longer time to transcribe than the cytokines. Under normal conditions, this is not problematic. Once a chronic infection has established, the rate at which IFN $\lambda$ 2 and IFN $\lambda$ 3 are produced is increased, and the equilibrium may be disrupted.

### 6.2.5 Comparison of Individual IFN $\lambda$ genes and products

IFN $\lambda$ 1 shows more similarities to IFN $\lambda$ 2 and IFN $\lambda$ 3 than to IFN $\lambda$ 4, which is the least similar to the other IFN $\lambda$ s, showing only ~30% amino acid identity with IFN $\lambda$ s 1 to 3 (Figure 6.10) [223]. The two most closely related members of the family are IFN $\lambda$ s 2 and 3, which share 96% homology at the amino acid sequence level.



**Figure 6.10: Amino acid conservation comparison between IFN $\lambda$ 1-4. Data from Laidlaw *et al* [223].**

IFN $\lambda$ 2 and IFN $\lambda$ 3 genes both have six exons [224] and are highly homologous in upstream and downstream flanking regions, as well as in the coding region. The structure of their promoters is very similar [225]. A functional copy of the IFN $\lambda$ 1 gene has been retained along with functional genes for both IFN $\lambda$ 2 and IFN $\lambda$ 3, despite the latter's translated products having almost identical amino acid sequences, which implies that all three proteins might be needed for optimal response. Alternatively, the existence of one or two redundant genes might be a stage in an ongoing evolutionary progression, by which gene functions are being continuously refined and discarded.

Alignment of the amino acid sequences of the human type III IFN family (Figure 6.11) reveals a highly conserved C-terminal PDZ-domain binding sequence, which is masked by a nine amino acid extension in IFN $\lambda$ 1. PDZ domains are a common structural domain of 80-90 amino-acids found in signalling proteins and are one of the most evolutionarily conserved protein-protein interaction domains found in both pro-and eukaryotes [226]. The alignment

also shows a potentially reactive tyrosine (Y160) residue present only in IFN $\lambda$ 2, where the other IFN $\lambda$ s have a basic residue (histidine, H or lysine, K). The protein-protein interaction databases MoDPepInt [227] and ELMS [228] predict that Y160 will mediate binding to Src homology 2 (SH2) domains in protein partners that include the suppressors of cytokine signaling SOCS2 and SOCS5. This is significant because the SOCS family act to inhibit JAK-STAT signaling.

When IFN $\lambda$ 2 and IFN $\lambda$ 3 are aligned without IFN $\lambda$ 1 and IFN $\lambda$ 4, the high level of conservation is easily viewed (Figure 6.12A). The secondary structural features above the alignments (Figures 6.11 and 6.12A) are based on a model of the IFNL2 protein generated using Raptor X [229] (Figure 6.12B).

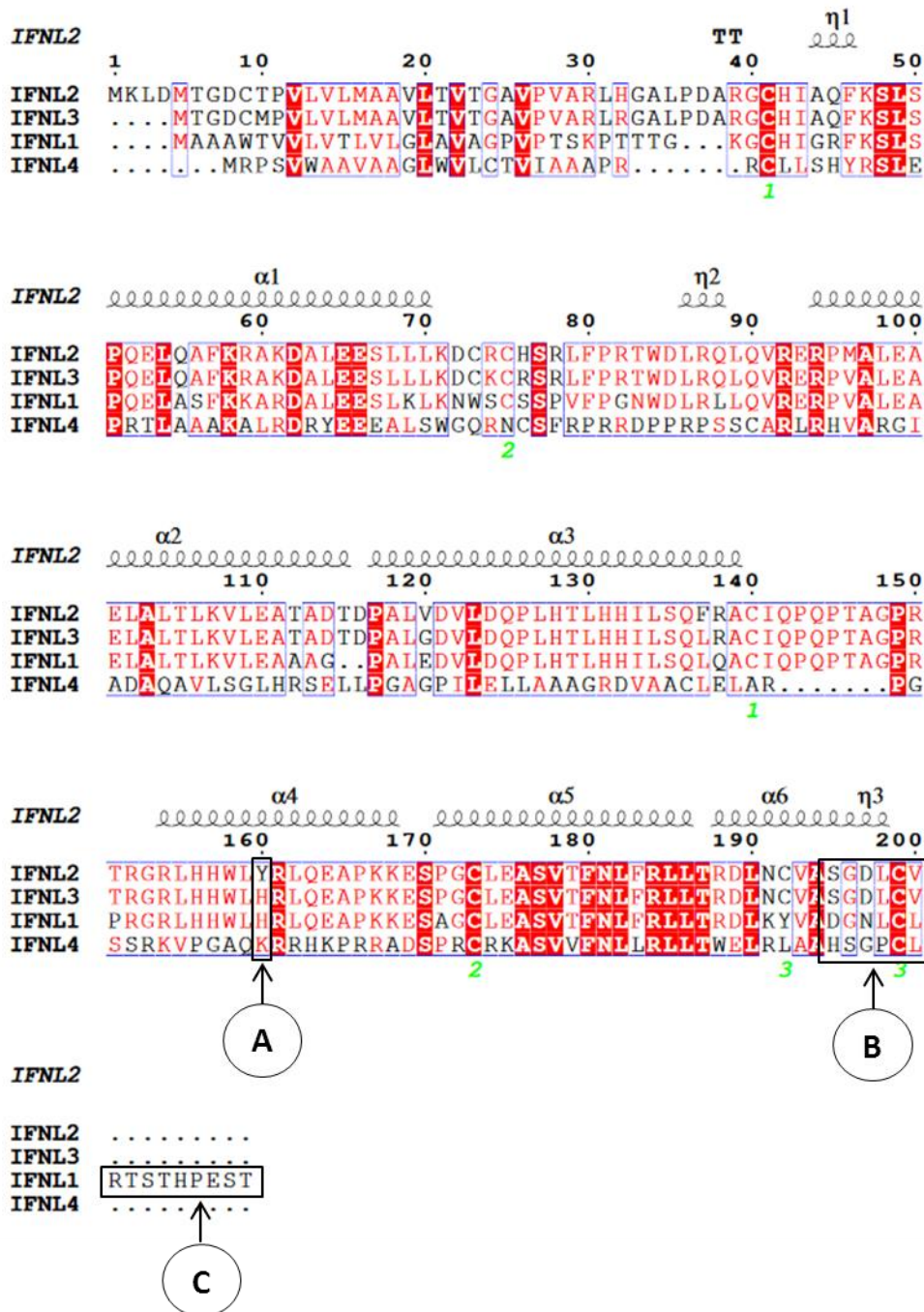
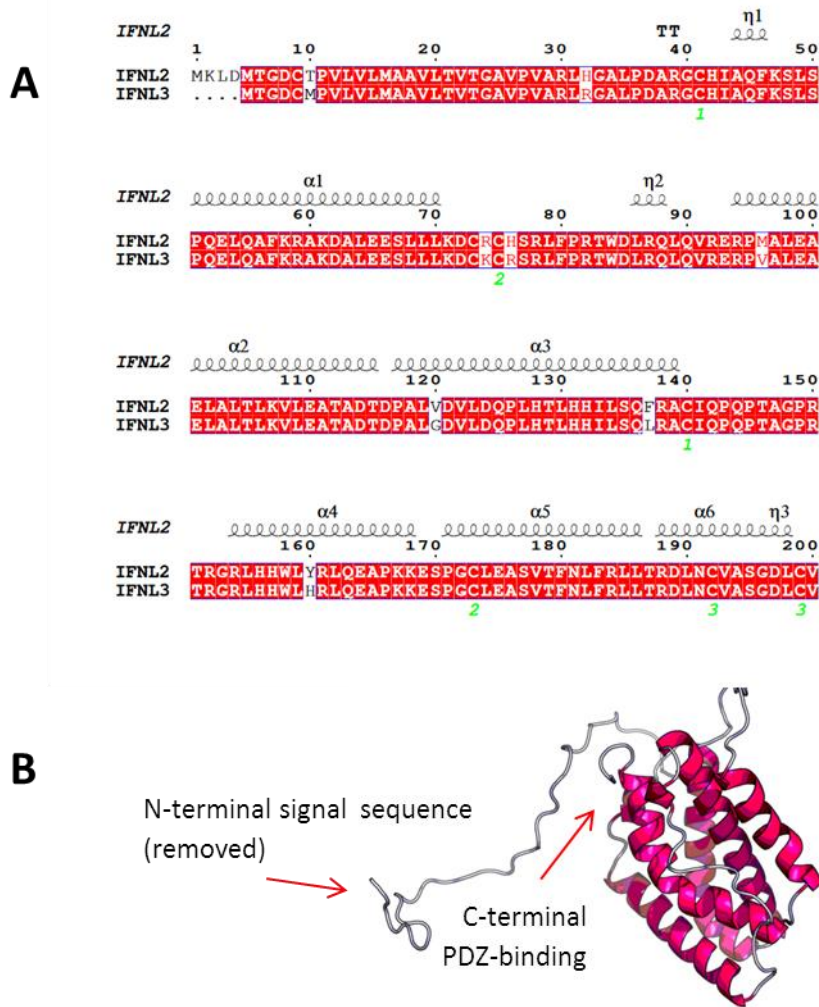


Figure 6.11: Alignment of the amino acid sequences of the human type III IFN family. The alignment shows a potentially reactive tyrosine (Y160) residue present only in IFN $\lambda$ 2 (A), where the other IFN $\lambda$ s contain a basic residue (histidine (H) or lysine (K)). Each sequence has a C-terminal PDZ-domain binding sequence (B), masked in IFN $\lambda$ 1 by a nine amino acid extension in IFN $\lambda$ 1 (C).



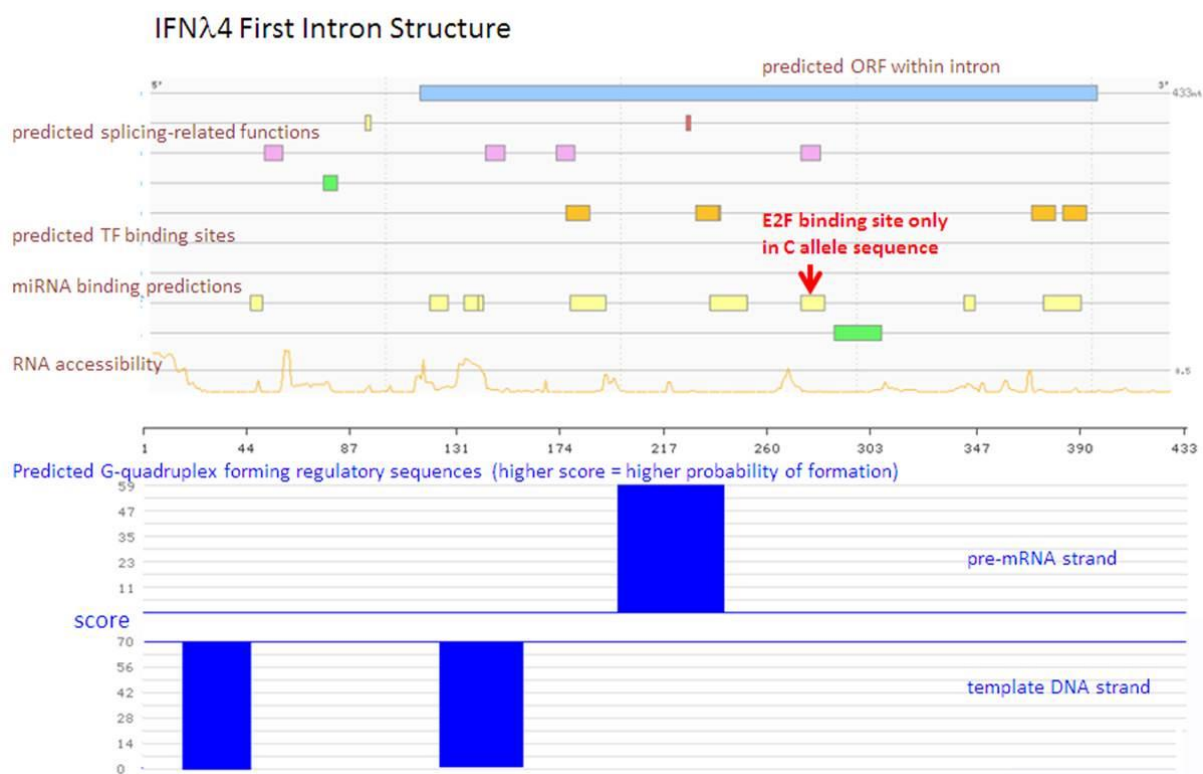
**Figure 6.12** Alignment of IFNL2 and IFNL3 amino acid sequences (A) and the corresponding model of IFNL2 protein generated using Raptor X (B) [229].

#### 6.2.6 Influence of rs12979860 on DNA and RNA structure

Besides encoding proteins, structural genes encode regulatory signals that switch their expression on and off, mainly as secondary and tertiary structures in noncoding regions. Other non-coding regions of the genome also contain switches that have longer-distance effects. Due to the outcomes influenced by the different SNP genotypes in HCV therapy, as well is in other settings (as described in Chapter 1), it is possible that rs12979860 C/T SNP acts to regulate *IFN $\lambda$*  gene expression, but the mechanisms by which it acts are not clear.

The rs12979860 SNP is located within what is now recognised as the first intron of the *IFN $\lambda$ 4* gene, ~ 19 kb and ~3 kb upstream of the transcription start sites of *IFN $\lambda$ 2* and *IFN $\lambda$ 3*, respectively (Figure 6.3 I).

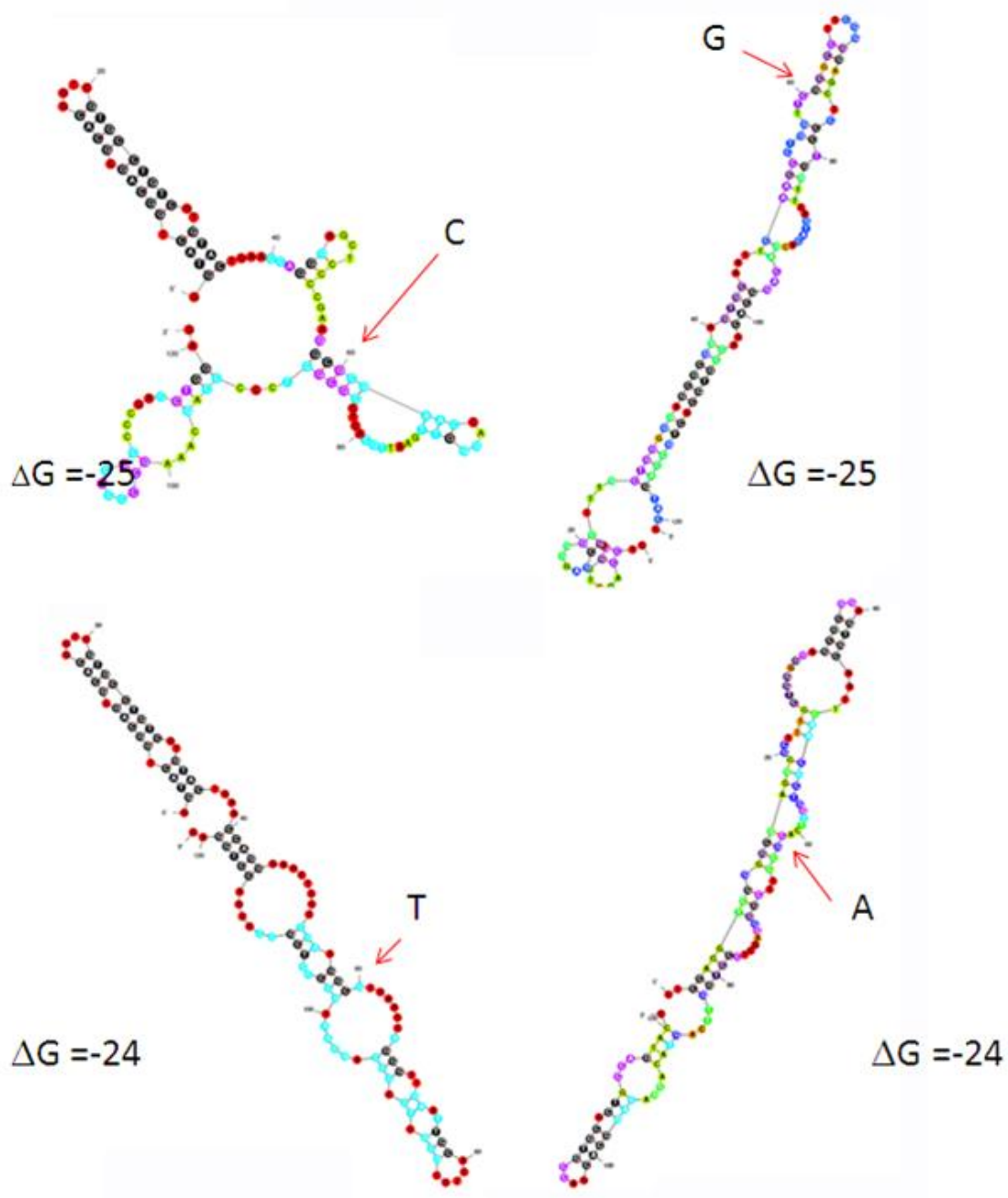
Searching the first intron of *IFN $\lambda$ 4* using RegRNA 2.0 (Figure 6.13) revealed the presence of a potential E2F binding transcription factor binding site in the C, but not the corresponding T allele sequence. The E2F family of transcription factors is evolutionarily ancient and has been highly conserved. E2F family members are ubiquitously expressed and can act as either transcriptional activators or repressors depending on the genetic context; an important role seems to be recruitment of other transcription factors to binding sites [230]. As the ancestral allele is thought to be the minor (T) allele, the T>C SNP might have been positively selected for because it allows greater genetic control over the expression of affected genes.



