AUTOPHAGY-RELATED GENE *ATG5* GENETIC VARIANTS AND HEPATITIS B VIRUS INFECTION SUSCEPTIBILITY

Thesis submitted in fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

By

AVNI VIJ

Department of Biotechnology and Bioinformatics
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT, DISTRICT SOLAN, H.P., INDIA
July 2018

DECLARATION

I hereby declare that the work reported in the Ph.D. thesis entitled **"Autophagy-related gene *ATG5* genetic variants and Hepatitis B Virus infection susceptibility"**, submitted at **Jaypee University of Information Technology, Waknaghat, India,** is an authentic record of my work carried out under the supervision of **Dr. Harish Changotra**. I have not submitted this work elsewhere for any other degree or diploma. I am fully responsible for the contents of my Ph.D. Thesis.

Avni Vij

Enrollment no: 136555

Department of Biotechnology and Bioinformatics

Jaypee University of Information Technology

Waknaghat, India-173234

Date:

CERTIFICATE

This is to certify that the work reported in the Ph.D. thesis entitled **"Autophagy-related gene *ATG5* genetic variants and Hepatitis B Virus infection susceptibility",** submitted by **Avni Vij** (enrollment no: 136555) at **Jaypee University of Information Technology, Waknaghat**, **India,** is a bonafide record of her original work carried out under my supervision. This work has not been submitted elsewhere for any other degree or diploma.

Supervisor:

Dr. Harish Changotra

Associate Professor

Department of Biotechnology and Bioinformatics

Jaypee University of Information Technology

Waknaghat, India-173234

Date:

ACKNOWLEDGEMENTS

As my thoughts roll down into the past, when I was undertaking this project, my heart overflows with emotions and gratitude towards people who believed in me and spent their time and energy for the same. I could have never reached the heights or explored the depths without the help, support, guidance and efforts of a lot of people.

First of all, I would thank ***the Almighty GOD*** *for guiding me throughout and bestowing on me the strength to make the right decisions with dignity, to know when to hold on and when to let go and teaching me to be humane at all times.*

"A teacher affects eternity; he can never tell where his influence stops." Such has been the effect of my guide and mentor, ***Dr. Harish Changotra****, Associate Professor, Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat. I have been amazingly fortunate to have a teacher who gave me the freedom to explore on my own, and at the same time bestowed upon me the guidance to recover when my steps faltered, who taught me how to question thoughts and express ideas. His patient guidance, generous advice, painstaking efforts and sharp and searching intellect have made it possible for me to complete this work.*

It has been a privilege to work at the Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology. I am deeply grateful to ***Dr. Sudhir Kumar,*** *Associate Professor, Acting Head-Department of Biotechnology and Bioinformatics, for providing me this platform and for his unending support and encouragement throughout my Ph.D. I would also like to express my grateful thanks to the JUIT administration,* ***Prof. (Dr.) Vinod Kumar****, Vice Chancellor, JUIT;* ***Prof. (Dr.) Samir Dev Gupta,*** *Dean (Academic and Research);* ***Maj Gen Rakesh Bassi, SM (Retd.),*** *Registrar & Dean of Students, for providing Junior Research Fellowship and all the necessary facilities and infrastructure to carry out my Ph.D. work.*

I would also like to thank ***Dr. Ajay Dusseja****, Department of Hepatology, PGIMER, Chandigarh for providing us the access to collect blood samples from the Hepatitis B infected patients required for the study.*

A special note of thanks and appreciation for my senior, ***Dr. Swapnil Jain*** *and my other lab mates,* ***Mr. Rohit Randhawa****,* ***Mrs. Ambika Sharma*** *and* ***Mr. Sanjay Singh****, for their constant support, encouragement and unending help throughout the course.*

I would like to thank my batch mates and dear friends, ***Ms. Ritika Verma*** *and* ***Ms. Poonam Katoch*** *for being there to share all the good and bad experiences throughout our Ph.D. duration, for always helping me and supporting me in times of need, for sailing me through the hardships and helping me achieve my goals and targets.*

A special note of thanks and appreciation for the technical and non-technical staff of the department, especially, ***Mr. Baleshwar, Mrs. Mamta, Mrs. Somlata, Mr. Ismail, Mr. Sonika*** *and* ***Mr. Ravi Kant*** *for their help and assistance.*

I cannot thank enough my father ***Mr. Pawan Vij,*** *my mother* ***Mrs. Parminder Vij*** *and my brother* ***Mr. Ankur Vij*** *who has been my strength and faith throughout my life, have always been there for me, always believed in me and loved me unconditionally. I wish and pray for their good health and happiness.*

I thank all those who have directly and indirectly helped me in my endeavours and I extend my apologies to those whom I have forgotten to name.

In the end, I would once again thank ***God*** *for always answering my prayers and showering his blessings upon me.*

Avni Vij

TABLE OF CONTENTS

LIST OF ACRONYMS & ABBREVIATIONS

ALT	Alanine aminotransferase
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
A-RFLP	Artificial Restriction Fragment Length Polymorphism
ARMS	Amplification Refractory Mutation System
ASH	Allele specific hybridization
ASPCR	Allele specific Polymerase Chain Reaction
AST	Aspartate aminotransferase
ATG	Autophagy-related
BDM-PUB	Bayesian Discriminant Method for Prediction of Ubiquitination sites
BLAST	Basic Local Alignment Search Tool
cccDNA	Closed circular DNA
CHB	Chronic Hepatitis B
cSNPs	Coding-Single Nucleotide Polymorphisms
CSS-PALM	Clustering and scoring strategy to predict palmitoylation sites
CUPSTAT	Cologne University Stability Analysis Tool
DASH	Dynamic Allele specific Hybridization
dbSNP	Data base SNP
DNA	Deoxyribonucleic acid
EDTA	Ethylene diamine tetra acetic acid
EM	Expectation maximum
EtBr	Ethidium bromide
FADD	Fas Associated Death Receptor
FEN	Flap endonuclease

g	Gram
GPS-CCD	Group based prediction system for calpain cleavage detection
GWAS	Genome Wide Association Study
HBcAb	Hepatitis B core Antibody
HBcAg	Hepatitis B core Antigen
HBsAb	Hepatitis B surface Antibody
HBsAg	Hepatitis B surface Antigen
HBV	Hepatitis B Virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
HLA	Human leucocyte antigen
H_p	Surrounding hydrophobicity
iSNPs	Intronic Single Nucleotide Polymorphisms
LD	Linkage disequilibrium
LFT	Liver function test
LRO	Long Range Order
mg	Milligram
MHCII	Major Histocompatibility Complex II
ml	Milliliter
MTCT	Mother to Child Transmission
NAFLD	Non-Alcoholic Fatty Liver Disease
NCBI	National Centre for Biotechnology Information
NMO	Neuromyelitis Optica
NPC	Nuclear Pore Complex
nsSNP	Non-synonymous Single Nucleotide polymorphism

NTCP	Sodium Taurocholate Cotransporting Polypeptide
ORF	Open Reading Frame
PBMCs	Peripheral Blood Mononuclear Cells
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PE	Phosphatidylethanolamine
PMN	Polymorphic mono nuclear cells
PolyPhen	Polymorphism Phenotyping
PTMs	Post Translation Modification sites
QTL	Quantitative Trait Loci
RBCs	Red Blood Cells
rcDNA	Relaxed circular DNA
RFLP	Restriction Fragment Length Polymorphism
RT	Room Temperature
SC	Stabilizing centre
SGOT	Serum glutamic oxaloacetic transaminases
SGPT	Serum glutamic pyruvic transaminases
SHB	Small Surface proteins of Hepatitis B virus
SIFT	Sorting Intolerant Form Tolerant
SLE	Systemic Lupus Erythematosus
SNP	Single Nucleotide Polymorphism
SSCPs	Single Strand Conformation Polymorphisms
sSNPs	Synonymous Single Nucleotide Polymorphisms
STD	Sexually Transmitted Disease
SVM	Support Vector Machine
TAE	Tris acetate-ethylene diamine tetra acetic acid

TE	Tris- ethylene diamine tetra acetic acid
VDA	Variant Detector Array
WBCs	White Blood Cells
WET-STAB	Weighted Decision Table Method for stability analysis
WHO	World Health Organization

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Hepatitis B is a severe liver inflammation caused by Hepatitis B virus (HBV). Host genetic factors influence HBV acquisition and outcome. Moreover, HBV has been demonstrated to utilize an autophagy process to promote its replication. Various autophagy-related (*ATG*) genes regulate the process at different stages and among them; *ATG5* is a key regulatory gene that promotes autophagosome formation by interacting with ATG12 and ATG16L1 to form a complex with E3 like activity. Single Nucleotide Polymorphisms (SNPs) in the *ATG5* gene have been strongly linked to a variety of diseases like asthma, Paget disease of bone and thyroid carcinoma. Moreover, a non-synonymous SNP (nsSNP), E122D in ATG5 has been shown to impair its interaction with ATG12 leading to the disruption of autophagosome formation. However, until now, SNPs in ATG5 gene have not been studied for their association with HBV infection susceptibility. In this study, we have predicted deleterious nsSNPs (M129V, A95D, and I65V) and analyzed their impact on the structure and function of the ATG5 protein through a computational approach. In addition, we selected non-coding SNPs (rs2245214, rs12212740, and rs510432) of ATG5 gene and genotyped those in 550 HBV infected patients and 250 healthy individuals. We found a strong association of the nsSNP M129V (rs34793250 (T/C), OR=3.35 (T vs. C), p=0.01) and the non-coding SNPs (rs2245214 (C/G), OR=1.30 (C vs. G), p=0.02 and rs510432 (G/A), OR=1.66 (GG vs. GA), p=0.01) with HBV infection susceptibility. Therefore, these SNPs could be utilized as biomarkers for HBV risk analysis in our population after validation. Further, the influence of these SNPs on the patients' treatment response could be investigated to develop better therapeutic strategies based on the host genotype.

Chapter 1

INTRODUCTION

Introduction

Hepatitis B is a severe liver infection which is caused by the hepatitis B virus (HBV). Initially, HBV causes a short-term infection called acute infection which gradually develops into chronic disease and later progresses to liver cirrhosis, hepatocellular carcinoma and may lead to death [1]. Globally, there are around 2 billion people infected with HBV among which 257 million people are the chronic carriers [2]. In India, there are more than 37 million HBV carriers [3] and according to the World Health Organization (WHO), every year, 1 lakh people die due to HBV related complications. Although vaccines, antiviral medications, and interferons against HBV are available, a better prediction of the risk, progression or outcome of the infection is required for better management of the disease [4-6]. Moreover, prediction of biomarkers that influence HBV treatment response also needs an investigation to develop more effective individualized therapies [7, 8]. For example, recent studies have identified various SNPs in IL28B to be associated with positive treatment response and viral clearance in HCV and HBV infected patients [9-13].

HBV is a DNA virus in the *Hepadnaviridae* family. The HBV virion is about 42 nm in diameter and consists of an outermost envelope which encloses an innermost nucleocapsid. The outer envelope is composed of Hepatitis B surface antigen (HBsAg) and the nucleocapsid is made of the Hepatitis B core antigen (HBcAg). The nucleocapsid also contains the viral polymerase and partially double-stranded DNA as its genome (3.2 kb) [14]. The short-term infection of HBV is called an acute infection during which the virus is usually cleared within first 6 months. However, if the infection lasts longer than six months, it is considered to be a chronic infection [15]. Various serologic markers exist including, HBsAg, an antibody against HBsAg (HBsAb), antibody to HBcAg (HBcAb), and IgM antibody against HBcAg, are used to determine the different phases of HBV infection (acute or chronic). The serologic information also helps to determine whether a patient is immune to HBV due to natural infection or due to prior vaccination [15].