**Figure 6.13** Structure of the *IFN $\lambda$ 4* first intron showing locations of potential G-quadruplex forming sequences in both the DNA template strand and the pre-RNA transcript in relation to potential regulatory sequences. There is a potential E2F binding transcription factor binding site in the C, but not the corresponding T allele sequence.

In contrast to mRNA, single stranded DNA (ssDNA) is present at sites of active transcription, replication and recombination, where the normally paired strands must separate to provide access for the appropriate enzyme complexes. The secondary structure of the separated strands can provide location signals as well as affecting polymerase access and transit. Point mutations may alter the secondary structure of nucleic acids. Figure 6.14, generated using m-Fold [231], shows the predicted secondary structures of both pairs of single stranded DNA (C-G or T-A) encompassing 60 bases either side of the rs12979860 SNP. Although the SNPs alter the structures' conformations quite dramatically, they do not significantly (based on the free energy differences) affect their relative stabilities. The presence and stability of secondary structures in ssDNA can influence the kinetics of formation of other secondary structures, including G-quadruplexes, in their vicinity and vice-versa. G-quadruplexes are structures in which intra- or inter-strand interactions are stabilised by stacked, planar arrays of four guanines. They occur in G-rich regions, which are distributed non-randomly throughout the genome, notably in telomeres, promoters, and at the 5'-ends of first introns [232, 233], where they are involved in regulation of gene expression and genome stability. G-quadruplexes can also form in RNA. The presence of a G-quadruplex can either inhibit or promote gene expression, depending on its location. Formation of G-quadruplexes in the genomic non-template strand may make the complementary template strand more accessible to RNA polymerases, whilst G-quadruplexes in the template strand may need to be resolved by specific helicases to allow RNA polymerase transit [234]. Similarly, G-quadruplexes in the primary (pre-mRNA) transcript can promote splicing, but if retained in the processed mRNA, they may reduce the translation rate by causing ribosomal pausing [235].





**Figure 6.14: Effect of rs12979860 SNP on local DNA structure. Arrows denote altered nucleotides. Predicted secondary structures are shown for both pairs of single stranded DNA (C-G, or T-A).**

### 6.2.7 Summary and Conclusions of Part 1

1. Although their protein products are comparable in size, the *IFN $\lambda$*  genes are very small compared to the sizes of the genes for their two receptors IFN $\lambda$ R1 and IL10RB (Figure 6.8). During severe acute infection or extended chronic infection the ability to regenerate receptors may become limiting.



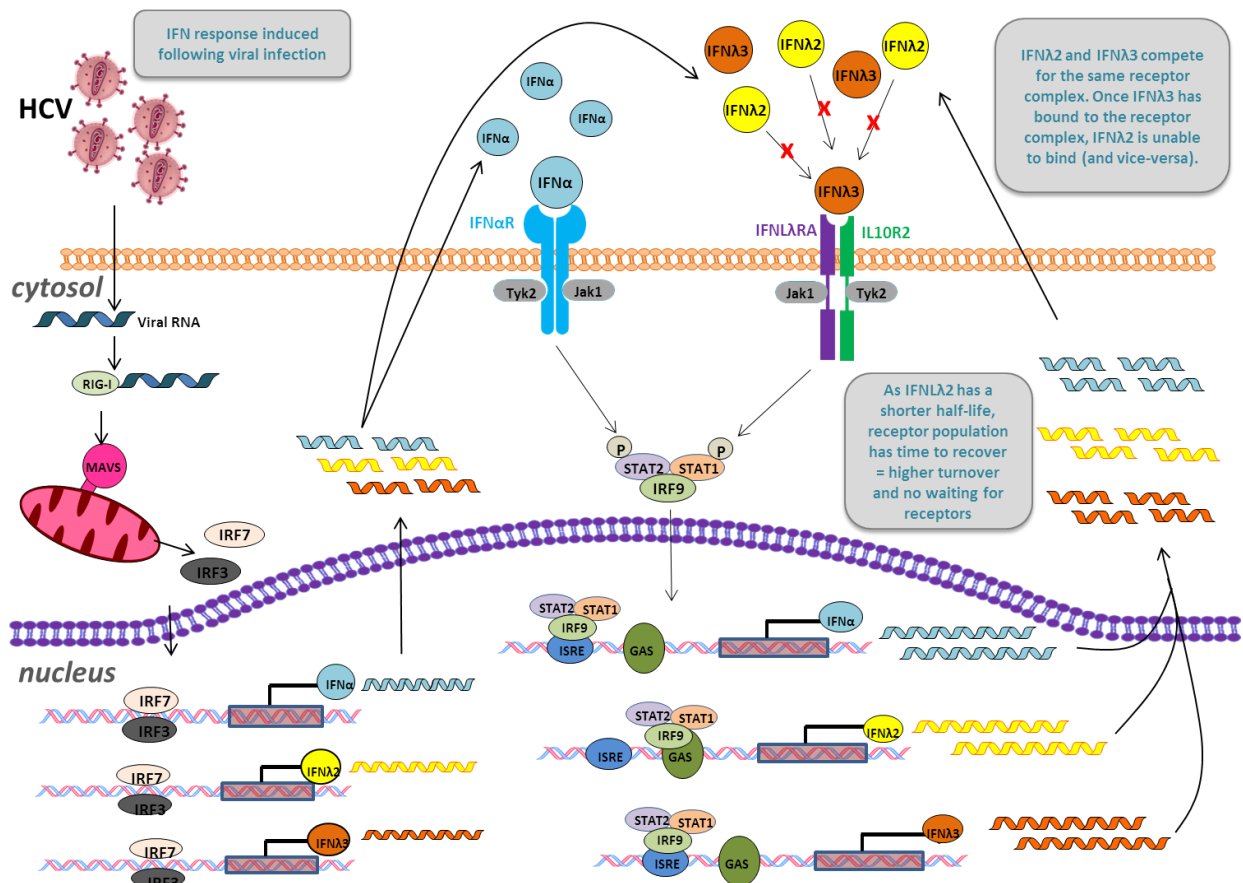
2. Despite close similarities, *IFN $\lambda$ s 1, 2 and 3* genes and their products show some significant differences which may account for observed difference in kinetics of expression and potency.
3. Each of the *IFN $\lambda$ s* signal via the same receptor pair (Figure 6.9), so if all are present, they must compete with each other (and with similar cytokines) for receptors. IL10RB is a ubiquitously expressed receptor subunit, but expression of *IFN $\lambda$ R1*, is limited to epithelial cells, hepatocytes, and dendritic cells [50]. Interestingly, it has been shown that, in chronic hepatitis C patients who are undergoing *IFN $\alpha$*  treatment, the expression of mRNA for *IFN $\lambda$ R1* is significantly greater in those who carry at least one minor (T) rs12979860 allele, which reduces response to treatment [236].
4. The rs12979860 C/T SNP, which has been shown to affect the *in vivo* and *in vitro* expression of *IFN $\lambda$ 3* [225] has theoretical potential to affect chromatin and pre-RNA structure, but exactly how these features are related to gene expression is still not clear and, deserves further investigation, especially in relation to the enigmatic *IFN $\lambda$ 4* gene and its products.

Why have multiple copies of functional *IFN $\lambda$*  genes been retained for millions of years? One likely possibility is that small but significant differences between the sequences and structures of each gene changes the way they are regulated. To date, the most studied SNPs that have been mapped to the *IFN $\lambda$*  locus and shown to affect the biological activities of its products have been located in non-coding regions. This is neither unusual nor unpredictable, since altering genetic control over a functional gene product is a safer option than risking (usually disadvantageous) changes in function that can result from even point mutations in coding sequences.

### 6.3: Part 2: Experimental

Conclusions from the bioinformatics study suggested that, because of their potential to alter both gene structure and to determine the presence or absence of a potential binding site for the transcription factor EF2, the SNPs at rs12979860 could alter the relative efficiency of transcription of the *IFN $\lambda$ 2* and *IFN $\lambda$ 3* genes (Figure 6.15). Although this idea could be tested using accurate quantitative assays for *IFN $\lambda$*  gene transcripts, interpretation of the results

may be difficult for a variety of reasons, including the known differences in half-lives of the mRNAs, with IFN $\lambda$ 2 mRNA having a shorter half-life than IFN $\lambda$ 3 mRNA [237]. Work described here used an alternative approach: assaying the plasma concentrations of IFN $\lambda$ 2 and IFN $\lambda$ 3 in parallel with genotyping in an attempt to discover whether a correlation between cytokine (i.e. secreted protein product) expression and rs12979860 genotype was detectable.



**Figure 6.15: Competition of IFN $\lambda$ 2 and IFN $\lambda$ 3 for receptor complexes.** IFN $\lambda$ 2 and IFN $\lambda$ 3 bind to the same receptor complex, under the same stimuli, and are thereby in competition with one another. In addition to the size differences of the cytokines and receptors mentioned earlier (Table 6.1), the half-life of the IFN $\lambda$ 2 mRNA is shorter than that of IFN $\lambda$ 3 mRNA, which will affect their time-dependent relative concentrations of their products. Consequently, the IFN $\lambda$ 3 response will last longer.

### 6.3.1 Methods

Commercially available IFN $\lambda$ 2 (IL28A; catalogue number OKAG00146) and IFN $\lambda$ 3 (IL28B; catalogue number OKAG00147) kits were sourced from Aviva Systems Biology (San Diego, USA). Each had a 96-well plate format, with 16 controls per plate. The remaining 80 wells were divided equally among IFN $\lambda$ 3 genotypes (27 CC, 26 CT, 27 TT patients). All samples

were sourced from those submitted for routine diagnostic IFN $\lambda$ 3 genotyping (see Chapter 2.2.1.3 for DNA extraction method, and 2.2.3 for IFN $\lambda$ 3 genotyping assay), and were tested using the ELISA kits according to manufacturer's protocols.

**6.3.1.1 IFN $\lambda$ 2 (IL28A) IFN $\lambda$ 3 (IL28B) ELISA procedure:**

Standards (8000 pg/ml – 125 pg/ml for the IFN $\lambda$ 2 assay; 2000 pg/ml – 31.25 pg/ml for the IFN $\lambda$ 3 assay) and negative controls were added to the first two rows of the ELISA plate. Patient samples were diluted 1:20 as per protocol, and were added to the remaining wells. The plate was then incubated at room temperature for two hours. Following this, biotin antibody, streptavidin solution, chromogenic substrate, and stop solution was added sequentially, with a series of washes between each step. The absorbance (OD) was measured at 450nm.

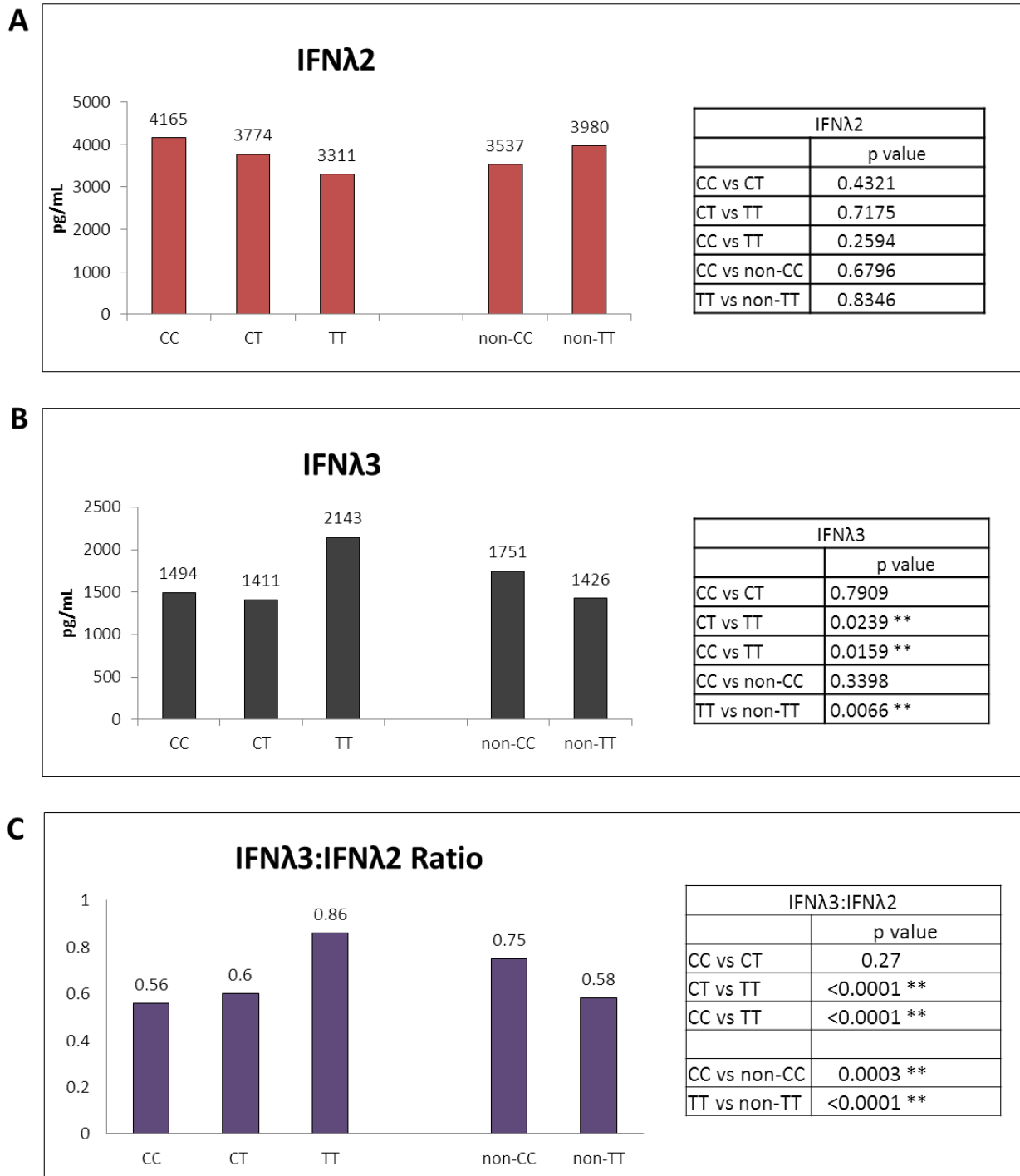
**6.3.2 Results**

Of the 80 samples subjected to the ELISA assay, usable data was obtained for 61 (75%) IFN $\lambda$ 2 and 58 (73%) IFN $\lambda$ 3. Furthermore, 51 (64%) were above the cut-off value for *both* IFN $\lambda$ 2 and IFN $\lambda$ 3 testing. The cytokine concentrations in remaining samples were below the limit of detection (Table 6.2). There were no females with the CT genotype and detectable IFN $\lambda$ 3, but for each of the remaining genotype combinations, the IFN $\lambda$ 3 concentrations were higher in females than in males. Due to the small sample size, further gender-based comparisons were not attempted.

CC	Total (n=27)	Male (n=23)	Female (n=4)
IFN $\lambda$ 2	21 (78 %)	19 (83 %)	2 (50 %)
IFN $\lambda$ 3	18 (67 %)	17 (74 %)	1 (25 %)
Both	17 (63 %)	16 (70 %)	1 (25 %)
CT	Total (n=26)	Male (n=20)	Female (n=6)
IFN $\lambda$ 2	19 (76 %)	15 (75 %)	4 (67 %)
IFN $\lambda$ 3	16 (62 %)	16 (80 %)	0 (0 %)
Both	14 (54 %)	14 (70 %)	0 (0 %)
TT	Total (n=27)	Male (n=15)	Female (n=12)
IFN $\lambda$ 2	20 (74 %)	11 (73 %)	9 (75 %)
IFN $\lambda$ 3	24 (89 %)	14 (93 %)	10 (83 %)
Both	20 (89 %)	11 (73 %)	9 (75 %)
Total	Total (n=80)	Male (n=58)	Female (n=22)
IFN $\lambda$ 2	61 (76 %)	46 (79 %)	15 (68 %)
IFN $\lambda$ 3	58 (73 %)	47 (81 %)	11 (50 %)
Both	51 (64 %)	41 (88 %)	10 (45 %)

**Table 6.2: Usable results from the two ELISA assays. The table shows the total number of samples with quantifiable cytokine concentrations for each genotype, stratified by gender.**

From the results of the IFN $\lambda$ 2 assay, a clear trend was evident, with the higher concentrations being associated with the C allele (CC>CT>TT), however the differences were not statistically significant (Figure 6.16A). However, there was a significant difference in plasma IFN $\lambda$ 3 concentrations between TT and all other (non-TT) groups (Figure 6.16B). In addition, the TT genotype showed a significantly higher IFN $\lambda$ 3:IFN $\lambda$ 2 ratio (Figure 6.16C).



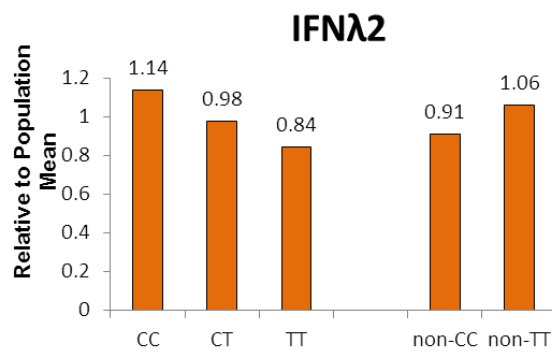
**Figure 6.16: IFN $\lambda$ 2 (A) and IFN $\lambda$ 3 (B) concentrations according to IFN $\lambda$ 3 rs12979860 genotype, and the ratio between them (C).**

### 6.3.2.1 Re-standardisation and Analysis

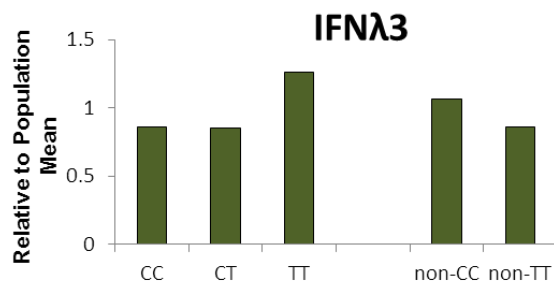
Although a genotype-dependent trend was evident, a lot of test sample ODs were close to or below the recommended kit calibration range. To overcome this problem, the results

were reanalyzed after standardization to make them comparable. This was achieved as follows (1) the minimum OD reading was subtracted from all test samples (2) the distribution of the adjusted OD values was found to be approximately normal, with the median approximately equal to the mean, with few outliers; (3) the mean value was calculated; (4) the adjusted OD values were standardized by expressing each result as a fraction of the population mean value, so that the result for each sample was converted into a dimensionless ratio. The ratio provides a relative, rather than absolute, concentration measurement.

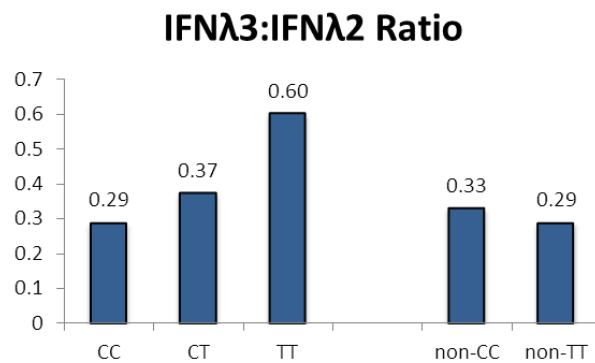
Analysis of the re-standardised data produced results similar to the initial analysis, with CC having the highest IFN $\lambda$ 2 concentration, and TT having the highest IFN $\lambda$ 3 concentration (Figure 6.17). There was no significance in the trend observed in IFN $\lambda$ 2 concentration, and the same groups showed significance differences in IFN $\lambda$ 3 concentrations. Although the significance values are slightly less after re-analysis, confidence is increased due to the greater sample size. The IFN $\lambda$ 3 to IFN $\lambda$ 2 ratios follow the same clear trend (Figure 6.17 C).

**A**

IFN $\lambda$ 2	
	p value
CC vs CT	0.3847
CT vs TT	0.4663
CC vs TT	0.1319
CC vs non-CC	0.1571
TT vs non-TT	0.1784

**B**

IFN $\lambda$ 3	
	p value
CC vs CT	0.9538
CT vs TT	0.0309 **
CC vs TT	0.0237 **
CC vs non-CC	0.1749
TT vs non-TT	0.0096 **

**C**

IFN $\lambda$ 3:IFN $\lambda$ 2 Ratio	
	p value
CC vs CT	0.3234
CT vs TT	0.0148 *
CC vs TT	<0.0001 **
CC vs non-CC	0.5108
TT vs non-TT	<0.0001 **

**Figure 6.17: IFN $\lambda$ 2 (A) and IFN $\lambda$ 3 (B) ELISA results re-standardised and compared according to IFN $\lambda$ 3 rs12979860 genotype, including the ratio between IFN $\lambda$ 2 and IFN $\lambda$ 3 (C).**

### 6.3.3 Discussion

The observed rs12979860 genotype-dependent differences in IFN $\lambda$ 2 and IFN $\lambda$ 3 production and the ratios of their plasma concentrations provides support for the hypothesis based on the bioinformatics part of this study, namely, that the C/T SNP rs12979860 could direct transcriptional traffic by affecting chromosome structure. It could be predicted that the difference would be difficult to detect because it would be time dependent and dynamic, depending on the presence of chronic or acute infection or inflammation and its duration, as well as drug treatment, and the individual's genetic background, and a multitude of other influential factors.

There are few publications that report serum or plasma concentrations of IFN $\lambda$ s, probably because they are usually close to the limits of detection, even in chronically-infected HCV patients [238]. In this study, 64% of samples gave a positive result in both the IFN $\lambda$ 2 and IFN $\lambda$ 3 assays. Past studies [43, 44, 238] have reported higher 'IL28' expression in favourable compared to non-favourable IFN $\lambda$ 3 genotypes, both by RT-PCR (for mRNA/DNA) and ELISA (protein). The clear difference in these studies is the use of a combined 'IL28' assay, rather than separate specific assays. One study found that IL29 is easier to detect by ELISA, and in order to detect concentrations of IFN $\lambda$ 2 and IFN $\lambda$ 3, a combined assay was performed, which found higher 'IL28' (ie, a combined IFN $\lambda$ 2 and IFN $\lambda$ 3 assay) serum concentration in homozygous IFN $\lambda$ 3 CC are higher than TT [238]. Both Suppiah *et al.* [43] and Tanaka *et al.* [44] measured IFN $\lambda$ 2 and IFN $\lambda$ 3 mRNA concentrations by RT-PCR, using a primer pair that is conserved in both genes, and found lower expression in non-favourable genotypes. As it was done in one combined assay, they did not distinguish between the two genes. A subsequent study of a Chinese population using a specific IFN $\lambda$ 3 assay showed higher serum concentrations in favourable IFN $\lambda$ 3 genotypes, which comprised 94% of the cohort [112].

When the ELISA results from this current study are pooled, the results are in agreement with these previous publications. However, additional information has been generated by the use of the two specific IFN $\lambda$ 2 and IFN $\lambda$ 3 assays. Although the sequences are almost homologous, the manufacturer claims the antibodies show specificity, and do not cross-



react with each other. The added value of using two specific assays showed that the presence of the T allele increases the plasma IFN $\lambda$ 3:IFN $\lambda$ 2 ratio (Figures 6.16 and 6.17) .

A trend between IFN $\lambda$ 2 serum concentrations was observed when comparing each rs12979860 genotype, with the addition of each beneficial C allele increasing the level. The opposite trend was observed in IFN $\lambda$ 3 serum concentrations, but only in the homozygote; the level in the heterozygote was the same as that of the CC homozygote. Similarly, a previous study found that rs8099917 (but not rs12979860) had a significant correlation with IFN $\lambda$ 3 serum concentrations in chronically infected HCV patients [239]. A difference in their study is with the heterozygote form exhibiting an intermediate serum level, whereas this study found the serum concentrations of IFN $\lambda$ 3 in the heterozygote to be aligned with the beneficial homozygote form. This is of marked interest, as the majority of studies into IFN $\lambda$ 3 to date compare CC to non-CC genotypes, whereas in this instance the difference lies in comparing TT to non-TT.

Concentrations of IFNs alter during treatment regimens. Rallon *et al* [240] showed using a commercial kit specific for IFN $\lambda$ 3 in HIV/HCV co-infected patients. Baseline IFN $\lambda$ 3 concentrations did not differ according to IFN $\lambda$ 3 genotype, but concentrations in CC patients increased more by week 4 of treatment than non-CC. They suggested that induction of IFN $\lambda$ 3 was occurring as a result of IFN $\alpha$  treatment in IFN $\lambda$ 3 CC patients [240]. It would be of interest to monitor concentrations of both IFN $\lambda$ 2 and IFN $\lambda$ 3 as patients progress through different stages of infection, treatment, and viral clearance. Although this was beyond the scope of the present project, it may help clarify the effect of exogenous IFN $\alpha$  therapy on endogenous IFN concentrations, and provides direction for future studies and collaborations. It would also be of considerable interest to include IFN $\lambda$ 4 expression in any further investigations; expression of this gene has been accepted as the causative to the difference seen in many IFN $\lambda$ 3 SNPs since its identification. However, at this stage, there are no commercial assays specifically for detecting IFN $\lambda$ 4 concentrations.

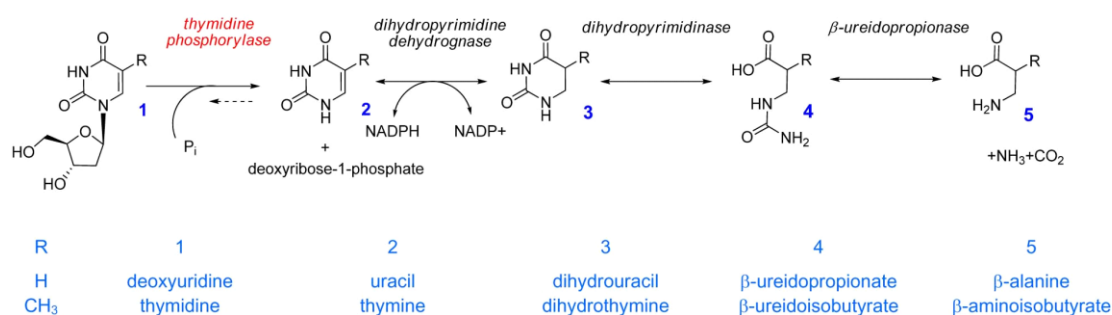
## 6.4 Conclusion

Previous correlation studies have grouped genotype comparisons as either 'CC' or 'non-CC', or as applicable in different SNPs. As such, the heterozygote form has continually been grouped in with the minor allele. Of interest, the major difference identified in this Chapter is within the TT genotype, and the concentration of the CT was the same as the CC in both IFN $\lambda$ 2 and IFN $\lambda$ 3. This should not be particularly surprising given that rs368234815 denotes expression ( $\Delta$ G) or repression (TT) of IFN $\lambda$ 4 is widely accepted as being the causative SNP responsible for the effect seen by rs12979860 genotype [55]; if the  $\Delta$ G at rs368234815 is only present in individuals with a TT at rs12979860, it is reasonable to compare the rs12979860 TT group against the non-TT group as it is these individuals who express IFN $\lambda$ 4. It is presently unclear why expression of IFN $\lambda$ 4 has been recently selected against, whilst IFN $\lambda$ 1, IFN $\lambda$ 2, and IFN $\lambda$ 3 have survived as functional genes. By utilizing information publically available, small but significant differences between the structural genes and their products have been identified; these differences in gene and transcript structure give clues as to the variance in turnover rates and function within the type III IFN family. This theoretical argument has been backed up by the experimental data showing different plasma concentrations of IFN $\lambda$ 2 and IFN $\lambda$ 3 according to rs12979860 genotype. Further investigations will seek confirmation of these findings, and may seek to explore alternative technologies.

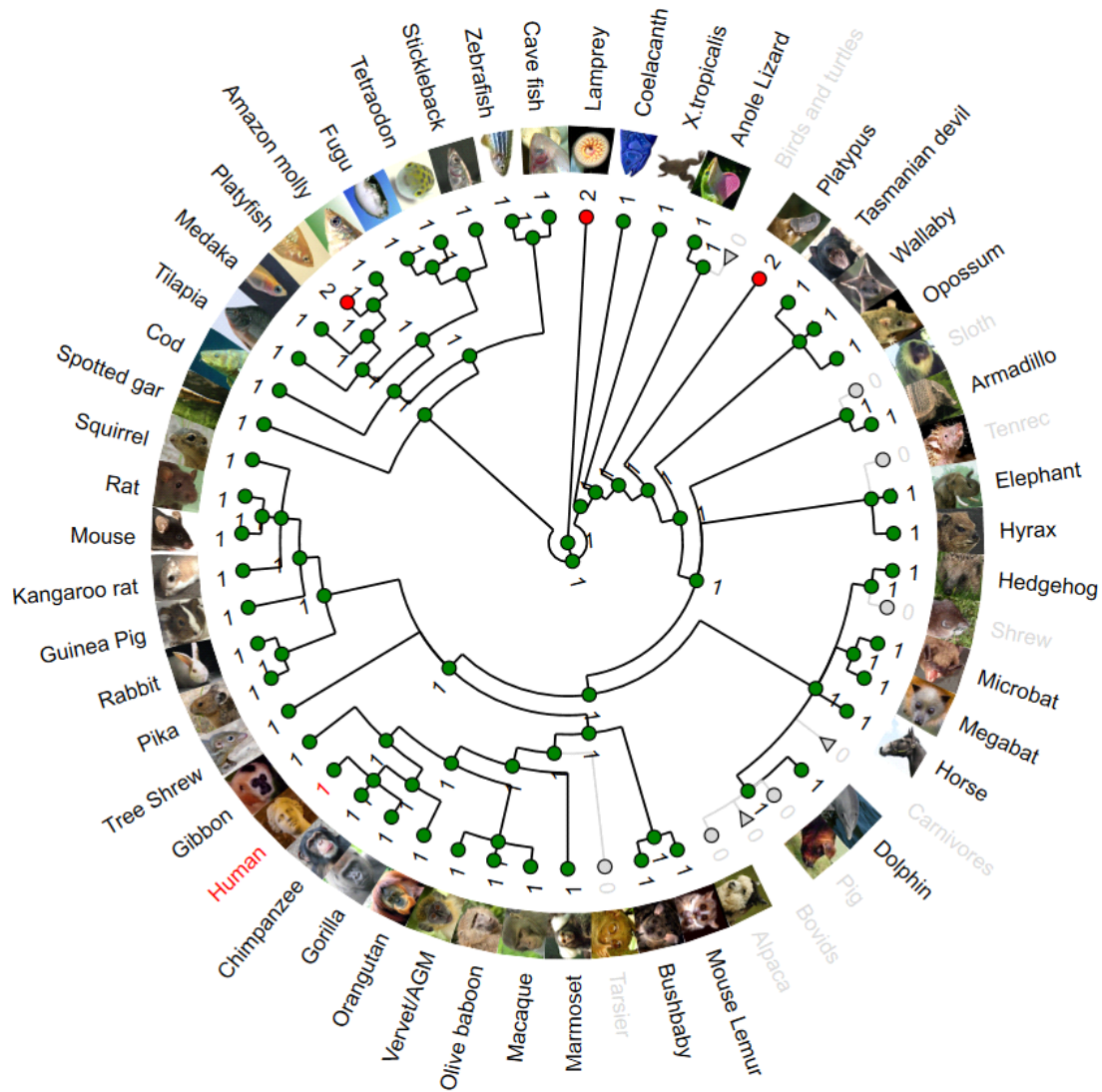
## Chapter 7: Association Between Thymidine Phosphorylase, Inflammation and Interferon $\lambda$

### 7.1 Introduction

Thymidine Phosphorylase (TP; EC 2.4.2.4) is an enzyme (482 amino acids; ~155 KDa) encoded by the *TYMP* gene, which maps to band q33 on the long arm of chromosome 22 [241]. TP initiates the catabolism of the pyrimidine deoxynucleosides thymidine (TdR), deoxyuridine (UdR) and several of their pharmacologically important analogues, which are used clinically as antiviral and anti-neoplastic agents [242] (Figure 7.1). The importance of TP is highlighted by its remarkable evolutionary conservation. With few exceptions, most vertebrate taxa encode a TP homologue (Figure 7.2). Human TP shares close sequence homology with TP of the most primitive vertebrates and even prokaryotes. For example, *Latimeria chalumnae* (the coelacanth) and *Escherichia coli* TP share 54% and 39% sequence identity, respectively, with human TP [243] (Figure 7.3).