Autophagy is a fundamental biological process that directs a broad range of physiological as well as pathological conditions in humans [16]. Autophagy has an important role in antimicrobial host defense against diverse viral infections [17]. However, HBV induces autophagy and exploit the process for its own replication [18]. Upregulation of autophagy by HBV has been demonstrated by various studies [19-21]. Studies suggest that HBV X protein

(HBx) and the small surface protein (SHBs) assist in the induction of the autophagy process. Therefore, antiviral drugs that target autophagy could be developed for the treatment of HBV infection [20, 22]. Moreover, any disruption/intervention at initial stages of this pathway could help to tackle the virus and resolve the infection.

Autophagy has been implicated in various human disorders such as cancer [23], certain neurodegenerative disorders [24], cardiomyopathies [25, 26], pancreatitis [27] and various infectious diseases [28]. Besides these, recent data suggest its involvement in other human conditions such as Asthma [29], skin disorders [30], dental diseases [31], Paget disease of bone [32], amyotrophic lateral sclerosis [33], and hereditary spastic paraparesis [34]. Various stress signals induce the process of autophagy that then initiates the formation of an isolation membrane and a phagophore. The phagophore further expands to form a double-membrane structure called an autophagosome. The maturing autophagosomes engulf cytoplasmic constituents like damaged cell organelles or degraded proteins and then fuses with the lysosome to form autolysosomes. The engulfed matter is finally degraded in the autolysosomes with the help of lysosomal enzymes [35, 36]. The process of autophagy is regulated by several autophagy-related (ATG) genes. 41 autophagy-related genes have been discovered in yeast and many of them have mammalian counterparts [37]

ATG5 is the key autophagy regulatory protein that conjugates to ATG12 and forms a complex with the multimeric protein ATG16L1. The ATG12–ATG5/ATG16L1complex acts as an E3-like enzyme, essential for the conjugation of microtubule-associated protein 1 light chain 3 (LC3) and phosphatidylethanolamine (PE) required for the formation of autophagosome [38, 39]. Liver-specific Atg5 gene knockout in transgenic mice has been shown to reduce HBV DNA levels by inhibiting autophagy [19, 21]. In another study, ATG5 levels have been found to be elevated in HBV patients. This elevated level of ATG5 may promote prolonged induction of the autophagy process which directly benefits HBV replication [40]. Recently, a number of genetic studies have focused on single nucleotide polymorphisms (SNPs) in autophagy-related genes to study their association with susceptibility to various diseases [29, 41-44]. In this context, SNPs such as rs12212740 and rs510432 of ATG5 have been linked to Asthma susceptibility and lung function [45, 46], rs573775 has been shown to be linked with Systemic lupus erythematosus [47], rs2245214 is found to be associated with Thyroid carcinoma and Paget disease of bone [32, 48].

Moreover, a recent study found a mutation E122D in ATG5 that impairs its interaction with ATG12 and leads to Ataxia development due to autophagy inhibition [49].

Host genetic factors influence HBV acquisition and play an important role in determining progression and outcome of the infection [50, 51]. Evidently, polymorphisms in immune related components such as Human leucocyte antigen (HLA), cytokines, chemokines, interferons and various immunoregulatory factors have been associated with susceptibility and outcome of HBV infection [52]. On the other hand, polymorphisms in autophagy-related genes have not been studied for their association with HBV infection until now.

Identification of disease-associated SNPs in a gene is the first step in population-based association studies. Although the experimental approach to identify disease-associated SNPs from neutral SNPs provides the best confirmation but these experimental strategies are time consuming and expensive. On the other hand, *in silico* methods provide fast, efficient and cost-effective detection and analysis of deleterious SNPs from a large number of neutral SNPs [53, 54]. Non-synonymous SNPs (nsSNPs) on the other hand, are more important as they occur in the protein coding region, which alters the encoded amino acid sequence. The substituted amino acid may affect the protein structure and function leading to disease development. Some of the common mechanisms by which an amino acid substitution can cause disease is by changing the protein stability or protein binding affinity affecting protein-protein interactions [55].

In our study, we have explored various *in silico* tools for the prediction of deleterious nsSNPs in ATG5. For this, we used, Sorting Intolerant From Tolerant (SIFT), Polymorphism Phenotype (PolyPhen2) and PredictSNP, for the prediction of nsSNPs that have a damaging impact over ATG5 protein function. Additionally, we have analyzed *ATG5* gene to find out its evolutionary connections with other species, important regulatory elements (RE), transcription factor binding sites (TFBS), post-translational modification sites (PTMs), linkage disequilibrium (LD) and its interactions with other proteins.

In order to analyze the effect of nsSNPs on the structure and function of the ATG5 protein, we followed another computational protocol where we have performed sequence-based analysis and structure-based analysis for pathogenicity, stability and conservation analysis. Also we have analyzed the structural variations of the protein complex upon mutation, computed the total energy, and predicted binding affinity change upon mutations. Therefore

in this study, we have explored polymorphisms in *ATG5* (deleterious nsSNPs and noncoding SNPs*)* and have genotyped them to find out their association with HBV susceptibility or progression.

Rationale

HBV modulates the process of autophagy in the host and exploits this process for its own replication. Any disruption/intervention at the initial stages of this pathway could help in virus replication inhibition and infection resolution. Moreover, host genetic factors influence HBV acquisition and may predispose individuals to infection. ATG5 is an essential autophagy protein that conjugates with ATG12 and ATG16L1 to form an E3 like complex that plays a pivotal role in autophagosome formation. *ATG5* deficient mice die within 1 day of birth during early neonatal starvation period, highlighting its role in regulating the process of autophagy. In this context, polymorphisms in ATG5 could be explored to find out their association with HBV infection, susceptibility or progression. Moreover, nsSNPs in ATG5 could disrupt its interactions or even destabilize its complex, which might result in varying levels of autophagy and possibly alter individual susceptibility to HBV infection. An understanding of such SNPs could help in elucidating novel mechanisms of disease pathogenesis, identification of biomarkers for risk analysis or treatment responses and could be further utilized for developing a more effective individualized therapy.

Considering these lacunae, the following are the objectives of this study.

Objective 1: To predict deleterious nsSNPs in the autophagy-related gene *ATG5.*

Objective 2: To investigate the impact of predicted nsSNPs on the ATG5 protein at sequence as well as structural levels.

Objective 3: To genotype the predicted deleterious nsSNPs as well as the selected non-coding SNPs in hepatitis B virus-infected patients and healthy individuals.

Chapter 2

REVIEW OF LITERATURE

2.1 Hepatitis B infection

Hepatitis B is a severe infection of the liver due to the hepatitis B virus (HBV). HBV-induced hepatitis is an acute (short-term) illness for a few people (<6 months), while for others; it can be chronic (long-term) infection (>6 months). The chronic Hepatitis gradually leads to the development of scars in the liver tissue leading to liver cirrhosis or development of a malignant tumour called hepatocellular carcinoma (HCC) or liver cancer [2]. The likelihood of developing chronic HBV infection depends upon the age of the host. It is more than 90 percent of infected infants (<1 year), 25–50 % of children (1-5 years), and 5% adults that develop chronic infection. Immunocompromised persons or people under some immunosuppressant therapy are always at a higher risk of developing chronic infection [15].

2.1.1 Epidemiology

Hepatitis B is a major health problem in the world. Globally, around 2 billion people have been infected with HBV [56]. According to the latest reports from the WHO, 257 million people are chronically infected with HBV in the world (Figure 2.1). Moreover, in 2015, hepatitis B resulted in 8,87,000 deaths due to the complications like liver cirrhosis and HCC [2]. On the basis of prevalence of HBV surface antigen (HBsAg; HBV endemicity), there are three regions with high (> 8% of the population carries HBsAg), intermediate (2-7%) or low endemicity (<2%). High endemic regions include the developing regions, South-East Asia, sub-Saharan of Africa, China, and Pacific islands, the Amazon basin and parts of the Middle East. South Asia, Eastern and southern Europe, Japan, Russia, and parts of South America are the intermediate endemic regions. While, North America, North Western Europe, and Australia are classified as low endemicity regions [57].

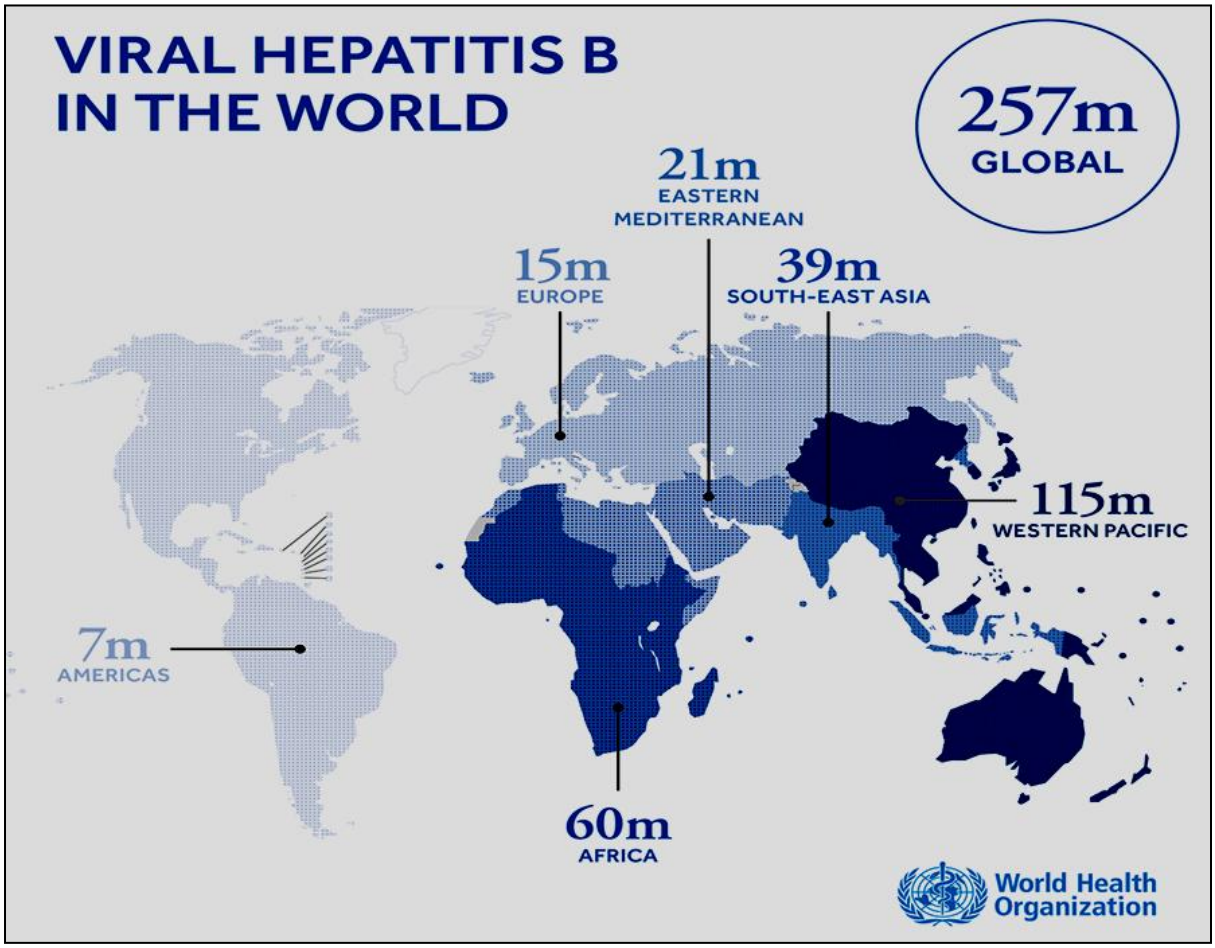

Figure 2.1: Geographical prevalence of HBV in the different regions of the world (Adapted from World health organization (WHO)) [2].

In India, more than 37 million people have been estimated to be the carriers of HBV and 15-25 % of HBsAg carriers are at risk of death due to liver cirrhosis and HCC. Moreover, 1 million infants out of 26 million infants that are born in any given year, hold a greater than normal life-long risk of presenting with a chronic infection [3].

2.1.2 Pathogenesis

The HBV is a small DNA virus that belongs to *Orthohepadnavirus* genus of *Hepadnaviridae* family. HBV is categorized into eight genotypes (A-H) and has a partially double-stranded and circular DNA (3.2 kilobase pairs) as its genome [14]. HBV infects hepatocytes in the liver leading to acute or chronic disease conditions. The actual infectious virion (Dane particle) is 42 nm in diameter. In the infected host, the virus circulates in high concentrations that is around 10^8 virions/ml of blood [58].

2.1.3 Structure of HBV

The outermost surface of the HBV virion comprised of the Hepatitis B surface antigen-HBsAg, which in part helps that forms a lipid enclosure. Next is the inner nucleocapsid that consists of HBV core antigen-HBcAg, HBV e antigen-HBeAg, DNA, and the viral polymerase, which also possesses reverse transcriptase activity.

2.1.4 HBV Genome organization

The nucleocapsid (core) of the virus consists of a partially double-stranded relaxed circular DNA (rcDNA) of 3.2 kb. The DNA has one complete coding strand called the negative strand and an incomplete non-coding strand called the positive strand. The viral polymerase is covalently attached at the 5′ end (negative strand). Furthermore, the viral genome consists of four open reading frames (S, C, P, and X) as shown in Figure 2.2. The ORF S (surface) encodes the surface envelope proteins (HBsAg; large, medium, small) and is divided into three regions, the S, pre-S1 and the pre-S2 region. The ORF C (core) encodes the e antigen (HBeAg) and structural proteins of the nucleocapsid (HBcAg). The C region has the core and pre-core regions. The ORF P (Pol) encodes the viral polymerase (DNA polymerase). The Pol region has 3 domains, namely; the terminal protein domain that has its function in the encapsidation and in the initiation of minus-strand synthesis; the reverse transcriptase (RT) domain that functions to catalyze genome synthesis; and the ribonuclease H domain that plays a role in degradation of pre-genomic RNA and in carrying out the viral replication. The ORF X encodes a multifunctional HBx protein. Studies show that the HBx protein has an important role in HBV infection progression and is also responsible for the oncogenic potential of HBV. HBsAg has various other functions in DNA repair, signal transduction events, transcriptional activation of genes and in protein degradation. However, the mechanisms of these activities of the protein are largely unknown. In addition, the HBV genome includes various other functionally important regions in the 5′ ends of the plus strand called the direct repeats (DR1 and DR2) (Figure 2.2). During the replication, these are required for the strand-specific DNA synthesis. En1 and En2 are the enhancer elements, which direct the expression of viral proteins in the liver. Also, there is a glucocorticoid-responsive element (GRE) sequence, a polyadenylation signal, and a posttranscriptional regulatory element has been described as a part of the HBV genome [14].

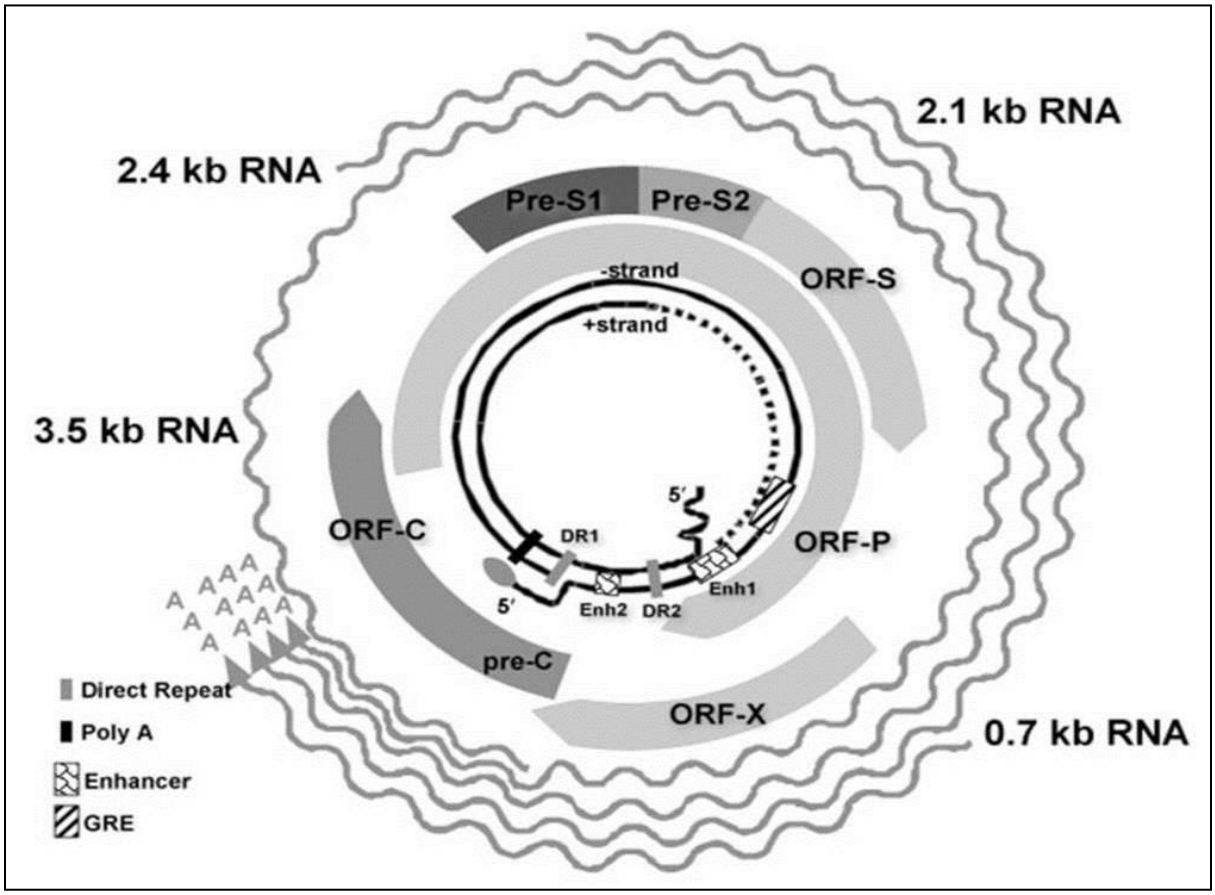

Figure 2.2: The HBV genome organization and other regulatory elements (Adapted from Liang, 2009) [14].

2.1.5 HBV Lifecycle

HBV gets transmitted from host to host by the contact with infected blood and other body fluids such. Therefore, the virus enters the liver of the host from the bloodstream. In liver tissue, the virus infects the hepatocytes through binding to its receptor. After its entry, the virus releases its genome as a relaxed circular DNA (rcDNA) at the nuclear pore. The rcDNA is then repaired into covalently closed circular DNA (cccDNA) by the viral and host cellular enzymes the nucleoplasm where it is repaired by viral and host cellular enzymes [59]. This cccDNA is then transcribed into different viral RNAs encoding the viral nucleocapsid proteins (C, pre-C), viral polymerase (P), surface proteins (large, medium, and small) and the transcriptional transacting protein (HBx). The pictorial presentation of HBV life cycle is shown in Figure 2.3.

2.1.5.1 Binding and entry

Initially, the virus particle attaches to the hepatocyte's surface by low-affinity reversible binding through its preS1 receptor. This primary binding of the virus to the hepatocyte is energy independent process that is required for productive HBV infection. After this, the virus particle binds irreversibly to an NTCP (sodium taurocholate co-transporting polypeptide) receptor on the hepatocyte's membrane. The preS1 domain of the surface protein of the virus is responsible for this binding to the receptor [60]. After the binding to the hepatocytes membrane, the virion enters the cell with the help of host factors by endocytosis.

2.1.5.2 Viral-nucleocapsid release and transportation

After the entry of the virus, the rcDNA is released into the cytoplasm (Figure 2.3). Next, the nuclear import receptors namely, importin α and β and the nuclear localization signaling (NLS) at the C-terminal of core protein interaction and help in transporting the genome (rcDNA) to the nucleus (nucleoplasm) through the nuclear pore complex (NPC) [61].

2.1.5.3 rcDNA repair-cccDNA formation

The viral rcDNA is repaired in the nucleus where the positive strand is completely synthesized with the viral polymerase utilizing a short RNA primer. After the complete synthesis, the positive and the negative strands of the DNA are ligated to form the cccDNA. This cccDNA is used as the template for transcribing different viral mRNAs.

2.1.5.4 Transcription

The transcription process is regulated by the host and the viral factors. Host factors include the various transcription factors, namely, CREB, STAT1 and STAT2, chromatin modifying enzymes namely, PCAF, HDAC1, and various hepatocyte nuclear factors. Moreover, the viral proteins such as HBx protein also participates in the transcription process by interacting with the promoter for all four ORFs for the synthesis of their respective mRNAs. The four main mRNAs are then exported to the cytoplasm again with the help of host and viral factors. Studies have demonstrated that others proteins like HBcAg and the nuclear export factor-1 (NFX1) are involved in transporting the viral pre-genomic RNA (pgRNA) to the cytoplasm [61].

2.1.5.5 Translation

The pgRNA is translated into the core protein and the viral polymerase, while, sub-genomic RNA translation occurs for the synthesis of the surface proteins (L, M, and S), as well as the multifunctional HBx regulatory protein.

2.1.5.6 Assembly of nucleocapsid

With the help of the viral core protein and polymerase, pgRNA is packaged to form the viral nucleocapsid. Once packaged, the genome is replicated by the reverse transcription of pgRNA utilizing the viral polymerase and eventually this process leads to the formation of the encapsidated, double-stranded rcDNA [62].

2.1.5.7 Nucleocapsid recycling or Nucleocapsid envelopment

The nucleocapsid containing the new viral genome (rcDNA) is reimported to the nucleus for the amplification and maintenance of the stable cccDNA pool. HBsAg levels have been shown to regulate the transport of the nucleocapsid to the nucleus with the decreased amount promoting the reimport. In addition, the nucleocapsid can move towards endoplasmic reticulum (ER) for its envelopment by the HBV envelope glycoproteins. The viral envelope proteins are embedded into the ER membrane after the process of translation. These proteins bud into the lumen of ER and are secreted as non-infectious sub-viral particles. Then again, when these surface proteins bind to the DNA-containing nucleocapsid inside the ER lumen, these interactions eventually allow the produced products to be secreted as infectious Dane particles through the endosomal sorting complex required for transport (ESCRT) pathway (Figure 2.3).