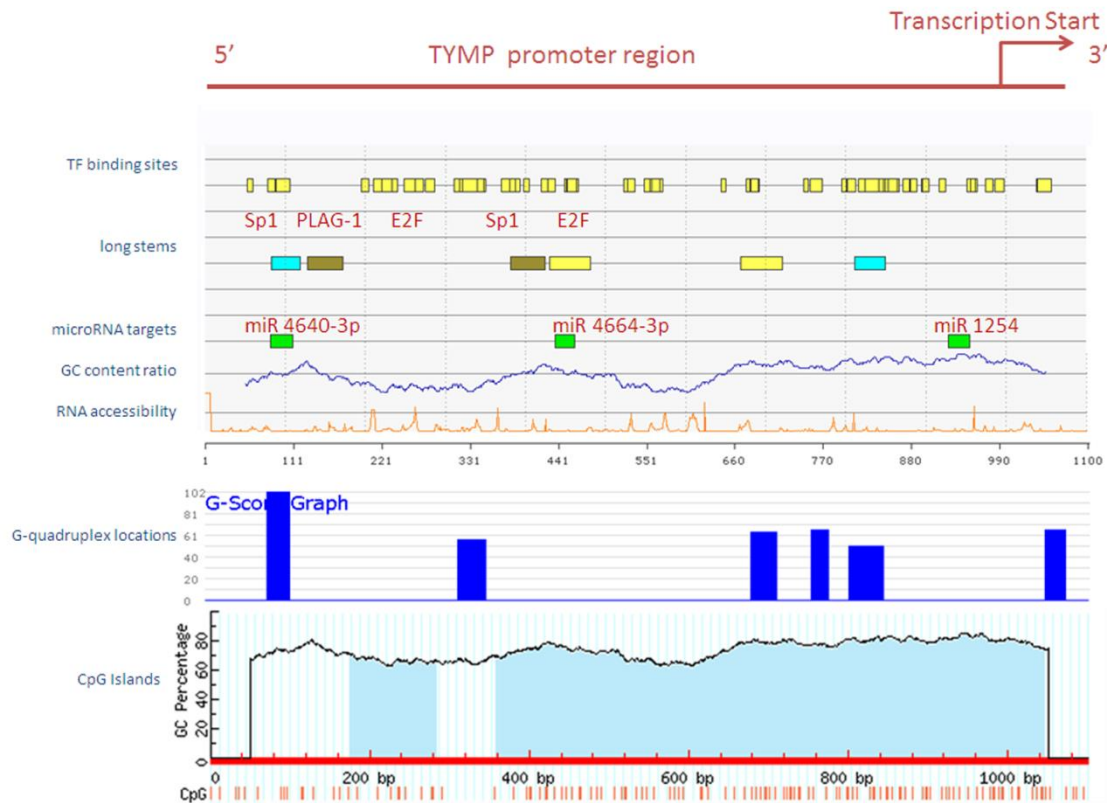


**Figure 7.1** TP catalyses the initial step in catabolism of the natural deoxynucleosides TdR and UdR as well as many of their pharmacologically important analogues. Although the reaction is theoretically reversible, steady state is rarely achieved in eukaryotes because of rapid consumption of deoxyribose-1-phosphate. In humans, the rate-limiting step in degradation of TdR and UdR is catalysed by dihydropyrimidine dehydrogenase, activity of which is more limited in some ethnic groups, particularly those with African ancestry.



**Figure 7.2:** Nearly all vertebrate genomes encode a single *TYMP* gene homologue. Intriguingly, the three taxa that encode two copies are located at major evolutionary nodes or "crisis" points. Birds seem to lack a TP homologue. (Figure imported from the Ensembl GeneBrowser; <http://asia.ensembl.org/index.html>.)

127



**Figure 7.4: The GC-rich TP proximal promoter region contains multiple binding sites for Sp1 as well as binding sites for AP-1, NFkB, ISRE and GAS sequences. These are all features associated with redox-sensitive regulation.**

### 7.1.1 Regulation of TYMP expression

TP activity generates reactive oxygen species (ROS) [244]. The *TYMP* proximal promoter region contains functional binding sites for multiple redox-sensitive transcription factors, including Sp1, NFkB, and AP-1 (Figure 7.4), which presents opportunities for auto-regulation. Sp1, NFkB, and AP-1 are involved in gene expression regulating survival processes including differentiation, apoptosis and DNA repair [245]. Sp1 can act as an activator or a repressor due to its capacity to regulate methylation [246]. It binds to CpG rich motifs, which are present in abundance immediately upstream of the *TYMP* transcription start site. The *TYMP* promoter also contains a functional interferon gamma-activated sequence (GAS) as well as Interferon Stimulated Response Elements (ISRE) [247]. Collectively, these characteristics provide evidence that TP is involved in, and is probably essential for rapid stress responses, which include the innate immune response.

*TYMP* is expressed ubiquitously, since it is essential for proper regulation of the dNTP supply [248]. There are eight isoforms of TP, and different transcripts are produced in different cell types. Expression levels also vary, controlled at the transcriptional level by methylation of cytosine residues in the promoter, which is commonly aberrant in tumours. For example, hyper-methylation has been reported in DLD-1 colon carcinoma cells, whilst, at the other extreme, complete demethylation occurs in breast carcinoma SKBR-3 cells [243]. *TYMP* expression is also be regulated at both transcriptional and translational levels by G-quadruplexes, which are thermodynamically stable structures that can form in both DNA and RNA and usually function to inhibit gene expression by blocking the transit of polymerases or ribosomes (see previous chapter) [249]. In addition, TP production is controlled post-transcriptionally by AU-rich element mediated mRNA decay [250]. These regulatory mechanisms are consistent with *TYMP* expression being either constitutive or inducible depending on tissue and cell type as well as on both developmental stage and the immediate local microenvironment.

### **7.1.2 Physiological role of TP: Regulation of Mitochondrial Replication**

The main physiological function of TP is to indirectly regulate the pyrimidine dNTP supply, by controlling the availability of TTP, which is critical because TTP is the ultimate intracellular regulator of other dNTP pools. Any type of dNTP pool imbalance is mutagenic, pro-carcinogenic and potentially lethal if sustained and, because it is relatively rapid and continuous, mitochondrial genome replication is very sensitive to changes in the dNTP supply and suffers the immediate consequences [251]. TTP excess can inhibit or prevent mitochondrial replication by causing dCTP starvation: mitochondrial deoxypyrimidine (“thymidine”) kinase (TK2) which salvages both deoxycytidine and TdR, is blocked by negative feedback from TTP [252]. TTP also inhibits the activity of ribonucleotide reductase, the rate-limiting enzyme for supply of the other dNTPs [253, 254] .

The long-term outcome of sustained inhibition of mitochondrial replication is a reduction in the number of mitochondrial genome copies, eventual depletion of the mitochondrial population and, since mitochondrial respiration supplies most of cells’ energy requirement in nearly all tissues, loss of bioenergetics capacity. Since they have the greatest energy requirements, neuromuscular tissues are most sensitive to mitochondrial depletion. The



consequences include strokes, seizures, diabetes as well as many others. The first reported mitochondrial depletion syndrome was mitochondrial neurogastrointestinal encephalomyopathy (MNGIE), an autosomal recessive disease which is caused by loss-of-function mutations affecting the *TYMP* gene, resulting in decreased capacity for TdR catabolism which causes TTP and dUTP overload. Heterozygotes have partially reduced enzyme function, which is sufficient in most cases to control TdR supply, implying that normal individuals have a large reserve capacity for TP activity [255].

Both TdR overload and TdR starvation are potentially lethal. TdR starvation induces a phenomenon known as thymineless death, a highly evolutionarily conserved programmed cell death pathway that occurs in almost all prokaryotes and eukaryotes [256]. Prolonged restriction of the TTP supply results in its replacement by dUTP, which initiates futile, energy-costly cycles of attempted DNA repair, higher mutation frequency and ultimately cell death [257]. TdR starvation causes the destruction of replication origins and complete breakdown of mitochondrial genomes, whereas the nuclear DNA of cells that have suffered prolonged TTP starvation contain multiple single-stranded breaks and gaps, which may be converted by futile attempts at repair or replication into lethal double-stranded breaks [258].

### 7.1.3 Induction of TP Activity by Interferons and Other Pro-Inflammatory Stimuli

Administration of interferon (IFN)- $\alpha$  has been showing to increase TP activity 2-3 fold in peripheral blood cells [259], as well as in tumour tissues in colorectal [260] and gastrointestinal cancer patients [261]. The elevated TP activity has been shown to occur within 1-2 hours of IFN- $\alpha$  administration [262]. In blood, the increase in TP following administration of a single dose of IFN- $\alpha$  is dose-dependent and may be sustained for up to thirteen days [263]. Type II interferon (IFN- $\gamma$ ), which is inducible by stress and other IFNs, can also upregulate *TYMP* expression [247]. As the standard treatment for chronic hepatitis C is pegylated IFN- $\alpha$ /Ribavirin (pegIFN- $\alpha$ /RBV), these are points of considerable interest. TP activity could be expected to increase during pegIFN- $\alpha$ /RBV treatment, but only if there was some unused reserve capacity for expression remaining after induction by an individual's endogenous immune response. It is therefore possible that baseline TP activity could be prognostic. Since all IFNs signal *via* the JAK-STAT pathway, it is not unreasonable to propose



that IFN $\lambda$ s, like IFNs- $\alpha$  and - $\gamma$ , induce TP expression. Results from previous reports and work described in the previous Chapter 6 implicates the rs12979860 genotype as an important mediator of immune and inflammatory responses, suggesting that it may affect *TYMP* expression at one or more levels (transcriptional, translational, or post-translational). The aim of work described in this chapter was to begin to investigate this possibility.

#### **7.1.4 Hypothesis and Aim**

IFNs are known to up-regulate TP activity. The alternate alleles at IFN $\lambda$ 3 rs12979860 appear to specify different levels of control over IFN $\lambda$  activity, which may measurably affect the expression of downstream targets, with *TYMP* being a strong candidate. TP activity, which is easily measured in whole blood [264], may correlate with IFN $\lambda$ 3 activity and rs12979860 genotype. Accordingly, the aim was to measure TP activity in parallel with IFN $\lambda$ 3 genotyping and to search for possible correlations.

## **7.2 Materials and Methods**

### **7.2.1 Blood Samples**

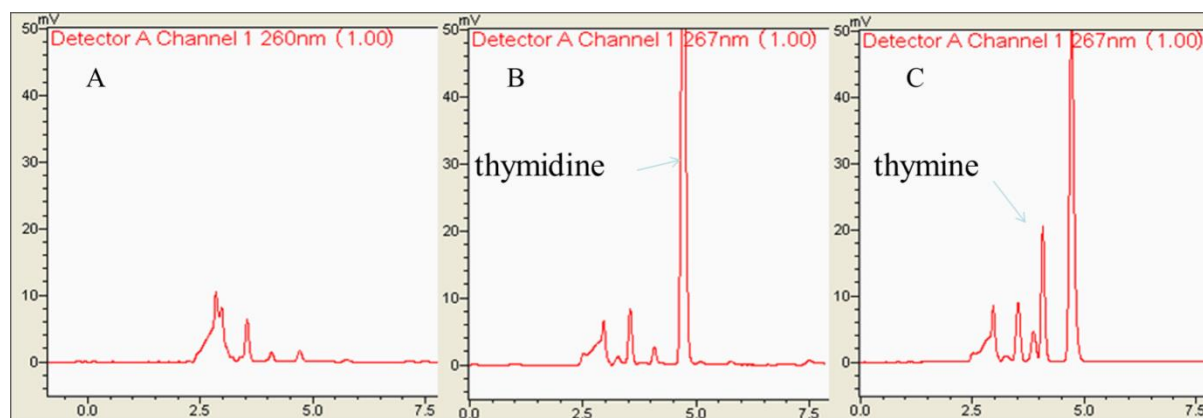
A total of 300 consecutive EDTA anti-coagulated blood samples submitted to VIDRL for routine IFN $\lambda$ 3 genotyping were assayed in parallel for TP activity. IFN $\lambda$ 3 genotypes were determined as previously described (Chapter 2.2.3) using cell-free DNA extracted from plasma with the Qiagen DNA MiniKit (Chapter 2.2.1.3).

### **7.2.2 Assay for Thymine Phosphorylase Activity**

A simple and robust HPLC-based assay was devised to measure TP activity in whole blood samples, based on a previously published method [264]. The protocol was almost identical except for two modifications which increased the assays' sensitivity and robustness: (1) the whole blood fraction in the assay mixture was halved and the TdR substrate was dissolved in sterile PBS so that the final TdR concentration in the incubation mixture was 450 $\mu$ M; (2) at the end of the incubation, the enzymatic reaction was terminated and plasma proteins

precipitated by adding an equal volume of 7% w:v perchloric acid instead of three volumes of 100% methanol.

Briefly, the assay mixture consisted of whole blood mixed with an equal volume of PBS with or without TdR substrate. For routine assays the total volume was 200 $\mu$ L, incubation temperature was 37°C, final TdR concentration was 450 $\mu$ M and the incubation time was 60 minutes (Figure 7.5). Under these conditions, dihydrothymine formation was negligible. At the end of the incubation period, blood cells were pelleted by centrifugation and perchloric acid was added to small aliquots of the supernatant to precipitate protein, which was removed by brief high speed centrifugation. Small aliquots (routinely 5 $\mu$ L) of the final supernatant were then subjected reverse-phase HPLC on a 4.6 x 250mm C<sub>18</sub> ODS reversed phase HPLC column (particle size 5 $\mu$ M) which was protected by a matching guard column (both supplied by Grace Davidson Discovery Systems, Baulkham Hills, NSW). The column temperature was 40°C. Columns were installed in a Shimadzu LCMS-2020 System, which included an integrated autosampler-injector system and UV detector. The reaction was approximately linear for at least an hour, with inter- and intra- assay variation consistently < 10%, comparable to the original method [264].

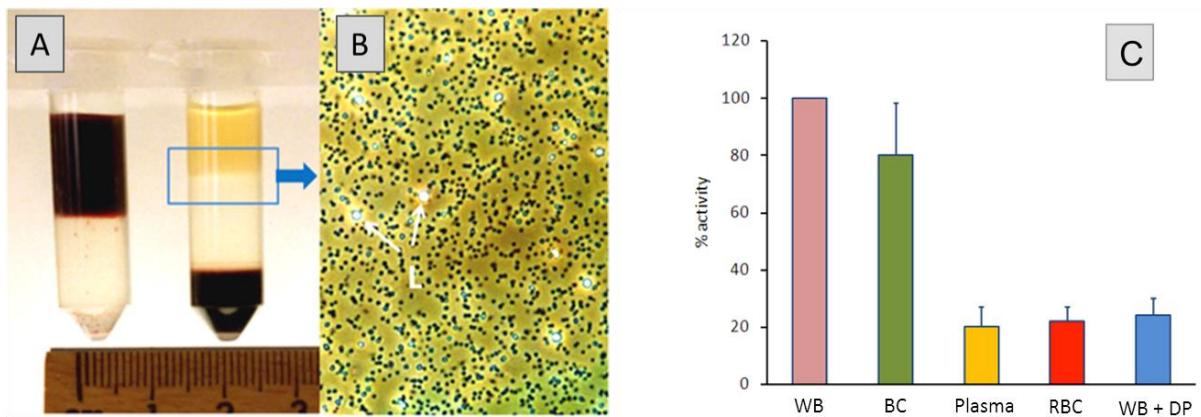


**Figure 7.5: HPLC Assay for TP Activity. (A), control (no substrate); (B) start of assay, no incubation; (C); end of assay, 60 min incubation.**

### 7.3 Results

TP activity is mainly intracellular and most is contained in platelets and small lymphocytes (Figure 7.6). Whole blood (WB) samples (with and without 4 $\mu$ M dipyridamole (DP), a

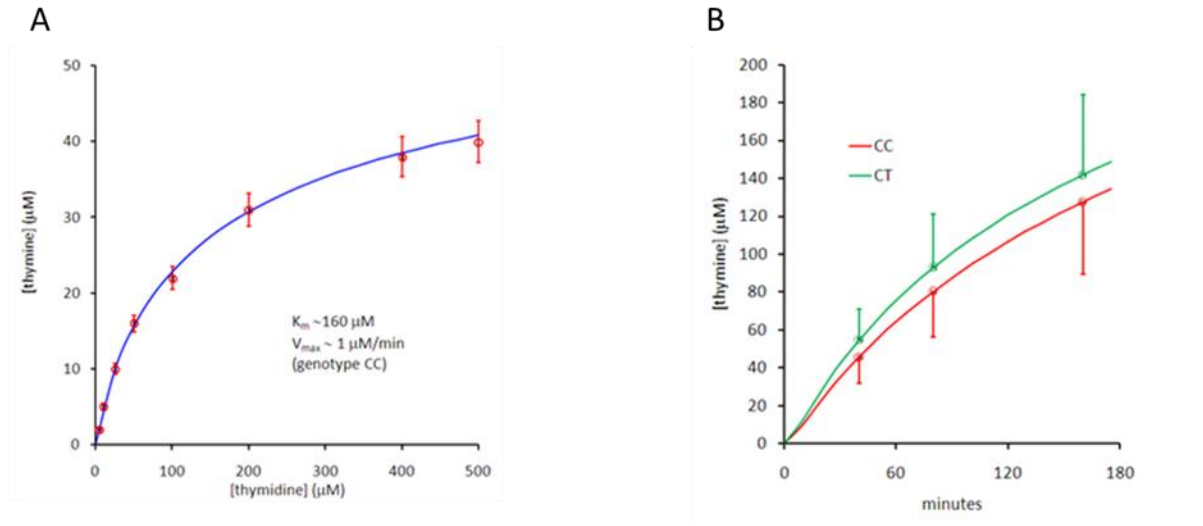
nucleoside transport blocker) and corresponding fractions (BC, plasma, RBC: red blood cells) were assayed in parallel for TP activity, which is expressed as a percentage of the WB total. Most TP activity was in the BC fraction (Figure 7.6 C).