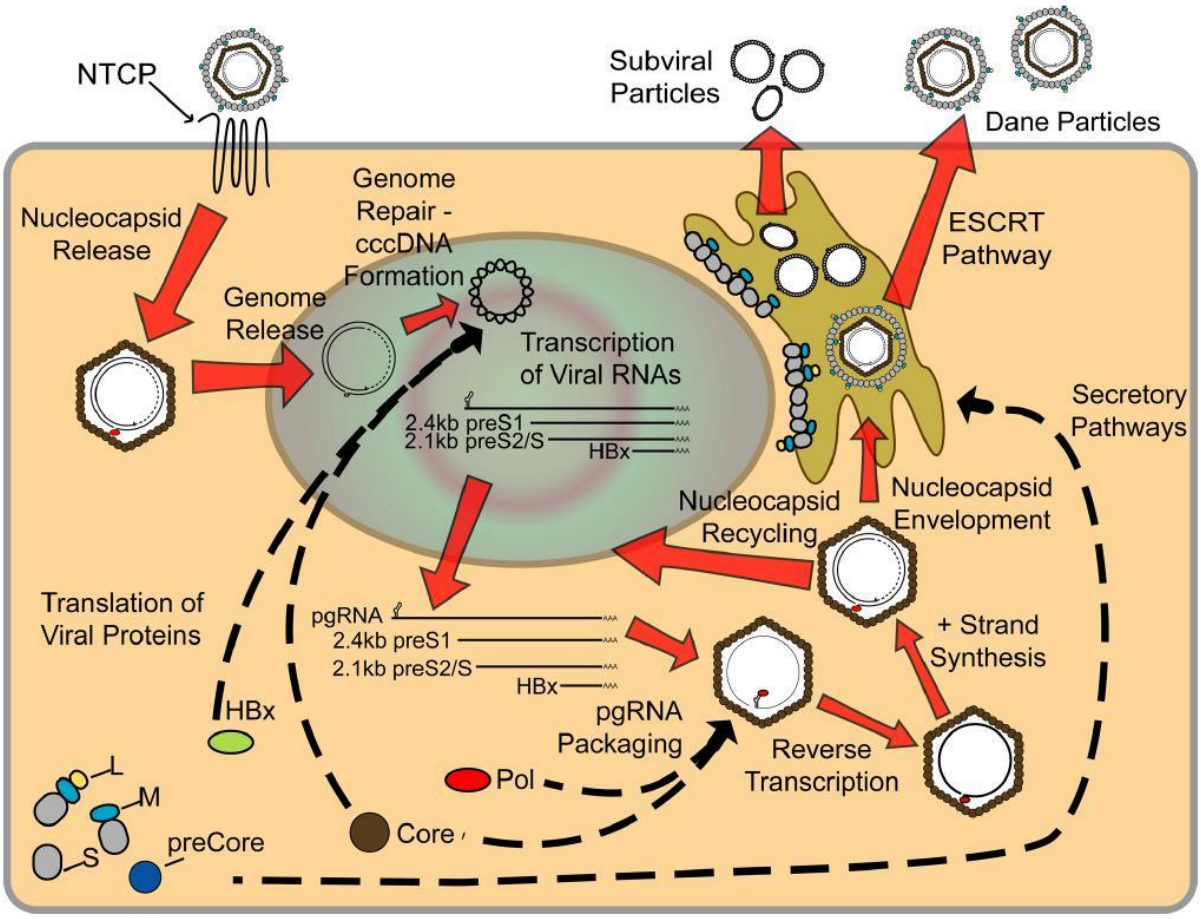

Figure 2.3: **Hepatitis B Virus life cycle (Adapted from Lamontagne et al., 2016)** [62]. In liver tissue, the virus binds to its receptor NTCP on the surface of hepatocytes. After its entry, rcDNA in the nucleocapsid of the virus is released at the nuclear pore and converted into cccDNA. This cccDNA is transcribed into various viral RNAs encoding viral nucleocapsid proteins, polymerase (P), surface proteins (L, M, and S) and transcriptional transacting proteins (X)

2.1.6 Diagnosis and Serology of HBV infection

The clinical symptoms of HBV infection much similar to the symptoms seen following infection with other viruses. Therefore, serologic testing of the patient is considered the best diagnosis of HBV infection. In the Hepatitis B serologic testing, several HBV specific antigens and antibodies are measured in the patient to identify different stages of HBV infection. Moreover, the tests also help to determine whether a patient has a risk of acquiring an infection or is already immunized to HBV due to prior infection or vaccination. The various serologic markers (or combination markers) along with the possible diagnosis and

interpretation is shown in Table 2.1.The serologic course in the patient during HBV infection is shown in Figure 2.4.

Hepatitis B surface antigen (HBsAg): The surface of HBV consists of this antigen (HBsAg). The antigen is detected in the serum of the patient infected with HBV during the acute and chronic stages. Moreover, its presence indicates active infection and therefore could be transmitting virus to the susceptible individuals. The HBV vaccine is designed using HBsAg.

Hepatitis B surface antibody (anti-HBs): This is the antibody produced against the viral surface antigen HBsAg. The presence of anti-HBs indicates immune response against the virus and hence there is the chance of patient recovery. Hence, anti-HBs is detected in the vaccinated person indicating his immunization status.

Total hepatitis B core antibody (anti-HBc): The host immune system produces this antibody against the viral core protein. The anti-HBc is detected during the acute stage of HBV infection. In addition, anti-HBc is seen in the case of a previous infection in the patient. The antibody may also be observed in the patient at the onset of symptoms. Also, the antibody persists in a patient for their lifetime.

IgM antibody to hepatitis B core antigen (IgM anti-HBc): The antibody is usually positive for the patient during the acute stage of HBV infection and is absent in the patient during the chronic stage of the infection. Therefore, its presence indicates a recent infection in the patient.

Table 2.1: The different serologic markers used for the diagnosis of HBV infection (Adapted from www.cdc.gov/hepatitis/hbv)

Serologic markers	Results	Diagnosis and Interpretation
HBsAg	Negative	Susceptible
anti-HBc	Negative	
anti-HBs	Negative	
HBsAg	Negative	Immune due to natural infection
anti-HBc	Positive	
anti-HBs	Positive	
HBsAg	Negative	Immune due to hepatitis B vaccination
anti-HBc	Negative	
anti-HBs	Positive	
HBsAg	Positive	Acutely infected
anti-HBc	Positive	
IgM anti-HBc	Positive	
anti-HBs	Negative	
HBsAg	Positive	Chronically infected
anti-HBc	Positive	
IgM anti-HBc	Negative	
anti-HBs	Negative	
HBsAg	Negative	Interpretation unclear; four possibilities:
anti-HBc	Positive	1. Resolved infection (most common)
anti-HBs	Negative	2. False-positive anti-HBc, thus susceptible
		3. "Low level" chronic infection
		4. Resolving acute infection

HBsAg: Hepatitis B surface antigen

anti-HBs: Hepatitis B surface antibody

anti-HBc: Total hepatitis B core antibody

IgM anti-HBc: IgM antibody to hepatitis B core antigen

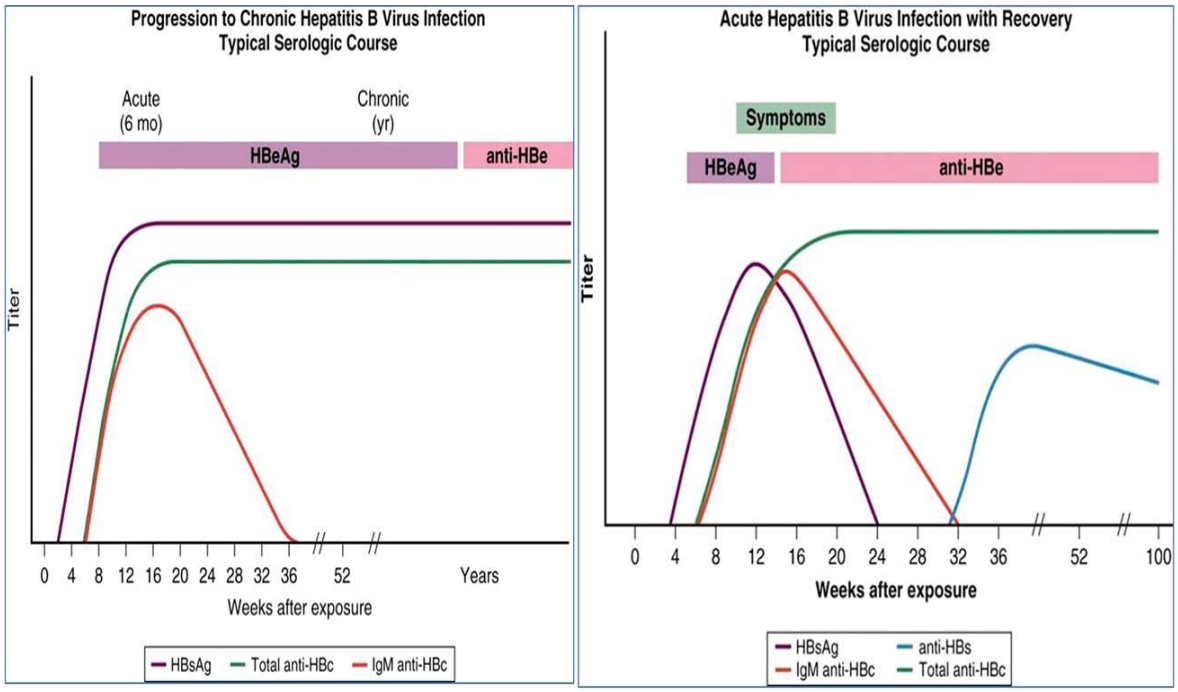

Figure 2.4 Serologic patterns during the acute and chronic stages HBV infection (adapted from (http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/hepatology/hepatitis-B/)

2.1.7 HBV transmission

HBV infection is mainly transmitted through the contact with infected blood and body fluids-semen and saliva. There is no evidence for the spread of HBV by contaminated food, water, feces or through the air.

There are three main routes of HBV transmission:-

- Perinatal Transmission

  HBV can transmit from infected mothers to the child (MTCT) during the perinatal period. MTCT is responsible for more than 33% of chronic HBV infections around the world [63]. The transmission of HBV from infected mothers to babies can occur through the placenta in the uterus, natal spread during delivery or through breast milk to the child after delivery [57].

- Sexual Transmission

  The HBV infection is considered to be a sexually transmitted disease (STD). Multiple sexual partners or sexual partners of injection drug users or prostitutes are highly susceptible to the infection [64].
- Parenteral Transmission

  HBV infection could be spread by the percutaneous route through reusing of needles, during infected blood transfusions or dialysis, acupuncture, piercing, and tattooing, sharing personal items like razors or toothbrushes. Thus blood donors and the health care workers are at high risk of getting the infection through parenteral transmission [65].

2.1.8 Treatment

Treatment of HBV is aimed at the elimination of infection, halting the process of liver damage, preventing the development of liver cancer, normalizing liver enzymes and improving symptoms.

These are following classes of treatment given to HBV patients

- Antivirals: In order to eliminate HBV by suppressing viral replication. Example: nucleoside analogs (lamivudine, entecavir and telbivudine) and nucleotide analogs (adefovir and tenofovir)
- Immune-modulators: To enhance the immune response of the host against the virus to help recover from infection. Example: interferon-alpha and pegylated interferon-alpha.

2.1.9 Prevention

Immunization against HBV through vaccination is considered to be the most effective way to prevent the infection and its complications.

- Plasma-derived vaccines: Hepavax B, Hepaccine-B
- Recombinant DNA yeast-derived or mammalian cells-derived vaccines: Hepavax-Gene, Engerix-B.

2.2 Role of Genetics in HBV infection

There are various factors that affect the risk or outcome of HBV infection, which could be categorized into, environmental factors (alcohol and aflatoxin exposure), host factors

(immunological and genetic) and viral factors (viral loads, genotype, and mutants). The host immunological response against HBV plays important role in defining the acquisition, progression, and outcome of the infection [66]. There are a number of studies that link the variations in human HLA genes with HBV infection susceptibility, progression, and outcome (Table 2.2) [51, 67-69].

Table 2.2: The list of polymorphisms in the human leukocyte antigen (HLA) class I and II genes that are associated with HBV clearance, persistence, and response to vaccination (Adapted from Chatzidaki et al., 2011) [51].

HLA complex	Association	Studied population
HLA class I		
A∗02,∗0301,∗11	Viral clearance	Adults
B∗08,∗18,∗35	HBV persistence	Adults
HLA class II		
HLA DR		
DRB1∗1	Response to HBV immunization	Infants, adults
DRB1∗03	Vertical transmission	Infants
	Response to vaccination	Infants
	Nonresponse to vaccination	Adults
DRB1∗07	Intrauterine infection	Infants
	Nonresponse to full vaccination	Adolescents
	Chronic infection	Adults
DRB1∗11	Response to immunization	Infants
	Viral persistence	Adults
DRB1∗13(01/02)	Viral clearance	Adults, children
DRB1∗15	Response to vaccination	Infants, adolescents,
HLA DQ		adults
DQB1∗02	Nonresponse to vaccination	Infants, adults
DQB1∗06	Response to vaccination	Adolescents

HBV-hepatitis B virus; HLA-human leucocyte antigen.

Moreover, polymorphisms in certain cytokines, chemokines and their receptors have also been demonstrated to influence the HBV infection outcome. In this context, a TNF-α, -238 promoter variant was shown to be significantly associated with HBV persistence and development of chronic conditions [70, 71]. Similarly, polymorphisms in Interleukin-28B (IL-28B) play an important role in hepatitis C virus clearance and prediction of treatment outcome [72]. The function of IL-28B polymorphisms in HBV infection is being explored for its association with infection susceptibility, outcome and treatment response [73]. The *IL28B* SNP rs12979860 has been shown to be strongly linked with HBV clearance in an Egyptian population [74]. Another study demonstrated a strong association of IL-28B polymorphisms with serologic response to interferon treatment in CHB patients [13]. Besides

these, IL-28B polymorphisms have also been shown to influence HBeAg seroclearance in CHB patients [75]. In another study, SNPs rs12979860, rs12980275 and rs8105790 in IL-28B, were demonstrated to be significantly related to HBV clearance in Saudi Arabian patients [76]. However, there are studies that have found no link between the IL-28 polymorphisms and HBV clearance [77-79].

Similarly, polymorphisms in other candidate genes have been studied for their role in HBV infection, including, IL10, IL19 and IL20 [80], TNFA (tumor necrosis factor alpha) [81], interferons (IFNAR1, 12IFN-g, IFN-g) [82-84], RANTES (regulated on activation, normal T cell expressed and secreted) [85], MBL2 (mannose binding lectin) [86], CCND2 (Cyclin D2) [87], VDR (vitamin D receptor) [88], CTLA4 (cytotoxic T-lymphocyte antigen 4) , estrogens [89], MTHFR (5,10-methylenetetrahydrofolate reductase) [90], TLR3 (Toll-like receptor 3) [91], TLR4, TLR5, TLR9 [92], as well as, UBE2L3 (Ubiquitin Conjugating Enzyme E2 L3) [93].

2.3 Autophagy (Macroautophagy)

Autophagy is an essential, self-degrading process that removes damaged cell organelles and proteins for cell renewal. The process plays an important role in human health and disease. At basal levels, the process of autophagy is beneficial as it maintains cellular homeostasis by eliminating and recycling damaged cell organelles and misfolded proteins. However, under stress conditions like nutrient deficiency, pathogen infection, hypoxia, radiation or anticancer drug treatment, the process gets amplified to bring about cellular adaptation and survival [35, 94, 95]. Deregulated autophagy has been demonstrated to be involved in various human ailments including liver diseases, infections, muscular disorders, cancer, neurodegenerative diseases, cardiac hypertrophy and diabetic nephropathy [28, 96-99]. The implication of autophagy in different human diseases highlights its clinical and biological significance. Autophagy could, therefore, be targeted to understand the etiology of various diseases. Moreover, studies have shown modulation of autophagy could be used as a therapeutic strategy to improve treatment response in breast cancer, and other infections [100, 101].

Autophagy appears to be regulated by a complex mechanism [102, 103]. The process of autophagy is mainly divided into four stages/phases, 1) Initiation, 2) Elongation, 3) Fusion and 4) Degradation. Various stress signals (mentioned above) induce and initiate the process with the formation of an isolation membrane. The isolation membrane further expands to

form phagophore. The phagophore membrane further elongates to form a double-membrane structure called an autophagosome. Next, the autophagosome while maturing, engulfs various cytoplasmic constituents like old cell organelles, damaged proteins, pathogens, etc, and fuses with the lysosome to form autolysosomes. The engulfed matter is degraded finally in the autolysosomes with the help of lysosomal enzymes. Numerous autophagy-related (ATG) genes regulate the various stages of the process of autophagy. Until now, 41 ATG genes have been discovered. The *ATG5* gene is an important gene that encodes for an ATG5 protein which regulates the elongation of the autophagosome. ATG5 covalently joins with ATG12 to form an ATG5-ATG12 conjugate, which further interacts non-covalently with ATG16L1 to form ATG12-ATG5/ATG16L1 complex. This ATG12-ATG5/ATG16L1 complex acts like E3 enzyme that promotes ATG8-phosphatidylethanolamine (PE) formation and autophagosome elongation.

2.3.1 Autophagy and HBV

It has been shown that HBV infection can induce autophagy as demonstrated by various studies [19-21]. Moreover, these studies demonstrate that the virus induces autophagy for its own replication and hence could be targeted to tackle the infection [19]. In another study, it was demonstrated that the HBV X protein (HBx) binds to PI3KC3 leading to the production of a high level of phosphatidylinositol-3-phosphate (PtdIns (3)P) in cells. This ultimately leads to an increased number of autophagic vacuoles, autophagosomes and autolysosomes [104]. Moreover, different genotypes of HBV have been shown to have different effects on autophagy [105]. Also, HBx has been shown to induce autophagy by activating death-associated protein kinase (DAPK) in a beclin1 dependent manner [106]. HBV has also been demonstrated to induce autophagy by another mechanism that is through its small surface protein (SHB). In this mechanism, HBV utilizes its various structural proteins to alter the process of autophagy. The SHBs induce autophagy by triggering ER stress, which in turn activates autophagy through unfolded protein responses (UPR) [107]. Additionally, HBV infection was demonstrated to utilize autophagy for its envelopment [20]. Therefore, the most specific mechanism used by HBV for its replication via autophagy induction is remains unclear. However, the involvement of autophagy in DNA replication and envelopment is believed to contribute enhanced viral replication.

The involvement of autophagosomes in viral replication in other viruses has also been demonstrated in various studies. HBV and HCV induce the formation of autophagosomes but prevent their fusion with lysosomes to inhibit degradation of the enclosed constituents. This mechanism helps the virus to replicate inside or on the surface of autophagosomes without being degraded or eliminated [108]. Figure 2.5 is a general representation of how HBV and HCV exploit the process of autophagy to promote their replication.

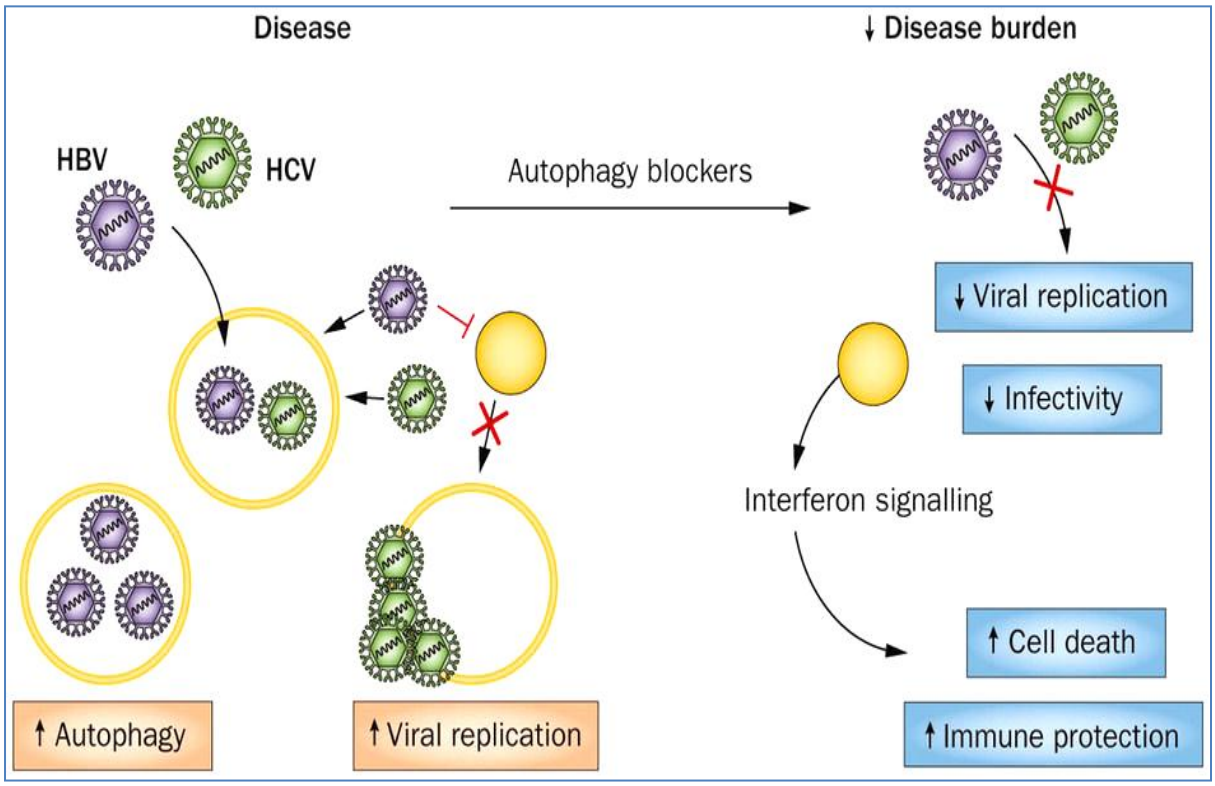

Figure 2.5: Autophagy exploitation by HBV (Adapted from Schneider and Cuervo, 2014) [108]. HBV and HCV take advantage of the process of autophagy by disrupting the fusion of lysosomes and autophagosomes in order to prevent their degradation in autophagolysosomes. However, with the use of autophagy blockers, viral replication and infectivity could be hampered and that would ultimately reduce the disease burden

2.4 Autophagy-related gene 5 (*ATG5*): An essential autophagy regulator

*ATG*5gene is located at the long arm of chromosome number 6 (6q21). The gene is 141.34 kb in size with 10 exons. This gene encodes autophagy protein 5, ATG5 which is 275 amino acids long. ATG5 (Lys 149) binds covalently to ATG12(C-terminal Gly 186) to form a functional ATG12~AATG5 conjugate with the help of the enzymatic reactions catalyzed by ATG7 and ATG10. Furthermore, the ATG12~ATG5 conjugate binds noncovalently to

ATG16L1 and forms a complex (ATG12~ATG5-ATG16L1) that acts as an E3 ligase facilitating an ATG8-PE (phosphatidylethanolamine) conjugation reaction required for autophagosome expansion [109-111]. ATG12 or ATG5-ATG16L1 alone do not function like an E3 enzyme for ATG8-PE ligation. Therefore ATG12~ATG5-ATG16L1 complex formation is essential for autophagosome biogenesis [111]. Moreover, in yeast as well as in mammalian cells, the inactivation of the ATG12 conjugation system leads to more dysfunctional autophagosome formation when compared to the effect of an inactivated ATG8 conjugation system, suggesting the key role that ATG12 plays in this system [39]. Also, in the autophagy process, the ATG12~ATG5 conjugate is considered to be the only E3-like enzyme for LC3 lipidation, in contrast to the various E3 ligases found in the ubiquitin-proteasome pathway [110]. Therefore, ATG5 has an indispensable role in accelerating the formation of the ATG12-ATG5 conjugate that is critical for the autophagosome formation.

ATG5 deficient mice die within 1 day of birth during the early neonatal starvation period highlighting the importance of the process of autophagy [112]. Another study demonstrated that mosaic deletion of *ATG5* in mice leads to the development of multiple benign tumors in the liver [113]. While podocyte-specific deletion of *ATG5* leads to glomerulosclerosis, accumulation of ubiquitin, p62 and mitochondria in aging mice revealing the influence of autophagy over glomerulopathy and mitochondrial maintenance. ATG5 is also essential for mitochondrial maintenance in the parasite *Leishmania*. Leishmania mutants lacking ATG5 were not able to differentiate normally and developed dysfunctional mitochondrion that resulted in reduced virulence of the mutants [114].