**Figure 7.6 Measurement of TP activity.** The cell (“buffy coat”; BC) fraction of density  $<1.077\text{g/ml}$ , isolated using Percoll mini-gradients [264] (A), consisted mainly of platelets and small lymphocytes (arrows in phase contrast microscope image, (B)). Blood samples (WB=whole blood; BC=buffy coat; RBC=red blood cells; WB+DP=whole blood +  $4\mu\text{M}$  dipyridamole) were assayed in parallel for TP (C).

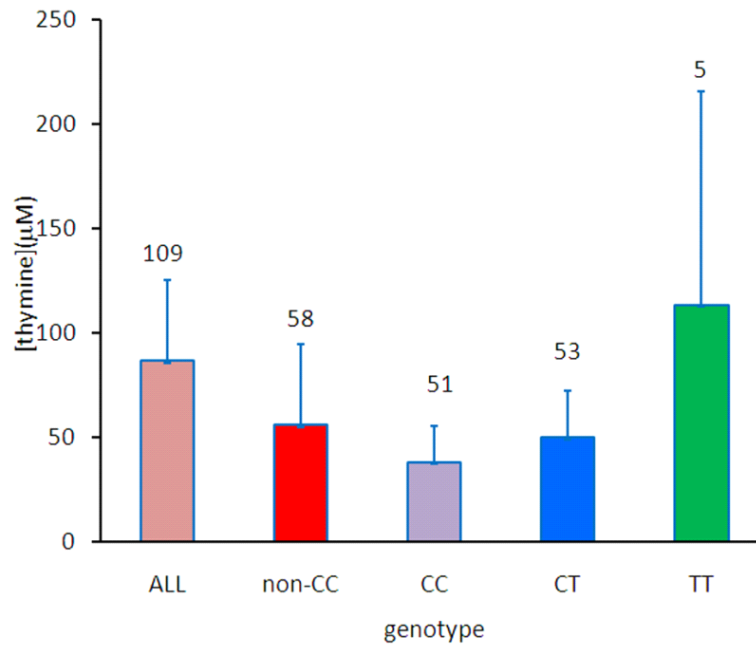
In blood, TP is mainly intracellular, so to minimise artefacts due to cell age-related lysis, results from blood samples that were more than three days old or were known to have abnormal blood cell counts, as well as others that showed visible evidence of cell lysis were not included in the analysis. These criteria eliminated thirty three assay results, leaving a total of 267, of which, unfortunately, only five were from TT homozygotes.

The TP assay was performed for different durations and with different substrate concentrations (Figure 7.7). Non-linear regression analysis gave apparent  $K_m$  of  $\sim 160\mu\text{M}$  and apparent  $V_{\text{max}}$  (expressed as the rate of increase in thymine concentration after 60 minutes) of  $\sim 1\mu\text{M}$  per min, corresponding to  $\sim 1\text{nM}$  TdR catabolised per minute/ml of blood. A greater than 7-fold range in both apparent  $K_m$  and  $V_{\text{max}}$  was observed. Although there was considerable overlap, the initial rate of TdR catabolism at saturation was generally greater in samples with the CT genotype than those with the CC genotype; 5 samples per genotype) (Figure 7.7 B). Sufficient numbers of genotype TT samples were not available to allow direct comparisons.



**Figure 7.7: Kinetics of TP activity in blood. (A)** Illustrates a typical result in which data points represent means and SDs from a single blood sample assayed in triplicate. **(B)** Illustrates TdR catabolism at saturation was generally greater in samples with the CT genotype than those with the CC genotype.

Despite the wide range in WB TP activity and substantial overlap in TP activity between genotypes (Figure 7.8), statistically significant ( $p < 0.05$  by student t test) genotype-dependent differences were observed. The relatively small overall sample size and in particular the lack of samples with “TT” genotype necessitates confirmatory testing using a larger population before definitive statements can be made regarding the relationship between IFN $\lambda$ 3 and TP expression.



**Figure 7.8: Relationship between TP activity and IFN $\lambda$ 3 genotype.**

#### 7.4 Discussion

A simple HPLC-based assay was devised and used to measure TP activity in whole blood. Assays results revealed an extremely wide range of TP activity, consistent with previous reports [242, 243, 264]. Correlation of TP assay results with results of IFN $\lambda$ 3 genotyping revealed a trend in TP activity in the order TT>CT>CC, supporting the hypothesis that the rs12979860 T allele is associated with increased pro-inflammatory activity. This postulate is supported by the observations that platelets are the main source of TP in blood [243] and that patients who are IFN $\lambda$ 3 TT homozygotes have a higher platelet counts than heterozygotes and CC homozygotes [265]. Additionally, sterile inflammatory conditions such as allergic reactions and asthma occur more frequently in IFN $\lambda$ 3 TT patients [118], and are also associated with higher TP levels [266]. The study was limited by the low number of IFN $\lambda$ 3 TT samples, a consequence of including only those samples which were no more than three days old. These results require confirmation by large-scale testing of fresh blood that has been collected and stored under standardized conditions, which would allow the additional benefit of gender analysis given larger numbers.

As Fukushima *et al* [267] observed, an increase in TP activity is dependent on the initial expression level; cells with a lower TP expression had a greater capacity to increase activity when compared to those with higher initial TP activity. Thus, if there is not the capacity for TP production to be further stimulated (e.g. in those patients with the T allele), then a drug such as IFN- $\alpha$  will lose some of its efficacy. This, along with the increase in TP activity following administration of IFN- $\alpha$  [259] is consistent with the observation that non-CC IFN $\lambda$ 3 patients do not respond to pegIFN- $\alpha$ /RBV as well as CC patients. IFN $\lambda$ 3 TT patients lack the ability to further up-regulate their TP expression, presumably due to promoter demethylation. This is consistent with that of other ISGs [61], further strengthening the link between IFN $\lambda$ 3 genotype and TP expression.

This study also provides some insight into the seemingly paradoxical observation that patients who carry the beneficial IFN $\lambda$ 3 C allele have higher baseline viral loads. Labarga *et al.* suggested that the C allele of rs12979860 may reflect lower immune activity [268], somewhat in agreement with what has been observed here (Figure 7.8). Although apparently disadvantaged with a higher level of viraemia, IFN $\lambda$ 3 CC patients are able to respond more efficiently to exogenous IFN- $\alpha$  therapy compared to those harbouring a non-CC genotype.

Revealed here is a correlation between whole blood TP activity and IFN $\lambda$ 3 genotype. The T allele at rs12979860, which is associated with greater baseline levels of IFN $\lambda$ 3 and poor response to HCV treatment, is also associated with increased TP activity. This implies that the presence of the T allele stimulates a strong initial immune response that reduces the dynamic range available for further stimulation by antiviral treatment. It follows that both IFN $\lambda$ 3 genotype and TP activity are potentially useful surrogate markers of immune activation and pro-inflammatory activity in general, not confined to HCV infection alone.

## Chapter 8: Final Discussion

Research into the associations of IFN $\lambda$ 3 and IFN $\lambda$ 4 has been vast in recent years, with the number of publications in excess of 1,650. Many of these articles have been reporting the correlation between various SNPs and outcome of HCV clearance, concentrating on different ethnicities, and a few delved into other viruses such as hepatitis B virus (HBV) and HIV. A small proportion of the studies have identified functional or biochemical differences, and novel SNPs. With the recent implementation of new treatment regimens for chronic HCV infection, the relevance of such correlations has been questioned. These correlations still exist, however they are lessened; research into this field is ongoing, and may prove to be beneficial to other viruses or immune responses beyond HCV clearance.

This work described in the thesis has answered a number of questions, both in terms of practical diagnostic testing, and in further investigations as to the reasons behind the observations.

The initial validation of the IFN $\lambda$ 3 assay involved assessing specimen types, and it was clear that plasma and serum samples provided sufficient genomic DNA to detect IFN $\lambda$ 3 genotypes. This was highly beneficial, as many patient samples received were for multiple tests (ie HCV genotype, HCV viral load, and IFN $\lambda$ 3 testing), and multiple blood tubes were not required.

Due to the nature of 'real world' diagnostics, some tests are retrospectively requested, and the practicality of this was assessed for this assay. Both buffy coat and plasma samples were found to be viable at -20°C for 1 and 2 years respectively, and plasma samples were not compromised by multiple freeze/thaw cycles. This has provided confidence that if requests are received retrospectively, the assay can determine genotypes appropriately.

In addition to confirming samples are still viable under different conditions, a diagnostic laboratory has the responsibility of ensuring the testing methods utilized are of the highest standard. It was with this in mind that alternative reagents were assessed following the validation process described. An alternative genotyping mastermix was assessed; the manufacturer reported it to be as efficient as others on the market, and at a lower cost. In this study, it was found to be sub-standard and it was later found that this product had been discontinued. A more successful study was conducted comparing the in-house method to a

commercially available assay; this assay was found to be both reliable and user-friendly. Being a commercial assay, the cost is significantly higher, but it was found to serve well as a confirmatory backup test. It is also important in terms of ongoing validation and quality control to be able to compare performance of an in-house assay on a regular basis, to satisfy both our own piece of mind, and that of external regulatory bodies.

The final modification to the protocol was to alter the volumes of the mastermix from the manufacturer's directions. The initial protocol states a mastermix volume of 13.75µl to be used; this is not achievable with the step-pipette available to our laboratory, and with the high throughput experienced for this test, it was important to investigate alternatives. The initial reduction trialed was not successful, explained by the manufacturer's stringent quality control. However, after fine-tuning the volumes, a subtle change to using a 13µl aliquot was successful.

Following the validation process, association studies were performed, the first of which investigated relationships between IFNλ3 genotype and HCV genotype 1a and 1b patients. A myriad of publications have reported that the greatest effect of IFNλ3 is seen in HCV genotype 1 and 4 patients, without probing further into the subtype level. As Australia has a high number of HCV genotype 1 patients, it was this subset which was investigated. A total of 1,707 patients had received both HCV genotyping and IFNλ3 genotyping at the time of investigation, of which 1,032 were HCV genotype 1a or 1b. It was found that the benefit of the C allele was greatest in HCV 1b patients, and was more so in female than male patients. The study was limited by the lack of clinical notes accompanying the specimens, and therefore many patients did not fit the inclusion criteria of the study. Another constraint was the lack of TT patients, which was possibly due to discontinuation of treatment, and therefore no post-therapy bleed was present, excluding the patient from the data set.

Although VIDRL is the reference laboratory for Victoria, we also receive requests from most States and Territories in Australia. Due to the high number of requests received for IFNλ3 SNP testing, data could be collated, and the frequencies and distributions of IFNλ3 genotypes between the different States and Territories could be analysed (with the exception of Western Australia, for which no samples were received). In the States inhabited by largely Caucasian populations (Queensland, Tasmania, South Australia), the CT genotype accounts for more than 50% of the patients respectively. In contrast, the predominant genotype in New South Wales was CC (50%), and in both the Northern



Territory and Victoria, CC and CT were roughly equal in numbers. Nationwide, the predominant genotype was CT, followed by CC, then TT. This profile is similar to that seen in North America and Western European countries, likely a reflection of migration and ethnic diversification.

With the population correlation studies completed, the direction of study moved on to explore reasons for the effect observed with respect to the IFN $\lambda$ 3 rs12979860 SNP and HCV clearance.

The initial step was to gain clues as to why there are multiple functioning IFN $\lambda$ s in human populations, and to identify the similarities and differences between them. There are three (or four in African populations) functioning *IFN $\lambda$*  genes, along with at least two pseudogenes. Each of the *IFN $\lambda$*  genes signal *via* the same receptors (IL10RB and IFN $\lambda$ R1), and may compete for binding efficiency. The two most closely related, IFN $\lambda$ 2 and IFN $\lambda$ 3 share >96% identity; a key difference identified is the presence of a tyrosine residue in IFN $\lambda$ 2, which appears to have replaced a histidine in the other IFN $\lambda$ s, allowing for additional post-transcriptional regulation.

The evolution of the ancestral T to the modern C allele at rs12979860 has at least two effects, as demonstrated here. Not only do the two variants differ in secondary structure, but the presence of the C allele creates a potential E2F transcription factor binding site, which can act as either a repressor or an activator. This may provide an additional layer of regulation, as compared to the T allele. To determine whether these differences affect transcription, specific ELISA assays were performed to measure IFN $\lambda$ 2 and IFN $\lambda$ 3 plasma concentrations, and to correlate the data with rs12979860 genotypes. Significant differences were observed in IFN $\lambda$ 3, with TT genotypes having the highest concentrations, and a trend was observed in the IFN $\lambda$ 2 assay, with TT having the lowest concentrations. The ratio differences between the two was also significant. Not only does this expose differences in response to viral infection and treatment response, but it also provides some insight into why multiple copies of a gene may be retained through evolution; different kinetics can trigger different responses from virtually identical gene products.

By investigating thymidine phosphorylase (TP) levels in pre-treatment samples of HCV infected individuals, it was determined that higher levels were present in patients with IFN $\lambda$ 3 TT genotype. It has been shown that TP increases following IFN $\alpha$  therapy, but if the threshold has already been reached in a non-medicated state, there is no further scope for

enhancement. This study has shown that the T allele is associated with a pro-inflammatory response, which will likely have implications in other disease or immune states.

The rs12979860 SNP has been well documented to have a significant impact on both spontaneous and treatment induced response to HCV infection, and is the strongest pre-treatment indicator for achieving an SVR. Although treatment options have expanded in recent years, and the correlations between treatment response and IFN $\lambda$ 3 genotype may appear to be lessened, they are not irrelevant. It is probable that this SNP, has a greater impact with a broad range of immune responses, and is likely implicated in many more diseases than is currently known. The work performed here strongly suggests that the impact of rs12979860 is much more far-reaching than in just HCV infection alone.

## References

1. Mohd Hanafiah, K., et al., *Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence*. Hepatology, 2013. **57**(4): p. 1333-42.
2. Lozano, R., et al., *Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010*. Lancet, 2012. **380**(9859): p. 2095-128.
3. Kirby Institute, *HIV, viral hepatitis and sexually transmissible infections in Australia. Annual Surveillance Report*. 2012.
4. Gidding, H.F., et al., *The epidemiology of hepatitis C in Australia: Notifications, treatment uptake and liver transplantations, 1997-2006*. Journal of Gastroenterology and Hepatology, 2009. **24**(10): p. 1648-1654.
5. Iwarson, S., *Transfusion transmitted non-A, non-B hepatitis*. Acta Anaesthesiol Scand Suppl, 1988. **89**: p. 13-5.
6. World Health Organization, *Global Health Sector Strategy On Viral Hepatitis 2016 - 2021* WHO Document Production Services, 2016.
7. Robertson, B., et al., *Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization*. International Committee on Virus Taxonomy. Arch Virol, 1998. **143**(12): p. 2493-503.
8. Cook, S. and E.C. Holmes, *A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission*. Arch Virol, 2006. **151**(2): p. 309-25.
9. Choo, Q.L., et al., *Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome*. Science, 1989. **244**(4902): p. 359-62.
10. Kuo, G., et al., *An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis*. Science, 1989. **244**(4902): p. 362-4.
11. Smith, D.B., et al., *Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource*. Hepatology, 2014. **59**(1): p. 318-27.
12. Simmonds, P., et al., *Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes*. Hepatology, 2005. **42**(4): p. 962-73.
13. Villano, S.A., et al., *Persistence of viremia and the importance of long-term follow-up after acute hepatitis C infection*. Hepatology, 1999. **29**(3): p. 908-14.
14. Mayerat, C., A. Mantegani, and P.C. Frei, *Does hepatitis B virus (HBV) genotype influence the clinical outcome of HBV infection?* Journal of Viral Hepatitis, 1999. **6**(4): p. 299-304.
15. Thio, C.L., *Host genetic factors and antiviral immune responses to hepatitis C virus*. Clin Liver Dis, 2008. **12**(3): p. 713-26, xi.