Autophagic machinery is also involved in the processing and presentation of intracellular antigens for the MHC II by dendritic cell [115, 116]. A study provided evidence for the essential requirement of *ATG5* in antigen presentation by dendritic cells in vivo [117]. Therefore, the autophagy machinery plays an important role in antigen presentation and hence, could be utilized for better vaccine design [118]. Furthermore, *ATG5* has shown to influence B cell development and maintenance [119]. Moreover, deletion of *ATG5* may result in abnormalities in T-cell development and proliferation [120, 121]. ATG5 deletion also interferes with normal adipocyte differentiation, which has been demonstrated both in a cell model as well as *in vivo* [122].

Autophagy also has a functional influence on the preimplantation development of embryos in mammals. The Oocyte-specific ATG5 knockout in mice leads to the disruption in the development of embryos, which were fertilized by ATG5-null sperm; whereas no disruption was observed in the development of embryos which were fertilized by wild-type sperm [123]. Also, another study demonstrates an important role of ATG5 in neurogenesis and organogenesis through functional analysis in zebrafish. Moreover, the study provides evidence for the essential requirement of ATG5 over ATG12 in the formation of ATG12–ATG5 conjugate in zebrafish development [124]. ATG5 also influences primary ciliogenesis by eliminating accumulated OFD1 (oral-facial-digital syndrome 1) from centriolar satellites through the autophagic pathway [125].

In mouse embryos, ATG5 is shown to regulate neural stem /progenitor cell differentiation [126]. Mice with ATG5 deficient neurons, exhibit degeneration that leads to the buildup of inclusion bodies in the cytoplasm [127]. On the other side, loss of ATG5 in Purkinje cells causes axonal swelling due to neurodegeneration in mice [128]. Besides this, another study demonstrates ATG5 to be essential for astrocyte differentiation in the developing mouse cortex through the modulation of STAT3 activity [129]. Elevated ATG5 levels also impaired neuronal development in mice [130].

Accumulating data indicates that ATG5 expression directly influences autophagy levels in mice, as the systemic deletion of ATG5 leads to complete disruption of the process [112]. Also, overexpression of ATG5 in mice has been shown to enhance autophagic activity [131]. Likewise, ATG5 overexpression is also linked to the pathogenesis of autoimmune demyelination, multiple sclerosis and HBV infection [132]. In addition to this point, ATG5 expression was also found to be elevated in macrophages of the spleen and kidney and in peripheral blood mononuclear cells (PBMCs) in patients with Systemic lupus erythematosus (SLE) [132]. In contrast, ATG5 expression level is found to be strongly down-regulated in patients with colorectal cancer [133].

A recent study demonstrates the prognostic significance of ATG5 expression levels in predicting favorable disease-free survival in patients with breast cancer [134]. Moreover, ATG5 knockdown reduces the oxidative stress (basal as well as drug-induced) due to camptothecin (CPT) treatment in osteosarcoma (OS) cells highlighting the effect of autophagy blockage on anticancer drug efficacy [135, 136]. Therefore, the process of

autophagy plays a significant role in health and human diseases and it could be modulated by altering ATG5 expression for therapeutic management/treatment of various human disorders.

The various roles of ATG5 have been pictorially represented in Figure 2.6.

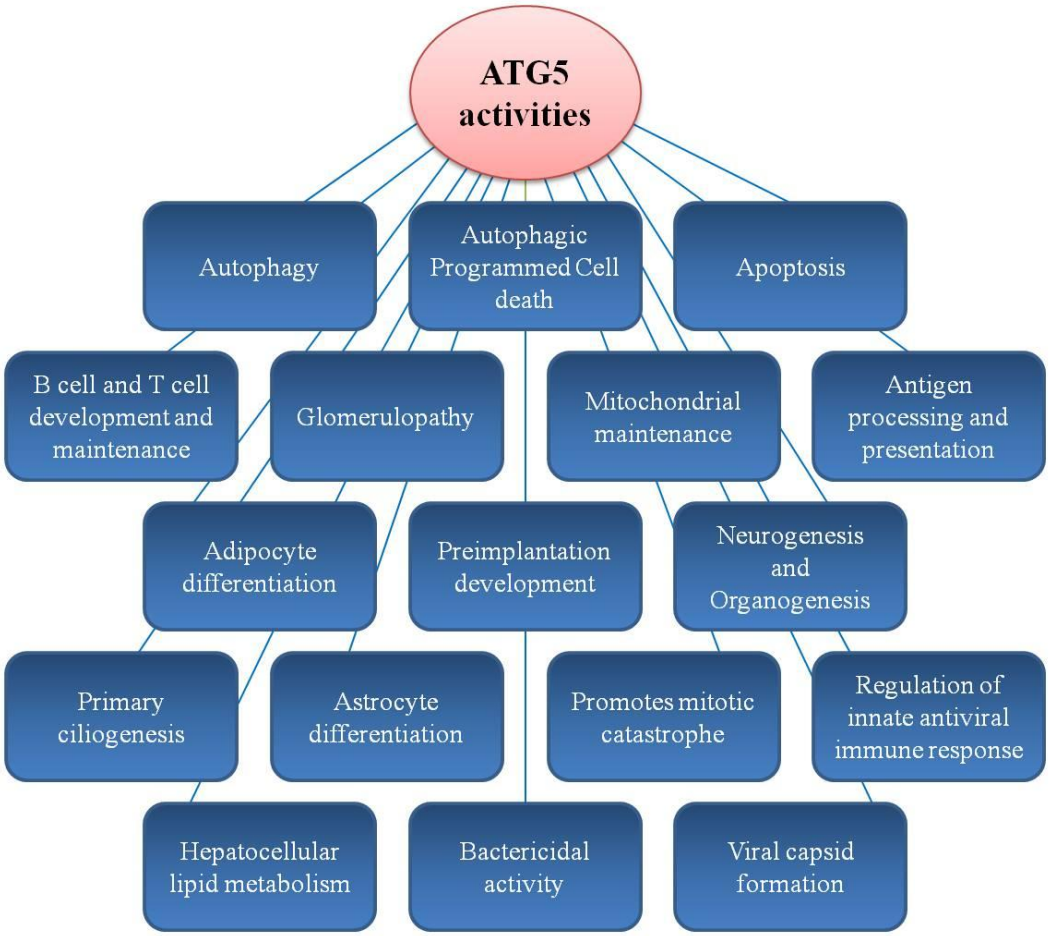

Figure 2.6: The figure represents the essential roles of ATG5 in various cellular processes highlighting its biological significance.

2.4.1 Autophagy-independent roles of ATG5

The autophagy genes code for autophagy-related proteins that act, coordinate and regulate the autophagy process at various phases. Interestingly, these proteins participate in additional cellular processes like cell survival and apoptosis, cell trafficking and signaling, and the process of transcription and translation [137]. In addition, a very recent study demonstrates the role of autophagy-related genes in fungal stress response, asexual development and virulence [138].

Likewise, ATG5 which was considered to be a cytosolic protein plays important role in the nucleus of a cell too. There is an evidence for the presence of ATG5 in the nucleus where it interacts with BIRC5/survivin (an apoptosis inhibitor) in response to DNA-damaging agents (anti-cancer drugs) and leads to mitotic catastrophe independent of autophagy [139, 140]. These dual roles of ATG5 in the cytoplasm, as well as the nucleus, emphasize its potential as a therapeutic target and thus should be considered when developing new anticancer treatments. ATG5 also has an essential role in providing cellular immunity to and clearance of intracellular pathogens like *Listeria monocytogenes* and *Toxoplasma gondii*, through GTPase trafficking independent of autophagy [141]. Moreover, ATG5 also function to protect against *Mycobacterium tuberculosis* (Mtb), by preventing (polymorphic mononuclear cells, PMN)-mediated immunopathology, highlighting the autophagy-independent contribution of ATG5 in controlling infection [142]. In contrast to this bactericidal activity, the conjugate ATG5-ATG12 has been shown to impede innate antiviral immune responses promoting viral replication in cells [143]. Furthermore, ATG5 after calpain-mediated cleavage increases the susceptibility of the cell to apoptosis establishing a link between the process of autophagy and apoptosis [144, 145]. Such additional activities of ATG5 emphasize its biological significance.

2.4.2 ATG5 interaction partners and its implications

The primary role of ATG5 is to conjugate with ATG12 and ATG16L1 to form a complex that behaves like an E3 enzyme and facilitate the attachment of phosphatidylethanolamine to the C-terminus of Atg8, which is the most crucial phase for autophagosome formation [146]. Different interaction partners such as TECPR1 and ATG3 have also been reported to interact with ATG5 and participate in autophagosome maturation process [147-149]. The interaction of ATG5 with BIRC5/survivin, which has already been discussed is a major finding as it exposed the presence of ATG5 in the nucleus, which was earlier considered to be a cytosolic protein only. Such interactions of ATG5 could be modulated and studied to provide better understanding of the various associated molecular links [139].

Recently, RACK1 (Receptor Activated C-Kinase 1) protein was found to be another interacting partner of ATG5 that helps to regulate the autophagy pathway [150]. RACK1 interacts with ATG5 and induces autophagy during starvation conditions and any hindrance in their interaction leads to autophagy inhibition highlighting the importance of their

interaction [150]. Another protein, Caveolin-1 (Cav-1), interacts with ATG5 and ATG12 and even their active ATG5-ATG12 conjugate to suppress/deactivate the functional activities of ATG5, and its conjugate ATG5-ATG12. This interaction has been demonstrated in lung epithelial cells and overall, results in autophagy inhibition [151]. Therefore, the Cav-1 and ATG5 interaction regulate the autophagy process and hence it could be modulated for better management of various lung disorders.

Moreover, the transcription and phosphorylation of ATG5 by p73 and p38 respectively, regulate the process of autophagy and are eventually involved in the maintenance of homeostasis in organisms following environmental stress. As observed in the studies, lack of p73 inhibits transactivation of ATG5 leads to disturbed hepatocellular lipid metabolism in mice explaining the significance p73-ATG5 axis in regulating lipid metabolism [152]. Similarly, another study revealed the role of p38 in controlling ATG5 activity through phosphorylation and consequently as a negative regulator of the maturation of autophagosome and hence autophagy [153]. Likewise, there are various important regulatory elements, transcription factors and their binding sites, and post-translational modification sites in ATG5 that may influence its activity. Such regulatory signatures of ATG5 could be targeted to alter its expression/activity for the treatment of different ailments [154].

Furthermore, ATG5 interacts with FADD (Fas-associated protein with death domain), leads to autophagic cell death by inducing IFN-gamma revealing the dual role of ATG5 in autophagy as well as in autophagic inducing cell death activity [155]. Besides this, ATG5 interaction with a protein, Rab33B, is required for HBV naked capsid formation and releases revealing the involvement of the process of autophagy in viral capsid biogenesis as well as exposing the non-autophagic function of ATG5 [156]. In the case of HCV, ATG5 interacts with the viral RNA polymerase (NS5B) during the onset of infection favoring viral replication [157]. Moreover, the ATG5-ATG12 conjugate is also shown to be involved in the negative regulation of the type I IFN production pathway by interacting with the retinoic acid-inducible gene I (RIG-I) and the IFN-promoter stimulator 1 (IPS-1), which results in obstruction of innate immune responses and consequently the promotion of viral replication [143, 158, 159]. Therefore, these studies highlight the significance and medical implications of ATG5 in viral infections and expose the various targets that could be exploited to develop new treatment strategies to overcome increasing viral disease burden.

The various interaction partners of ATG5 have been pictorially represented in Figure 2.7 along with their associated molecular pathways to emphasize the possible role for the key autophagy gene ATG5 in the cell and how viruses might manipulate this process to expose potential targets for modulation and therapeutic intervention.