16. Northcott, M.J., et al., *Prevalence of transfusion-acquired hepatitis C in an Australian bleeding disorders population*. Haemophilia, 2013. **19**(6): p. 847-52.
17. Shepard, C.W., L. Finelli, and M.J. Alter, *Global epidemiology of hepatitis C virus infection*. The Lancet Infectious Diseases, 2005. **5**(9): p. 558-567.
18. World Health Organization, *WHO Global Database on Blood Safety Summary Report 2011*. 2011.
19. Gibney, K.B., et al., *The profile of health problems in African immigrants attending an infectious disease unit in Melbourne, Australia*. Am J Trop Med Hyg, 2009. **80**(5): p. 805-11.
20. Dore, G.J., et al., *Epidemiology of hepatitis C virus infection in Australia*. Australian Family Physician, 2003. **32**(10): p. 796-798.
21. Grebely, J., et al., *Barriers associated with the treatment of hepatitis C virus infection among illicit drug users*. Drug Alcohol Depend, 2008. **93**(1-2): p. 141-7.
22. Strathdee, S.A., et al., *Factors associated with interest in initiating treatment for hepatitis C virus (HCV) infection among young HCV-infected injection drug users*. Clinical Infectious Diseases, 2005. **40**: p. S304-S312.
23. Hellard, M., R. Sacks-Davis, and J. Gold, *Hepatitis C treatment for injection drug users: a review of the available evidence*. Clin Infect Dis, 2009. **49**(4): p. 561-73.
24. Rolls, D.A., et al., *Modelling hepatitis C transmission over a social network of injecting drug users*. J Theor Biol, 2012. **297**: p. 73-87.
25. Hellard, M., et al., *The impact of injecting networks on hepatitis C transmission and treatment in people who inject drugs*. Hepatology, 2014. **60**(6): p. 1861-70.
26. Hellard, M., N. Crofts, and J. Hocking, *Hepatitis C virus among inmates in Victorian correctional facilities*. . 2002.
27. Roberts, E.A. and L. Yeung, *Maternal-infant transmission of hepatitis C virus infection*. Hepatology, 2002. **36**(5 Suppl 1): p. S106-13.
28. Benova, L., et al., *Vertical transmission of hepatitis C virus: systematic review and meta-analysis*. Clin Infect Dis, 2014. **59**(6): p. 765-73.
29. Martinot-Peignoux, M., et al., *Twelve weeks posttreatment follow-up is as relevant as 24 weeks to determine the sustained virologic response in patients with hepatitis C virus receiving pegylated interferon and ribavirin*. Hepatology, 2010. **51**(4): p. 1122-6.
30. Jaeckel, E., et al., *Treatment of acute hepatitis C with interferon alfa-2b*. N Engl J Med, 2001. **345**(20): p. 1452-7.
31. Fried, M.W., et al., *PEGINTERFERON ALFA-2a PLUS RIBAVIRIN FOR CHRONIC HEPATITIS C VIRUS INFECTION*. N Engl J Med, 2012. **347**(13): p. 975-982.
32. Fellay, J., et al., *ITPA gene variants protect against anaemia in patients treated for chronic hepatitis C*. Nature, 2010. **464**(7287): p. 405-8.
33. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
34. Bush, W.S. and J.H. Moore, *Chapter 11: Genome-wide association studies*. PLoS Comput Biol, 2012. **8**(12): p. e1002822.
35. Lander, E.S., *Initial impact of the sequencing of the human genome*. Nature, 2011. **470**(7333): p. 187-97.

36. Poelmans, G., et al., *AKAPs integrate genetic findings for autism spectrum disorders*. Transl Psychiatry, 2013. **3**: p. e270.
37. NCBI. *dbSNP Short Genetic Variations*. 2 October 2013]; Available from: [http://www.ncbi.nlm.nih.gov/projects/SNP/snp\\_summary.cgi](http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi).
38. Hirschhorn, J.N. and M.J. Daly, *Genome-wide association studies for common diseases and complex traits*. Nat Rev Genet, 2005. **6**(2): p. 95-108.
39. Gabriel, S.B., et al., *The structure of haplotype blocks in the human genome*. Science, 2002. **296**(5576): p. 2225-9.
40. Need, A.C. and D.B. Goldstein, *Next generation disparities in human genomics: concerns and remedies*. Trends Genet, 2009. **25**(11): p. 489-94.
41. Ge, D., et al., *Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance*. Nature, 2009. **461**(7262): p. 399-401.
42. Rauch, A., et al., *Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study*. Gastroenterology, 2010. **138**(4): p. 1338-45, 1345 e1-7.
43. Suppiah, V., et al., *IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy*. Nat Genet, 2009. **41**(10): p. 1100-4.
44. Tanaka, Y., et al., *Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C*. Nat Genet, 2009. **41**(10): p. 1105-9.
45. Kotenko, S.V., et al., *IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex*. Nat Immunol, 2003. **4**(1): p. 69-77.
46. Sheppard, P., et al., *IL-28, IL-29 and their class II cytokine receptor IL-28R*. Nat Immunol, 2003. **4**(1): p. 63-8.
47. Abe, H., et al., *IL28 variation affects expression of interferon stimulated genes and peg-interferon and ribavirin therapy*. J Hepatol, 2011. **54**(6): p. 1094-101.
48. Donnelly, R.P. and S.V. Kotenko, *Interferon-lambda: a new addition to an old family*. J Interferon Cytokine Res, 2010. **30**(8): p. 555-64.
49. Osterlund, P., et al., *Gene expression and antiviral activity of alpha/beta interferons and interleukin-29 in virus-infected human myeloid dendritic cells*. J Virol, 2005. **79**(15): p. 9608-17.
50. de Weerd, N.A. and T. Nguyen, *The interferons and their receptors--distribution and regulation*. Immunol Cell Biol, 2012. **90**(5): p. 483-91.
51. McHutchison, J.G., et al., *Peginterferon alfa-2b or alfa-2a with ribavirin for treatment of hepatitis C infection*. N Engl J Med, 2009. **361**(6): p. 580-93.
52. Eslam, M., et al., *IFNL3 polymorphisms predict response to therapy in chronic hepatitis C genotype 2/3 infection*. J Hepatol, 2014. **61**(2): p. 235-41.
53. Thomas, D., et al., *Genetic variation in IL28B and spontaneous clearance of hepatitis C virus*. Nature, 2009. **461**: p. 798-802.
54. Di Lello, F.A., et al., *Effects of the genetic pattern defined by low-density lipoprotein receptor and IL28B genotypes on the outcome of hepatitis C virus infection*. Eur J Clin Microbiol Infect Dis, 2013. **32**(11): p. 1427-35.
55. Prokunina-Olsson, L., et al., *A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus*. Nat Genet, 2013. **45**(2): p. 164-71.

56. Real, L.M., et al., *IFNL4* ss469415590 variant shows similar performance to rs12979860 as predictor of response to treatment against Hepatitis C Virus genotype 1 or 4 in Caucasians. *PLoS One*, 2014. **9**(4): p. e95515.
57. Stattermayer, A.F., et al., *Polymorphisms of interferon-lambda4 and IL28B - effects on treatment response to interferon/ribavirin in patients with chronic hepatitis C*. *Aliment Pharmacol Ther*, 2014. **39**(1): p. 104-11.
58. Lu, Y.F., et al., *Interferon-lambda4 is a cell-autonomous type III interferon associated with pre-treatment hepatitis C virus burden*. *Virology*, 2015. **476**: p. 334-40.
59. Asselah, T., et al., *Liver gene expression signature to predict response to pegylated interferon plus ribavirin combination therapy in patients with chronic hepatitis C*. *Gut*, 2008. **57**(4): p. 516-24.
60. Chen, L., et al., *Hepatic Gene Expression Discriminates Responders and Nonresponders in Treatment of Chronic Hepatitis C Viral Infection*. *Gastroenterology*, 2005. **128**(5): p. 1437-1444.
61. Sarasin-Filipowicz, M., et al., *Interferon signaling and treatment outcome in chronic hepatitis C*. *Proc Natl Acad Sci U S A*, 2008. **105**(19): p. 7034-9.
62. Urban, T.J., et al., *IL28B genotype is associated with differential expression of intrahepatic interferon-stimulated genes in patients with chronic hepatitis C*. *Hepatology*, 2010. **52**(6): p. 1888-96.
63. Dill, M.T., et al., *Interferon-induced gene expression is a stronger predictor of treatment response than IL28B genotype in patients with hepatitis C*. *Gastroenterology*, 2011. **140**(3): p. 1021-31.
64. Honda, M., et al., *Hepatic ISG expression is associated with genetic variation in interleukin 28B and the outcome of IFN therapy for chronic hepatitis C*. *Gastroenterology*, 2010. **139**(2): p. 499-509.
65. Terczynska-Dyla, E., et al., *Reduced IFNlambda4 activity is associated with improved HCV clearance and reduced expression of interferon-stimulated genes*. *Nat Commun*, 2014. **5**: p. 5699.
66. O'Brien, T.R., L. Prokunina-Olsson, and R.P. Donnelly, *IFN-lambda4: the paradoxical new member of the interferon lambda family*. *J Interferon Cytokine Res*, 2014. **34**(11): p. 829-38.
67. Key, F.M., et al., *Selection on a variant associated with improved viral clearance drives local, adaptive pseudogenization of interferon lambda 4 (IFNL4)*. *PLoS Genet*, 2014. **10**(10): p. e1004681.
68. Kinzie, J.L., et al., *African Americans with genotype 1 treated with interferon for chronic hepatitis C have a lower end of treatment response than Caucasians*. *Journal of Viral Hepatitis*, 2001. **8**(4): p. 264-269.
69. Conjeevaram, H.S., et al., *Peginterferon and ribavirin treatment in African American and Caucasian American patients with hepatitis C genotype 1*. *Gastroenterology*, 2006. **131**(2): p. 470-7.
70. Layden-Almer, J.E., et al., *Viral dynamics and response differences in HCV-infected African American and white patients treated with IFN and ribavirin*. *Hepatology*, 2003. **37**(6): p. 1343-50.
71. Sterling, R.K., et al., *A Comparison of the Spectrum of Chronic Hepatitis C Virus Between Caucasians and African Americans*. *Clinical Gastroenterology and Hepatology*, 2004. **2**(6): p. 469-473.

72. Wiley, T.E., J. Brown, and J. Chan, *Hepatitis C infection in African Americans: its natural history and histological progression*. Am J Gastroenterol, 2002. **97**(3): p. 700-6.
73. Nguyen, M.H., et al., *Role of ethnicity in risk for hepatocellular carcinoma in patients with chronic hepatitis C and cirrhosis*. Clin Gastroenterol Hepatol, 2004. **2**(9): p. 820-4.
74. Eslam, M., et al., *Interferon-lambda rs12979860 genotype and liver fibrosis in viral and non-viral chronic liver disease*. Nat Commun, 2015. **6**: p. 6422.
75. Papafragkakis, H., et al., *Depression and pegylated interferon-based hepatitis C treatment*. International Journal of Interferon, Cytokine and Mediator Research, 2012: p. 25.
76. Stewart, B., et al., *Anxiety and depression in Australian chronic hepatitis C outpatients: prevalence and predictors*. Australas Psychiatry, 2012. **20**(6): p. 496-500.
77. Evon, D.M., et al., *Social support and clinical outcomes during antiviral therapy for chronic hepatitis C*. J Psychosom Res, 2011. **71**(5): p. 349-56.
78. Lotrich, F.E., et al., *IL28B Polymorphism Is Associated with Both Side Effects and Clearance of Hepatitis C During Interferon-Alpha Therapy*. J Interferon Cytokine Res, 2010.
79. Holmes, J.A., P.V. Desmond, and A.J. Thompson, *Does IL28B genotyping still have a role in the era of direct-acting antiviral therapy for chronic hepatitis C infection?* J Viral Hepat, 2012. **19**(10): p. 677-84.
80. Kiser, J.J. and C. Flexner, *Direct-acting antiviral agents for hepatitis C virus infection*. Annu Rev Pharmacol Toxicol, 2013. **53**: p. 427-49.
81. Thompson, A.J. and J.G. McHutchison, *Will IL28B polymorphism remain relevant in the era of direct-acting antiviral agents for hepatitis C virus?* Hepatology, 2012. **56**(1): p. 373-81.
82. Shimada, N., et al., *alpha-Fetoprotein is a surrogate marker for predicting treatment failure in telaprevir-based triple combination therapy for genotype 1b chronic hepatitis C Japanese patients with the IL28B minor genotype*. J Med Virol, 2014. **86**(3): p. 461-72.
83. Miyamura, T., et al., *IFNL4 ss469415590 Variant Is Associated with Treatment Response in Japanese HCV Genotype 1 Infected Individuals Treated with IFN-Including Regimens*. Int J Hepatol, 2014. **2014**: p. 723868.
84. Meissner, E.G., et al., *IFNL4-DeltaG genotype is associated with slower viral clearance in hepatitis C, genotype-1 patients treated with sofosbuvir and ribavirin*. J Infect Dis, 2014. **209**(11): p. 1700-4.
85. Camma, C., et al., *Cost-effectiveness of boceprevir or telaprevir for untreated patients with genotype 1 chronic hepatitis C*. Hepatology, 2012. **56**(3): p. 850-60.
86. Lange, C.M. and S. Zeuzem, *Perspectives and challenges of interferon-free therapy for chronic hepatitis C*. J Hepatol, 2013. **58**(3): p. 583-92.
87. Poordad, F., et al., *Boceprevir for Untreated Chronic HCV Genotype 1 Infection*. N Engl J Med, 2011. **364**(13): p. 1195-1206.
88. Wiese, M., et al., *Outcome in a hepatitis C (genotype 1b) single source outbreak in Germany—a 25-year multicenter study*. Journal of Hepatology, 2005. **43**(4): p. 590-598.

89. Wiese, M., et al., *Low frequency of cirrhosis in a hepatitis C (genotype 1b) single-source outbreak in germany: a 20-year multicenter study.* Hepatology, 2000. **32**(1): p. 91-6.
90. Tillmann, H.L., et al., *A polymorphism near IL28B is associated with spontaneous clearance of acute hepatitis C virus and jaundice.* Gastroenterology, 2010. **139**(5): p. 1586-92, 1592 e1.
91. Hoofnagle, J.H., *Course and outcome of hepatitis C.* Hepatology, 2002. **36**(5 Suppl 1): p. S21-9.
92. Rao, H.Y., et al., *IL28B genetic variants and gender are associated with spontaneous clearance of hepatitis C virus infection.* J Viral Hepat, 2012. **19**(3): p. 173-81.
93. Zhang, X.H., et al., *Prevalence of IL-28B and ITPA genotypes in Chinese Han population infected persistently with hepatitis C virus genotype 6 or HCV-1.* J Med Virol, 2013. **85**(7): p. 1163-9.
94. Tillmann, H.L., et al., *Beneficial IL28B genotype associated with lower frequency of hepatic steatosis in patients with chronic hepatitis C.* J Hepatol, 2011. **55**(6): p. 1195-200.
95. Stattermayer, A.F., et al., *Association of the IL28B genotype with insulin resistance in patients with chronic hepatitis C.* J Hepatol, 2012. **57**(3): p. 492-8.
96. Rojas, A., et al., *Hepatitis C virus infection alters lipid metabolism depending on IL28B polymorphism and viral genotype and modulates gene expression in vivo and in vitro.* J Viral Hepat, 2014. **21**(1): p. 19-24.
97. Raglow, Z., et al., *IL28B genotype and the expression of ISGs in normal liver.* Liver Int, 2013. **33**(7): p. 991-8.
98. Allam, S.R., et al., *The association of IL28B polymorphism and graft survival in patients with hepatitis C undergoing liver transplantation.* PLoS One, 2013. **8**(1): p. e54854.
99. Duarte-Rojo, A., et al., *The course of posttransplant hepatitis C infection: comparative impact of donor and recipient source of the favorable IL28B genotype and other variables.* Transplantation, 2012. **94**(2): p. 197-203.
100. Coto-Llerena, M., et al., *Donor and recipient IL28B polymorphisms in HCV-infected patients undergoing antiviral therapy before and after liver transplantation.* Am J Transplant, 2011. **11**(5): p. 1051-7.
101. Cisneros, E., et al., *Increased risk of severe hepatitis C virus recurrence after liver transplantation in patients with a T allele of IL28B rs12979860.* Transplantation, 2012. **94**(3): p. 275-80.
102. Fukuhara, T., et al., *Variants in IL28B in liver recipients and donors correlate with response to peg-interferon and ribavirin therapy for recurrent hepatitis C.* Gastroenterology, 2010. **139**(5): p. 1577-85, 1585 e1-3.
103. Biggins, S.W., et al., *Differential effects of donor and recipient IL28B and DDX58 SNPs on severity of HCV after liver transplantation.* J Hepatol, 2013. **58**(5): p. 969-76.
104. Firpi, R.J., et al., *CC genotype donors for the interleukin-28B single nucleotide polymorphism are associated with better outcomes in hepatitis C after liver transplant.* Liver Int, 2013. **33**(1): p. 72-8.
105. Charlton, M.R., et al., *Interleukin-28B polymorphisms are associated with histological recurrence and treatment response following liver*



- transplantation in patients with hepatitis C virus infection*. Hepatology, 2011. **53**(1): p. 317-24.
106. Lange, C.M., et al., *Impact of donor and recipient IL28B rs12979860 genotypes on hepatitis C virus liver graft reinfection*. J Hepatol, 2011. **55**(2): p. 322-7.
  107. Taylor, A.L., P. Gibbs, and J.A. Bradley, *Acute graft versus host disease following liver transplantation: the enemy within*. Am J Transplant, 2004. **4**(4): p. 466-74.
  108. Assi, M.A., et al., *Graft-vs.-host disease in lung and other solid organ transplant recipients*. Clin Transplant, 2007. **21**(1): p. 1-6.
  109. Markey, K.A., K.P. MacDonald, and G.R. Hill, *Impact of cytokine gene polymorphisms on graft-vs-host disease*. Tissue Antigens, 2008. **72**(6): p. 507-16.
  110. Holmes, J.A., et al., *IL28B genotype is not useful for predicting treatment outcome in Asian chronic hepatitis B patients treated with pegylated interferon-alpha*. J Gastroenterol Hepatol, 2013. **28**(5): p. 861-6.
  111. Kandemir, O., et al., *Chronic hepatitis B and IL28B rs12979860 polymorphism: preliminary study*. Mol Biol Rep, 2013. **40**(11): p. 6189-94.
  112. Li, W., et al., *Expression and gene polymorphisms of interleukin 28B and hepatitis B virus infection in a Chinese Han population*. Liver Int, 2011. **31**(8): p. 1118-26.
  113. Zhang, Q., et al., *IFNL3 (IL28B) polymorphism does not predict long-term response to interferon therapy in HBeAg-positive chronic hepatitis B patients*. J Viral Hepat, 2013.
  114. Petta, S., et al., *IL28B and PNPLA3 polymorphisms affect histological liver damage in patients with non-alcoholic fatty liver disease*. J Hepatol, 2012. **56**(6): p. 1356-62.
  115. Tian, C., et al., *Variant in PNPLA3 is associated with alcoholic liver disease*. Nat Genet, 2010. **42**(1): p. 21-3.
  116. Romeo, S., et al., *Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent*. Int J Obes (Lond), 2010. **34**(1): p. 190-4.
  117. Parczewski, M., et al., *IL28B CC genotype is associated with higher all-cause mortality in antiretroviral-treated HIV-infected patients*. AIDS Res Hum Retroviruses, 2012. **28**(12): p. 1640-6.
  118. Gaudieri, S., et al., *Genetic Variations in IL28B and Allergic Disease in Children*. PLoS One, 2012. **7**(1): p. e30607.
  119. Chen, C., et al., *IL-28B genetic variant is associated with the risk of schizophrenia in the Chinese Han population*. DNA Cell Biol, 2012. **31**(6): p. 988-92.
  120. Fruntes, V. and F. Limosin, *Schizophrenia and viral infection during neurodevelopment: a pathogenesis model?* Med Sci Monit, 2008. **14**(6): p. RA71-77.
  121. Suppiah, V., et al., *IL28B, HLA-C, and KIR variants additively predict response to therapy in chronic hepatitis C virus infection in a European Cohort: a cross-sectional study*. PLoS Med, 2011. **8**(9): p. e1001092.
  122. Eslam, M., et al., *FibroGENE: A gene-based model for staging liver fibrosis*. J Hepatol, 2016. **64**(2): p. 390-8.

123. Mandel, P. and P. Metais, *Les acides nucleiques du plasma sanguin chez l'homme*. CR Acad Sci Paris, 1948. **142**: p. 241-243.
124. Ray Kim, W., *Global epidemiology and burden of hepatitis C*. Microbes Infect, 2002. **4**(12): p. 1219-25.
125. Gonzalez-Masia, J.A., D. Garcia-Olmo, and D.C. Garcia-Olmo, *Circulating nucleic acids in plasma and serum (CNAPS): applications in oncology*. Onco Targets Ther, 2013. **6**: p. 819-32.
126. Zolotukhina, T.V., N.V. Shilova, and E.Y. Voskoboeva, *Analysis of cell-free fetal DNA in plasma and serum of pregnant women*. J Histochem Cytochem, 2005. **53**(3): p. 297-9.
127. Comert, F., et al., *Evaluation of hepatitis C virus RNA stability in room temperature and multiple freeze-thaw cycles by COBAS AmpliPrep/COBAS TaqMan HCV*. Diagn Microbiol Infect Dis, 2013. **75**(1): p. 81-5.
128. Lee, T.-H., et al., *Quantitation of genomic DNA in plasma and serum samples: higher concentrations of genomic DNA found in serum than in plasma*. Transfusion, 2001. **41**: p. 276-282.
129. Madisen, L., et al., *DNA banking: the effects of storage of blood and isolated DNA on the integrity of DNA*. Am J Med Genet, 1987. **27**(2): p. 379-90.
130. Baleriola, C., et al., *Stability of hepatitis C virus, HIV, and hepatitis B virus nucleic acids in plasma samples after long-term storage at -20 degrees C and -70 degrees C*. J Clin Microbiol, 2011. **49**(9): p. 3163-7.
131. Mychaleckyj, J.C., et al., *Buffy coat specimens remain viable as a DNA source for highly multiplexed genome-wide genetic tests after long term storage*. Journal of Translational Medicine, 2011. **9**.
132. Dring, M.M., et al., *Innate immune genes synergize to predict increased risk of chronic disease in hepatitis C virus infection*. Proceedings of the National Academy of Sciences, 2011. **108**(14): p. 5736-5741.
133. Comstock, G.W., et al., *Effects of repeated freeze-thaw cycles on concentrations of cholesterol, micronutrients, and hormones in human plasma and serum*. Clin Chem, 2001. **47**(1): p. 139-42.
134. Jennings, C., D.M. Weaver, and A.W. Kruski, *Effects of freeze-thawing on determinations of cholesterol and high-density lipoproteins in baboon sera*. Clin Chem, 1979. **25**(3): p. 490.
135. Cuhadar, S., et al., *The effect of storage time and freeze-thaw cycles on the stability of serum samples*. Biochemia Medica, 2013: p. 70-77.
136. Ririe, K.M., R.P. Rasmussen, and C.T. Wittwer, *Product differentiation by analysis of DNA melting curves during the polymerase chain reaction*. Anal Biochem, 1997. **245**(2): p. 154-60.
137. Lay, M.J. and C.T. Wittwer, *Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR*. Clin Chem, 1997. **43**(12): p. 2262-7.
138. Proff, C., M.A. Rothschild, and P.M. Schneider, *Low volume PCR (LV-PCR) for STR typing of forensic casework samples*. International Congress Series, 2006. **1288**: p. 645-647.
139. Fereidouni, S.R., et al., *Saving resources: avian influenza surveillance using pooled swab samples and reduced reaction volumes in real-time RT-PCR*. J Virol Methods, 2012. **186**(1-2): p. 119-25.
140. Westring, C.G., et al., *Validation of reduced-scale reactions for the Quantifiler Human DNA kit*. J Forensic Sci, 2007. **52**(5): p. 1035-43.

141. Bessekri, M.W., et al., *Comparative study on the effects of reduced PCR reaction volumes and increased cycle number, on the sensitivity and the stochastic threshold of the AmpFISTR Identifiler® Plus kit*. Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e306-e307.
142. Gaines, M.L., et al., *Reduced Volume PCR Amplification Reactions Using the AmpFISTR® Profiler Plus™ Kit*. J Forensic Sci, 2002. **47**(6).
143. David, G. and P. Buckle, *A questionnaire survey of the ergonomic problems associated with pipettes and their usage with specific reference to work related upper limb disorders*. Applied Ergonomics, 1997. **28**(4): p. 257-262.
144. Bjorksten, M.G., B. Almby, and E.S. Jansson, *Hand and Shoulder Ailments among Laboratory Technicians Using Modern Plunger-Operated Pipettes*. Applied Ergonomics, 1994. **25**(2): p. 88-94.
145. Hoang, X.S., et al., *P035 A simple and reliable method for genotyping IL28B polymorphism in patients infected with hepatitis C virus*. Cytokine, 2012. **59**(3): p. 529.
146. Nattermann, J., et al., *Genetic variation in IL28B and treatment-induced clearance of hepatitis C virus in HIV-positive patients with acute and chronic hepatitis C*. J Infect Dis, 2011. **203**(5): p. 595-601.
147. Bes, M., et al., *IL28B genetic variation and hepatitis C virus-specific CD4(+) T-cell responses in anti-HCV-positive blood donors*. J Viral Hepat, 2012. **19**(12): p. 867-71.
148. Sievert, W., et al., *A systematic review of hepatitis C virus epidemiology in Asia, Australia and Egypt*. Liver Int, 2011. **31 Suppl 2**: p. 61-80.
149. McCaw, R., et al., *Hepatitis C virus genotypes in Australia*. Journal of Viral Hepatitis, 1997. **4**(5): p. 351-357.
150. Di Bisceglie, A.M. and J.H. Hoofnagle, *Optimal therapy of hepatitis C*. Hepatology, 2002. **36**(5 Suppl 1): p. S121-7.
151. Antonucci, G., et al., *The effect of age on response to therapy with peginterferon alpha plus ribavirin in a cohort of patients with chronic HCV hepatitis including subjects older than 65 yr*. Am J Gastroenterol, 2007. **102**(7): p. 1383-91.
152. Koruk, S., *The follow-up results with sustained virologic response in chronic hepatitis C patients in Şanlıurfa/Turkey*. Journal of Microbiology and Infectious Diseases, 2012. **2**(1): p. 141-20.
153. Hepburn, M.J., et al., *Differences in treatment outcome for hepatitis C among ethnic groups*. Am J Med, 2004. **117**(3): p. 163-8.
154. Muir, A.J., J.D. Bornstein, and P.G. Killenberg, *Peginterferon alfa-2b and ribavirin for the treatment of chronic hepatitis C in blacks and non-Hispanic whites*. N Engl J Med, 2004. **350**(22): p. 2265-71.
155. Bressler, B.L., et al., *High body mass index is an independent risk factor for nonresponse to antiviral treatment in chronic hepatitis C*. Hepatology, 2003. **38**(3): p. 639-44.
156. Mangia, A., et al., *An IL28B polymorphism determines treatment response of hepatitis C virus genotype 2 or 3 patients who do not achieve a rapid virologic response*. Gastroenterology, 2010. **139**(3): p. 821-7, 827 e1.
157. Rivero-Juarez, A., et al., *Association between the IL28B genotype and hepatitis C viral kinetics in the early days of treatment with pegylated*

- interferon plus ribavirin in HIV/HCV co-infected patients with genotype 1 or 4.* J Antimicrob Chemother, 2012. **67**(1): p. 202-5.
158. Grebely, J., et al., *Adherence to treatment for recently acquired hepatitis C virus (HCV) infection among injecting drug users.* J Hepatol, 2011. **55**(1): p. 76-85.
  159. Dev, A.T., et al., *Southeast Asian patients with chronic hepatitis C: the impact of novel genotypes and race on treatment outcome.* Hepatology, 2002. **36**(5): p. 1259-65.
  160. Nguyen, L.H. and M.H. Nguyen, *Systematic review: Asian patients with chronic hepatitis C infection.* Aliment Pharmacol Ther, 2013. **37**(10): p. 921-36.
  161. Sinn, D.H., et al., *Association of a single nucleotide polymorphism near the interleukin-28B gene with response to hepatitis C therapy in Asian patients.* J Gastroenterol Hepatol, 2011. **26**(9): p. 1374-9.
  162. Liu, C.-H. and J.-H. Kao, *IL28B Genotype on HCV Infection in Asia.* Current Hepatitis Reports, 2013. **12**(3): p. 149-156.
  163. Gower, E., et al., *Global epidemiology and genotype distribution of the hepatitis C virus infection.* J Hepatol, 2014.
  164. Messina, J.P., et al., *Global distribution and prevalence of hepatitis C virus genotypes.* Hepatology, 2014.
  165. Thompson, A.J., et al., *Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus.* Gastroenterology, 2010. **139**(1): p. 120-9 e18.
  166. Bourliere, M., et al., *Epidemiological changes in hepatitis C virus genotypes in France: evidence in intravenous drug users.* J Viral Hepat, 2002. **9**(1): p. 62-70.
  167. Raptopoulou, M., et al., *Significant epidemiological changes in chronic hepatitis C infection: results of the nationwide HEPNET-GREECE cohort study.* Hippokratia, 2011. **15**(1): p. 26-31.
  168. Savvas, S.P., et al., *Changes in epidemiological patterns of HCV infection and their impact on liver disease over the last 20 years in Greece.* J Viral Hepat, 2005. **12**(5): p. 551-7.
  169. Esteban, J.I., S. Sauleda, and J. Quer, *The changing epidemiology of hepatitis C virus infection in Europe.* J Hepatol, 2008. **48**(1): p. 148-62.
  170. Hellard, M.E., J.S. Hocking, and N. Crofts, *The prevalence and the risk behaviours associated with the transmission of hepatitis C virus in Australian correctional facilities.* Epidemiology and Infection, 2004. **132**(3): p. 409-415.
  171. Teutsch, S., et al., *Incidence of primary hepatitis C infection and risk factors for transmission in an Australian prisoner cohort.* BMC Public Health, 2010. **10**: p. 633.
  172. Inoue, G., et al., *Hepatitis C virus clearance is prominent in women in an endemic area.* Journal of Gastroenterology and Hepatology, 2000. **15**(9): p. 1054-1058.
  173. Bakr, I., et al., *Higher clearance of hepatitis C virus infection in females compared with males.* Gut, 2006. **55**(8): p. 1183-7.

174. Hofer, H., et al., *Spontaneous viral clearance in patients with acute hepatitis C can be predicted by repeated measurements of serum viral load*. Hepatology, 2003. **37**(1): p. 60-4.
175. European Association for Study of, L., *EASL Clinical Practice Guidelines: management of hepatitis C virus infection*. J Hepatol, 2014. **60**(2): p. 392-420.
176. Gerlach, J.T., et al., *Acute hepatitis C: high rate of both spontaneous and treatment-induced viral clearance*. Gastroenterology, 2003. **125**(1): p. 80-88.
177. Jensen, D.M., et al., *Early identification of HCV genotype 1 patients responding to 24 weeks peginterferon alpha-2a (40 kd)/ribavirin therapy*. Hepatology, 2006. **43**(5): p. 954-60.
178. Pellicelli, A.M., et al., *HCV genotype 1a shows a better virological response to antiviral therapy than HCV genotype 1b*. BMC Gastroenterol, 2012. **12**: p. 162.
179. Poordad, F., et al., *Factors that predict response of patients with hepatitis C virus infection to boceprevir*. Gastroenterology, 2012. **143**(3): p. 608-18 e1-5.
180. Ascione, A., *Boceprevir in chronic hepatitis C infection: a perspective review*. Ther Adv Chronic Dis, 2012. **3**(3): p. 113-21.
181. Pickett, B.E., R. Striker, and E.J. Lefkowitz, *Evidence for separation of HCV subtype 1a into two distinct clades*. J Viral Hepat, 2011. **18**(9): p. 608-18.
182. Zeuzem, S., et al., *Faldaprevir and deleobuvir for HCV genotype 1 infection*. N Engl J Med, 2013. **369**(7): p. 630-9.
183. Davis, G.L., et al., *Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C*. Hepatology, 2003. **38**(3): p. 645-52.
184. Ladero, J.M., et al., *"12 weeks' stopping rule" in the treatment of genotype 1 chronic hepatitis C: Two prognostic categories under the same label?* Scandinavian Journal of Gastroenterology, 2008. **43**(8): p. 979-983.
185. Neukam, K., et al., *Prediction of response to pegylated interferon plus ribavirin in HIV/hepatitis C virus (HCV)-coinfected patients using HCV genotype, IL28B variations, and HCV-RNA load*. J Hepatol, 2012. **56**(4): p. 788-94.
186. Younossi, Z.M., et al., *The impact of IL28B genotype on the gene expression profile of patients with chronic hepatitis C treated with pegylated interferon alpha and ribavirin*. J Transl Med, 2012. **10**: p. 25.
187. Akkarathamrongsin, S., et al., *Early viral kinetics during hepatitis C virus genotype 6 treatment according to IL28B polymorphisms*. World J Gastroenterol, 2014. **20**(30): p. 10599-605.
188. Visco-Comandini, U., et al., *No impact of interleukin-28B polymorphisms on spontaneous or drug-induced hepatitis delta virus clearance*. Dig Liver Dis, 2014. **46**(4): p. 348-52.
189. Howell, C.D., et al., *Single nucleotide polymorphism upstream of interleukin 28B associated with phase 1 and phase 2 of early viral kinetics in patients infected with HCV genotype 1*. J Hepatol, 2012. **56**(3): p. 557-63.
190. Stenkvist, J., A. Sonnerborg, and O. Weiland, *HCV RNA decline in chronic HCV genotype 2 and 3 during standard of care treatment according to IL28B polymorphism*. J Viral Hepat, 2013. **20**(3): p. 193-9.