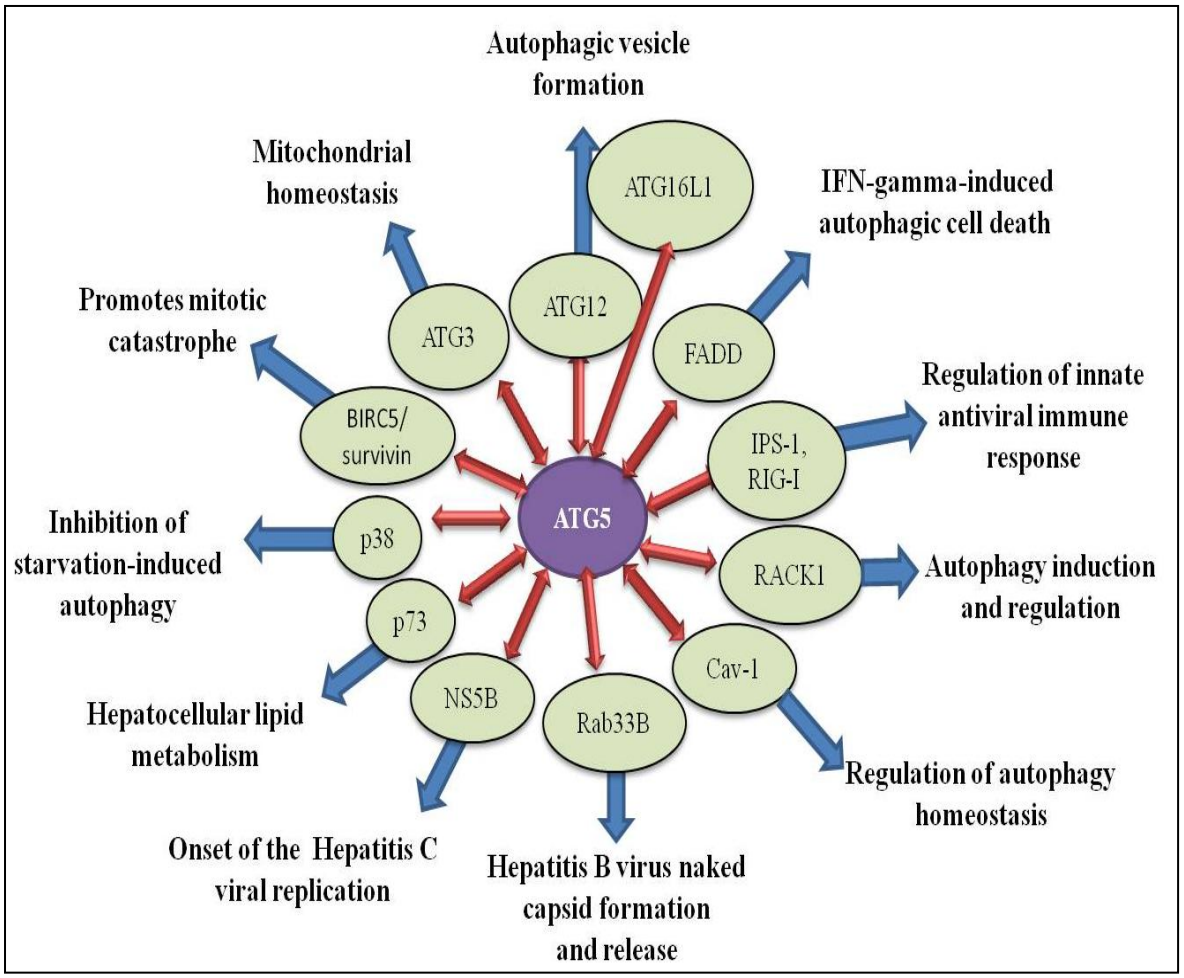

Figure 2.7: Different interaction partners of ATG5 required for accomplishment of various essential functions and pathways implicated in various human disorders.

2.5 Single Nucleotide Polymorphisms (SNPs) and its implications

Single nucleotide polymorphism (SNP) is a variation of a single nucleotide base at a specific locus in DNA. The possible nucleotide type on that position of SNP is called 'allele'. The polymorphism may have two (bi-allelic), three (tri-allelic) or four alleles (tetra-allelic). The bi-allelic polymorphism is most common with the minor allele frequency more than 1% whereas the tri-allelic and tetra-allelic polymorphisms are very rarely found. The most

frequent allele at the SNP position is called the 'wild-type' allele while the second most frequent is called the 'mutant' allele [160].

SNPs are the type of genetic variation which is most commonly found, comprising nearly $1/1000^{th}$ of the average human genome (3.3×10^9 bp). There about 10 million SNPs found in the human genome and on an average, a SNP occurs once in every 300 nucleotides.

The nucleotide base substitution in a SNP can be either a transition (between purines (A>G) or between pyrimidines (C>T) or Transversion (between purine and a pyrimidine).

The SNPs can be found in the coding region (cSNPs) or the non-coding region (intronic SNPs-iSNPs). SNPs in the coding region can be of two types: Synonymous SNPs (sSNPs) and non-synonymous SNPs (nsSNPs). For sSNPs, the base substitution does not lead to a change in an amino acid or protein sequence. Therefore, these are also called silent mutations. nsSNPs on the other hand, lead to changes in amino acids resulting in the changes to the protein; they are also called missense mutations. Another type of nsSNP is called non-sense mutation in which the substituted amino acid results into a termination codon [161].

2.5.1 SNP detection techniques

There are around 1 million SNPs which have been identified to date. For SNP detection, a set of biochemical reactions are carried out that determines the exact location of the suspected SNP and thus to find out its identity. There are four common methods used for SNP detection/identification:

2.5.1.1 Single-strand conformation polymorphisms (SSCPs) method

In this method, the DNA containing the target SNP is amplified. The amplified product is denatured and then run on a non-denaturing polyacrylamide gel. The fragments undergo various conformational changes during the run and would achieve a particular secondary structure based on their sequence. The fragments containing the target SNP are easily differentiated on the basis of their different migration pattern. Furthermore, the confirmation is performed by sequencing. SSCP is a simple technique with 75-95% success rate (variable). This technique is widely used, is labor intensive and has low throughput [162, 163].

2.5.1.2 Heteroduplex analysis

The Heteroduplex analysis depends upon the recognition of the duplexes formed after the reannealing of the denatured PCR product. The Heteroduplex could be easily analyzed over a gel, which can resolve the conformational changes due to sequence variations. These sequence variations lead to conformational changes in the double helix that directly affect its mobility over gel electrophoresis. This method of SNP detection is simple, low cost with a high success rate (95-100%) [164, 165].

2.5.1.3 Direct DNA sequencing

Through the better availability and improvement of DNA polymerases and other sequencing reagents, DNA sequencing is considered to be a superior strategy for SNP detection. By the means of automated sequencing machines, DNA sequencing could be achieved automatically. Through one experiment much of the information about the target SNP including its type, location, etc could be easily obtained. These advantages for this method over other methods, overcomes the limitations of high-quality samples (PCR products) and expensive instrumentation [166].

2.5.1.4 Variant detector arrays (VDAs)

The VDA technology, SNP identification is done by hybridization of the PCR products to the oligonucleotides, which are arrayed on a silicon glass DNA chip. The difference between the hybridization strength of matched and unmatched oligonucleotides is measured for target SNP identification. A lot of DNA sequences could be rapidly scanned using VDA technique [164].

2.5.2 Genotyping SNPs

Detection of the genotype of known individual SNPs is called Genotyping. SNPs genotyping is carried out by using two major methods, the traditional methods or the high throughput methods. The traditional methods include gel-based methods based on enzymatic digestion or PCR, for example: Restriction fragment length polymorphism (RFLP) and amplification refractory mutation system (ARMS). On the other hand, there are high throughput methods namely, allele discrimination methods (Allele-Specific Hybridization, Allele-Specific Single-Base Primer Extension), High-throughput assay chemistry (Flap endonuclease-FEN discrimination, Oligonucleotides ligation).

2.5.2.1 Traditional Methods

Traditional methods include the gel-based methods for SNP genotyping**.** RFLP and ARMS are the two traditional methods for SNP genotyping discussed below.

2.5.2.1.1 Polymerase chain reaction-Restriction fragment length polymorphism (PCR-RFLP)

PCR-RFLP is one of the first approaches used for SNP genotyping. In this method, the PCR is carried out to amplify the target DNA sequence containing the SNP. Then, with the help of restriction enzymes, the amplified product is digested. The digested fragments are resolved by agarose gel electrophoresis. The banding pattern is observed for each sample, which is specific to the type of genotype of the individual at a particular locus [167].

2.5.2.1.2 Amplification refractory mutation system (ARMS)

ARMS method is a PCR-based method for genotyping SNPs. It is also known as allele-specific Polymerase chain reaction (ASPCR). In this method, four different primers are generated (two forward primers and two reverse primers). The primers are designed to amply the target sequence containing the SNP but amplification occurs only when the target allele is present in the sequence. Moreover, the resultant PCR products for both the alleles would vary in their lengths and therefore could be easily resolved by electrophoresis [168, 169].

2.5.2.2 High throughput techniques

High throughput techniques include SNPs genotyping methods aimed at large scale association studies that include a number of target SNPs and a large population set. The genotyping methods are based on the most common high throughput techniques and they are discussed below.

2.5.2.2.1 Allele-specific hybridization (ASH)

For this method, two probes (oligonucleotides) are designed specific to the two different alleles in the target sequence. These probes bind to the target sequence only when there is a perfect match. During the reaction (under optimum conditions) even a single mismatch prevents the annealing of the allelic probe to the target sequence. As there is no use of restriction enzymes, the hybridization method is considered to be the simplest method of SNP

genotyping [166]. Examples of high throughput techniques based on ASH include the 5' nuclease or the TaqMan assay, enhanced version of ASH - dynamic allele-specific hybridization (DASH), molecular beacon probes and the scorpion analysis [170].

2.5.2.2.2 Enzymatic cleavage- Flap endonuclease (FEN) discrimination

Unlike RFLP, the cleavage-based assay known as the Invader assay has high throughput potential for SNP genotyping. The Invader assay is an enzyme based method to detect SNPs with high specificity. The method utilizes two oligonucleotide probes (the invader probe and the allele-specific probe) and an enzyme flap endonuclease (FEN). The allele-specific probe and the invader probe hybridize with the target sequence in the presence of the SNP to form a tripartite structure that is then recognized and cleaved by a FEN enzyme called Cleavase. However, no cleavage occurs in the absence of a SNP. The cleavage can be detected by fluorescence techniques or by mass spectrometry [171]. Other detection systems for high throughput SNP genotyping technologies include capillary electrophoresis, DNA arrays and pyrosequencing [166].

2.5.3 Applications of SNPs

SNPs are the most abundant genetic polymorphisms that can affect an individual's susceptibility to different diseases or infection with different pathogens. SNPs also influence the effect of various drugs or vaccines in an individual. With the advancement in SNP detection/genotyping technology and its accessibility, more SNPs are being linked to human diseases. SNPs play important role in various aspects of biomedical research and pharmacogenomics discussed in the following section.

2.5.3.1 Biomedical research

Various genetic association studies are helpful in predicting the link of a genetic variant with the disease risk, progression, outcome or even treatment response.

2.5.3.2 Genome-wide association study

In general, a genome-wide association study (GWAS), mainly explores associations between the SNPs and major human diseases. In GWAS, the entire genome is investigated in order to identify SNPs or other DNA variants associated with a disease. The findings from a GWAS could elucidate the pathophysiology of various diseases and facilitate the development of

personalized medicine in which a patient is given treatment based on his genotype. For example, a GWAS study identified a polymorphism (3 kb upstream of IL28B) to be significantly associated with treatment response (PegIFN and RBV) in patients suffering from chronic HCV (genotype 1) infection [9, 172].