191. McCarthy, J.J., et al., *Replicated association between an IL28B gene variant and a sustained response to pegylated interferon and ribavirin*. Gastroenterology, 2010. **138**(7): p. 2307-14.
192. Rallon, N.I., et al., *Differences in virological response to peginterferon-alpha plus ribavirin in HIV-positive patients coinfectd with HCV subtypes 1a or 1b*. J Acquir Immune Defic Syndr, 2012. **60**(2): p. 117-23.
193. Vispo, E., et al., *Different impact of IL28B polymorphisms on response to peginterferon-alpha plus ribavirin in HIV-positive patients infected with HCV subtypes 1a or 1b*. J Clin Virol, 2012. **55**(1): p. 58-61.
194. Australian Bureau of Statistics. *Australia's population by country of birth*. 2014 2.9.2014]; Available from: <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/3412.0Chapter12011-12%20and%202012-13>.
195. Phillips, J., M. Klapdor, and J. Simon-Davies, *Migration to Australia since federation a guide to the statistics*, 2010, Commonwealth of Australia.
196. Offer, S.M., et al., *A DPYD variant (Y186C) in individuals of african ancestry is associated with reduced DPD enzyme activity*. Clin Pharmacol Ther, 2013. **94**(1): p. 158-66.
197. Hu, X., et al., *Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer's disease*. PLoS One, 2011. **6**(2): p. e16616.
198. Pasaniuc, B., et al., *Enhanced statistical tests for GWAS in admixed populations: assessment using African Americans from CARE and a Breast Cancer Consortium*. PLoS Genet, 2011. **7**(4): p. e1001371.
199. *Hep-C-Strategy2014*.
200. Razali, K., et al., *Modelling the hepatitis C virus epidemic in Australia*. Drug Alcohol Depend, 2007. **91**(2-3): p. 228-35.
201. *Hepatitis C Virus Projections Working Group estimates and projections of the hepatitis c virus epidemic in australia 2006*.
202. Lindh, M., et al., *IL28B polymorphisms determine early viral kinetics and treatment outcome in patients receiving peginterferon/ribavirin for chronic hepatitis C genotype 1*. J Viral Hepat, 2011. **18**(7): p. e325-31.
203. Roberts, S.K., et al., *Distribution of interferon lambda-3 gene polymorphisms in Australian patients with previously untreated genotype 1 chronic hepatitis C: Analysis from the PREDICT and CHARIOT studies*. J Gastroenterol Hepatol, 2014. **29**(1): p. 179-84.
204. Davis, J.S., A.C. Kulatunga, and K. Hajkiewicz, *Outcomes for Indigenous and non-Indigenous patients who access treatment for hepatitis C in the Top End of the Northern Territory*. Med J Aust, 2013. **199**(1): p. 23.
205. Horyniak, D., et al., *Establishing the Melbourne Injecting Drug User Cohort Study (MIX): rationale, methods, and baseline and twelve-month follow-up results*. Harm Reduct J, 2013. **10**: p. 11.
206. Crofts, N., et al., *Blood-borne virus infections among Australian injecting drug users: implications for spread of HIV*. Eur J Epidemiol, 1994. **10**(6): p. 687-94.
207. Butler, T.G., et al., *Hepatitis B and C in New South Wales prisons: prevalence and risk factors*. Med J Aust, 1997. **166**(3): p. 127-30.
208. Stiff, A. and W. Carson Iii, *Investigations of interferon-lambda for the treatment of cancer*. J Innate Immun, 2015. **7**(3): p. 243-50.

209. Ohno, S.L., *Evolution by Gene Duplication*. 1970: Springer-Verlag.
210. Canestro, C., et al., *Impact of gene gains, losses and duplication modes on the origin and diversification of vertebrates*. Semin Cell Dev Biol, 2013. **24**(2): p. 83-94.
211. Salvucci, E., *Microbiome, holobiont and the net of life*. Crit Rev Microbiol, 2016. **42**(3): p. 485-94.
212. Villarino, A.V., et al., *Mechanisms of Jak/STAT signaling in immunity and disease*. J Immunol, 2015. **194**(1): p. 21-7.
213. Krause, C.D. and S. Pestka, *Cut, copy, move, delete: The study of human interferon genes reveal multiple mechanisms underlying their evolution in amniotes*. Cytokine, 2015. **76**(2): p. 480-95.
214. Zerbino, D.R., et al., *The ensembl regulatory build*. Genome Biol, 2015. **16**: p. 56.
215. Oti, M., et al., *CTCF-mediated chromatin loops enclose inducible gene regulatory domains*. BMC Genomics, 2016. **17**(1): p. 252.
216. Lasfar, A., et al., *Characterization of the mouse IFN-lambda ligand-receptor system: IFN-lambdas exhibit antitumor activity against B16 melanoma*. Cancer Res, 2006. **66**(8): p. 4468-77.
217. Fox, B.A., P.O. Sheppard, and P.J. O'Hara, *The role of genomic data in the discovery, annotation and evolutionary interpretation of the interferon-lambda family*. PLoS One, 2009. **4**(3): p. e4933.
218. Hoopes, J.I., et al., *APOBEC3A and APOBEC3B Preferentially Deaminate the Lagging Strand Template during DNA Replication*. Cell Rep, 2016. **14**(6): p. 1273-82.
219. Szklarczyk, D., et al., *STRING v10: protein-protein interaction networks, integrated over the tree of life*. Nucleic Acids Res, 2015. **43**(Database issue): p. D447-52.
220. Wack, A., E. Terczynska-Dyla, and R. Hartmann, *Guarding the frontiers: the biology of type III interferons*. Nat Immunol, 2015. **16**(8): p. 802-9.
221. McWilliam, H., et al., *Analysis Tool Web Services from the EMBL-EBI*. Nucleic Acids Res, 2013. **41**(Web Server issue): p. W597-600.
222. Egli, A., et al., *The impact of the interferon-lambda family on the innate and adaptive immune response to viral infections*. Emerg Microbes Infect, 2014. **3**(7): p. e51.
223. Laidlaw, S.M. and L.B. Dustin, *Interferon lambda: opportunities, risks, and uncertainties in the fight against HCV*. Front Immunol, 2014. **5**: p. 545.
224. Bocker, C., et al., *Evolutionary divergence and functions of the human interleukin (IL) gene family*. Hum Genomics, 2010. **5**(1): p. 30-55.
225. Chinnaswamy, S., *Genetic variants at the IFNL3 locus and their association with hepatitis C virus infections reveal novel insights into host-virus interactions*. J Interferon Cytokine Res, 2014. **34**(7): p. 479-97.
226. Nourry, C., S.G. Grant, and J.P. Borg, *PDZ domain proteins: plug and play!* Sci STKE, 2003. **2003**(179): p. RE7.
227. Kundu, K., et al., *MoDPepInt: an interactive web server for prediction of modular domain-peptide interactions*. Bioinformatics, 2014. **30**(18): p. 2668-9.
228. Gould, C.M., et al., *ELM: the status of the 2010 eukaryotic linear motif resource*. Nucleic Acids Res, 2010. **38**(Database issue): p. D167-80.

229. Kallberg, M., et al., *Template-based protein structure modeling using the RaptorX web server*. Nat Protoc, 2012. **7**(8): p. 1511-22.
230. Iaquinta, P.J. and J.A. Lees, *Life and death decisions by the E2F transcription factors*. Curr Opin Cell Biol, 2007. **19**(6): p. 649-57.
231. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction*. Nucleic Acids Res, 2003. **31**(13): p. 3406-15.
232. Eddy, J. and N. Maizels, *Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes*. Nucleic Acids Res, 2008. **36**(4): p. 1321-33.
233. Bochman, M.L., K. Paeschke, and V.A. Zakian, *DNA secondary structures: stability and function of G-quadruplex structures*. Nat Rev Genet, 2012. **13**(11): p. 770-80.
234. Rhodes, D. and H.J. Lipps, *G-quadruplexes and their regulatory roles in biology*. Nucleic Acids Res, 2015. **43**(18): p. 8627-37.
235. Endoh, T. and N. Sugimoto, *Mechanical insights into ribosomal progression overcoming RNA G-quadruplex from periodical translation suppression in cells*. Sci Rep, 2016. **6**: p. 22719.
236. Duong, F.H., et al., *IFN-lambda receptor 1 expression is induced in chronic hepatitis C and correlates with the IFN-lambda3 genotype and with nonresponsiveness to IFN-alpha therapies*. J Exp Med, 2014. **211**(5): p. 857-68.
237. McFarland, A.P., et al., *The favorable IFNL3 genotype escapes mRNA decay mediated by AU-rich elements and hepatitis C virus-induced microRNAs*. Nat Immunol, 2014. **15**(1): p. 72-9.
238. Langhans, B., et al., *Interferon-lambda serum levels in hepatitis C*. Journal of Hepatology, 2011. **54**: p. 859-865.
239. Al-Qahtani, A., et al., *Correlation between genetic variations and serum level of interleukin 28B with virus genotypes and disease progression in chronic hepatitis C virus infection*. J Immunol Res, 2015. **2015**: p. 768470.
240. Rallon, N.I., et al., *Impact of IL28B gene polymorphisms on interferon-lambda3 plasma levels during pegylated interferon-alpha/ribavirin therapy for chronic hepatitis C in patients coinfectd with HIV*. J Antimicrob Chemother, 2012. **67**(5): p. 1246-9.
241. Stenman, G., et al., *Regional localization of the human platelet-derived endothelial cell growth factor (ECGF1) gene to chromosome 22q13*. Cytogenet Cell Genet, 1992. **59**(1): p. 22-3.
242. Kamatani, N., et al., *Purine and Pyrimidine Metabolism*. 2013: p. 1-38.
243. Bronckaers, A., et al., *The dual role of thymidine phosphorylase in cancer development and chemotherapy*. Med Res Rev, 2009. **29**(6): p. 903-53.
244. Tabata, S., et al., *Thymidine phosphorylase enhances reactive oxygen species generation and interleukin-8 expression in human cancer cells*. Oncol Rep, 2012. **28**(3): p. 895-902.
245. Fujioka, S., et al., *NF-kappaB and AP-1 connection: mechanism of NF-kappaB-dependent regulation of AP-1 activity*. Mol Cell Biol, 2004. **24**(17): p. 7806-19.
246. Graff, J.R., et al., *Mapping Patterns of CpG Island Methylation in Normal and Neoplastic Cells Implicates Both Upstream and Downstream Regions in de Novo Methylation*. The Journal of Biological Chemistry, 1997. **272**(35): p. 22322-22329.



247. Goto, H., et al., *Interferon gamma-dependent induction of thymidine phosphorylase/platelet-derived endothelial growth factor through gamma-activated sequence-like element in human macrophages*. *Cancer Res*, 2001. **61**(2): p. 469-73.
248. Focher, F. and S. Spadari, *Thymidine phosphorylase: a two-face Janus in anticancer chemotherapy*. *Curr Cancer Drug Targets*, 2001. **1**(2): p. 141-53.
249. Huppert, J.L., *Structure, location and interactions of G-quadruplexes*. *FEBS J*, 2010. **277**(17): p. 3452-8.
250. Chen, L.C., et al., *Thymidine phosphorylase mRNA stability and protein levels are increased through ERK-mediated cytoplasmic accumulation of hnRNP K in nasopharyngeal carcinoma cells*. *Oncogene*, 2009. **28**(17): p. 1904-15.
251. Mathews, C.K. and S. Song, *Maintaining precursor pools for mitochondrial DNA replication*. *FASEB J*, 2007. **21**(10): p. 2294-303.
252. Gonzalez-Vioque, E., et al., *Limited dCTP availability accounts for mitochondrial DNA depletion in mitochondrial neurogastrointestinal encephalomyopathy (MNGIE)*. *PLoS Genet*, 2011. **7**(3): p. e1002035.
253. Pontarin, G., et al., *Mammalian ribonucleotide reductase subunit p53R2 is required for mitochondrial DNA replication and DNA repair in quiescent cells*. *Proc Natl Acad Sci U S A*, 2012. **109**(33): p. 13302-7.
254. Reichard, P., *Ribonucleotide reductase and deoxyribonucleotide pools*. *Basic Life Science*, 1985. **31**: p. 33-45.
255. Nishino, I., A. Spinazzola, and M. Hirano, *Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder*. *Science*, 1999. **283**(5402): p. 689-92.
256. Khodursky, A., E.C. Guzman, and P.C. Hanawalt, *Thymineless Death Lives On: New Insights into a Classic Phenomenon*. *Annu Rev Microbiol*, 2015. **69**: p. 247-63.
257. Sangurdekar, D.P., et al., *Thymineless death is associated with loss of essential genetic information from the replication origin*. *Mol Microbiol*, 2010. **75**(6): p. 1455-67.
258. Kuong, K.J. and A. Kuzminov, *Disintegration of nascent replication bubbles during thymine starvation triggers RecA- and RecBCD-dependent replication origin destruction*. *J Biol Chem*, 2012. **287**(28): p. 23958-70.
259. de Bruin, M., et al., *Upregulation of platelet derived endothelial cell growth factor/thymidine phosphorylase by interferon alpha*. *Nucleosides Nucleotides Nucleic Acids*, 2004. **23**(8-9): p. 1367-70.
260. Takebayashi, Y., et al., *The correlation of thymidine phosphorylase activity with the expression of interleukin 1 $\alpha$ , interferon  $\alpha$  and interferon  $\gamma$  in human colorectal carcinoma*. *Cancer Letters*, 1995. **95**: p. 57-62.
261. Nakayama, Y., et al., *Expression levels of thymidine phosphorylase (TP) and dihydropyrimidine dehydrogenase (DPD) in patients with gastrointestinal cancer*. *Anticancer Res*, 2005. **25**(6A): p. 3755-61.
262. Makower, D., et al., *Interferon induces thymidine phosphorylase/platelet-derived endothelial cell growth factor expression in vivo*. *Clin Cancer Res*, 1997. **3**(6): p. 923-9.
263. Braybrooke, J.P., et al., *Induction of thymidine phosphorylase as a pharmacodynamic end-point in patients with advanced carcinoma treated*

- with 5-fluorouracil, folinic acid and interferon alpha*. British Journal of Cancer, 2000. **83**(2): p. 219-224.
264. Shaw, T., R.H. Smillie, and D.G. MacPhee, *The role of blood platelets in nucleoside metabolism: assay, cellular location and significance of thymidine phosphorylase in human blood*. Mutat Res, 1988. **200**(1-2): p. 99-116.
  265. Sato, M., et al., *IL28B minor allele is associated with a younger age of onset of hepatocellular carcinoma in patients with chronic hepatitis C virus infection*. J Gastroenterol, 2014. **49**(4): p. 748-54.
  266. Yamamoto, H., et al., *Platelet derived endothelial cell growth factor/thymidine phosphorylase enhanced human IgE production*. Allergol Int, 2011. **60**(1): p. 79-85.
  267. Fukushima, M., et al., *Induction of thymidine phosphorylase by interferon and taxanes occurs only in human cancer cells with low thymidine phosphorylase activity*. Cancer Lett, 2002. **187**(1-2): p. 103-10.
  268. Labarga, P., et al., *Association between IL28B gene polymorphisms and plasma HCV-RNA levels in HIV/HCV-co-infected patients*. AIDS, 2011. **25**(6): p. 761-6.