2.5.3.3 Candidate gene approach

In order to conduct the genetic association study, the candidate gene approach targets pre-specified genes that are linked to the disease. Therefore, such SNPs in the selected genes are genotyped to find out their association with the disease susceptibility, progression or outcome in the patients [173, 174]. Such polymorphisms in the selected genes are hypothesized to increase an individuals' risk of developing the disease or in protecting the patient from the disease. Moreover, the susceptible genetic variants could be further investigated to study their role in treatment/drug response [175, 176].

2.5.3.4 Linkage disequilibrium analysis

Linkage disequilibrium (LD) is the correlation of SNPs that are positioned in proximity to each other. LD analysis of the genome provides information the population history, the breeding system and the pattern of the geographic subdivision. LD of a particular pair of loci or in a genomic region also reveals the history of natural selection, gene conversion, mutation and other forces that cause gene-frequency evolution. Overall the LD can provide information about the evolutionary history that could be used as a basis for association studies and for mapping genes in humans and in other species [177]. LD could also be used to locate genes that affect quantitative traits (QTL) [178]

2.5.3.5 Forensics

Due to various advantages of SNPs like their abundance in the genome, low mutation rate, the availability of easier and faster methods of detection they represent excellent markers in forensic research. Moreover, small DNA samples could be utilized for a forensic investigation [179].

2.6 *ATG5* polymorphisms and various human disorders

Recent genome wide association studies reveal genetic polymorphism within autophagy related genes to be convincingly associated with human diseases like Crohn's disease [41, 180, 181], autoimmune disease like systemic lupus erythematosus (SLE) [182], infectious diseases [42, 183, 184] neurodegenerative diseases [185], allergy [186, 187], psoriasis [43], Paget disease of bone [32], Behçet's Disease and VKH Syndrome [188], nonalcoholic fatty liver disease (NAFLD) [189], Huntington's disease [190] and cancer [44, 191-194]. These studies add to our understanding of the genetic basis of various human diseases by the identification of certain variants in genes, as well as the involvement of a biological pathway like autophagy in disease pathogenesis, which could be targeted to develop a novel treatment strategy or to enhance existing treatment responses in patients.

There are a number of studies that demonstrate the association of polymorphisms of a crucial gene like ATG5 with the susceptibility to various human disorders. In a recent study, 106774459T>A variant has been identified in a Parkinson's disease patient, which significantly enhanced ATG5 gene promoter activity. ATG5 expression was also observed to be elevated in those patients [125]. Moreover, the SNPs rs12212740 and rs510432 in ATG5 have been linked to asthma pathogenesis. The expression levels of ATG5 were also found to be elevated in asthma patients [46, 186]. Recent studies demonstrate the association of SNP rs2245214 with increased susceptibility to thyroid carcinoma [48] and Paget disease of bone [32]. Moreover, the SNP which has previously been observed to be associated with SLE [47] has recently been shown to be linked to Behçet's disease [188].

Recently, a study reveals the involvement of autophagy in the pathogenesis of Neuromyelitis Optica (NMO) by demonstrating the association of a genetic variant in ATG5 rs548234 with the disease. However, the study finds no association of this variant with multiple sclerosis susceptibility/pathogenesis [195]. The same SNP rs548234 at PRDM1-ATG5 intergenic region has also been linked to SLE in the Chinese population [196]. Table 2.3 lists the genetic polymorphisms in ATG5 studied to be linked with different diseases in human.

More recently, Kim et al., [49], identified a mutation (E122D) in ATG5 that impairs its conjugation with ATG12 leading to autophagy inhibition and subsequent development of ataxia highlighting the impact of mutations on protein function and ultimately disease

manifestation. Overall, these genetic studies highlight the therapeutic importance of autophagy which could be genetically modulated by targeting ATG5.

Table 2.3: List of ATG5 genetic variants/single nucleotide polymorphisms associated with different diseases.

Gene variant/ Polymorphism	Disease association	Reference
106774459T>A	Parkinson's Disease	[197]
rs12212740	Asthma	[46]
rs573775	Behcet's disease	[188]
rs510432	Asthma	[186]
rs573775	Systemic Lupus Erythematosus (SLE)	[47]
rs2245214	Thyroid carcinoma	[48]
rs2245214	Paget disease of bone (PDB)	[32]
rs548234	Neuromyelitis Optica	[195]
PRDM1-ATG5 rs548234	Systemic Lupus Erythematosus (SLE)	[196]

2.7 Computational analysis

SNPs within the coding region are considered to be important as they are likely to cause an amino acid substitution that could have a deleterious or damaging impact over the protein structure and function leading to disease development. In order to predict such nsSNPs, a number of computational tools are available.

2.7.1 Databases for data retrieval

Through advancement in high-throughput technology and next-generation sequencing techniques a huge amount of human genomic information has been generated. There are a number of resources from which desired data could be retrieved for further research and analysis. List of databases which could be explored are shown in the Table 2.4.

Table 2.4: List of databases for retrieval of genomic, SNP or nsSNP data

Database	Website
NCBI	https://www.ncbi.nlm.nih.gov/
dbSNP	https://www.ncbi.nlm.nih.gov/projects/SNP/
UniPROT/SWISS-PROT	http://www.uniprot.org/
RCSB PDB	https://www.rcsb.org/pdb/home/home.do
Ensembl	http://www.ensembl.org/index.html
OMIM	http://www.omim.org/
HGMD	http://www.hgmd.cf.ac.uk/ac/index.php
HGVbase	http://hgvbase.cgb.ki.se
PMD	http://pmd.ddbj.nig.ac.jp
dbNSFP	http://sites.google.com/site/jpopgen/dbNSFP

2.7.2 Tools for prediction of post-translation modification sites (PTM) sites

There are various web-based tools available for the prediction of post-translation modification sites (PTMs) for phosphorylation, ubiquitinylation, and palmitoylation and Calpain cleavage. The list of tools used in our study are shown in Table 2.5.

Table 2.5: List of tools for prediction of post-translational modification (PTM) sites

Tool	Website	Feature/methods used
NetPhos	http://www.cbs.dtu.dk/services/NetPhos/	Artificial Neural Network (ANN) based method
CSS-PALM	http://csspalm.biocuckoo.org/online.php	Clustering and scoring strategy (CSS) algorithm
BDM-PUB	http://bdmpub.biocuckoo.org/	Bayesian Discriminant Method.
GPS-CCD	http://ccd.biocuckoo.org/.	Matrix Mutation (MaM), GPS 2.0 algorithm.

2.7.3 Tools for predicting the functional impact of nsSNPs

As nsSNPs result in a change in the encoded amino acid, these can affect the human physiology in many ways. The substituted amino acid may inactivate enzymes by disrupting their functional sites or they may alter splice sites leading to the production of defective gene products [198]. Moreover, nsSNPs may lead to alterations in protein structure, stability, folding and binding affinity [199, 200]. Overall, they may result in loss of functional properties of the protein leading to the development of a disease. There are several computational approaches to predict the function of nsSNPs. The list of tools for the analysis of the effect of nsSNPs is shown in Table 2.6.

Table 2.6: List of computational tools for the analysis of the impact of nsSNPs on protein structure and function

Tool	Website	Features/Methods used for prediction
SIFT	http://sift.jcvi.org/	Statistical method using PSSM with Dirichlet priors
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/	Naïve Bayes approach coupled with entropy-based discretization
PredictSNP	https://loschmidt.chemi.muni.cz/predictsnp/	Machine learning methods
MutPred	http://mutpred1.mutdb.org/	Support Vector Machine and Random Forest
I-Mutant2.0	http://folding.biofold.org/i-mutant/i-mutant2.0.html	Support Vector Machine
iPTREE-STAB	http://203.64.84.190:8080/IPTREEr/iptree.htm.	Rule-Based Decision Tree
WET-STAB	http://bioinformatics.myweb.hinet.net/wetstab.htm	Weighted decision table method (WET)
DUET	http://biosig.unimelb.edu.au/duet/	Support Vector Machines (SVM)
CUPSTAT	http://cupsat.tu-bs.de/	Use of mean force potentials and stepwise multiple regression to develop prediction model
SRide	http://sride.enzim.hu/	Prediction is based on (A) surrounding hydrophobicity (B) a quantitative measure of the number of long-range residue–residue contacts, (C) stabilization centers and (D) conservation score.
The ConSurf server	http://consurf.tau.ac.il/2016/	Calculation of conservation scores using empirical Bayesian method.
PDBsum	http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html	Uses PROMOTIF program, PROCHECK, RasMol scripts, PROSITE patterns for detailed analysis of protein structures.
LS-SNP/PDB	http://ls-snp.icm.jhu.edu/ls-snp-pdb/	Generates High-quality protein graphics rendered with UCSF Chimera molecular visualization software. An automated, high-throughput build pipeline that systematically maps human nsSNPs onto Protein Data Bank structures
TM-align	https://zhanglab.ccmb.med.umich.edu/TM-align/	TM-score rotation matrix and Dynamic Programming (DP
BeAtMuSiC	http://babylone.ulb.ac.be/beatmusic	Set of statistical potentials derived from known protein structures

Chapter 3

METHODOLOGY

3.1 *In silico* analysis

3.1.1 Data mining

National Center for Biotechnology Information (NCBI) was explored for retrieving the protein sequence for human ATG5 (Protein ID: NP_004840), the 3D protein structure for ATG5 was obtained from Protein Data Bank (PDB ID: 4GDK). The NCBI and database SNP (dbSNP) was also surveyed for obtaining SNP information of ATG5.

3.1.2 Prediction of deleterious nsSNPs

The computational methodology followed to achieve this objective is shown in the form of a flow diagram (Figure 3.1). According to dbSNP, human ATG5 possesses around ***7,358*** Single nucleotide polymorphisms. This SNP data set was selected as a batch report and submitted to Sorting Intolerant From Tolerant (SIFT) for prediction of deleterious nsSNPs [201]. We have used various additional *in silico* tools that include Polymorphism Phenotype (PolyPhen2) [202, 203], PredictSNP [204] to predict the impact of nsSNPs on the ATG5 protein function. The impact of the change in amino acid is predicted as tolerated or damaging/deleterious. The *in silico* tools calculate the tolerance index and probability scores which directly depicts the effect of substituted amino acid on the function of the protein.

3.1.3. Linkage disequilibrium and haplotype analysis

The International HapMap project [205] was explored to obtain the genotype data for ATG5 for the CEU (CEPH—Utah Residents with Northern and Western European Ancestry) population to carry out linkage disequilibrium analysis. Haploview version 4.2 was used for linkage disequilibrium (LD) analysis. D′ (linkage disequilibrium) and r2 (correlation coefficient) were calculated and analyzed to find out the genetic association, extent of recombination and correlation between the different loci. Further, haplotypes were generated using Haploview 4.2

3.1.4. Prediction of post-translational modification (PTM) sites

Various PTM sites, namely, phosphorylation, ubiquitination, palmitoylation, and calpain-cleavage were predicted in the ATG5 protein by using different web-based tools. For the identification of phosphorylation sites, we used the NetPhos algorithm. The serine (S), threonine (T) and tyrosine (Y) residues were therefore predicted as the various

phosphorylation sites in ATG5. A tool based on the clustering and scoring strategy for predicting palmitoylation (CSS-PALM) was used to obtain various palmitoylation sites [206]. Another tool based on the Bayesian Discriminant Method for prediction of ubiquitination sites (BDM-PUB), was utilized for identifying ubiquitination sites in ATG5. Additionally for obtaining calpain cleavage sites in ATG5, a Group-based Prediction System-Calpain Cleavage Detector (GPS-CCD) 1.0 tool was used [207]. In order to annotate the variously predicted nsSNPs and PTM sites, we used IBS tool (illustrator of biological sequences) [208] to create a schematic diagram representing ATG5 protein structure.

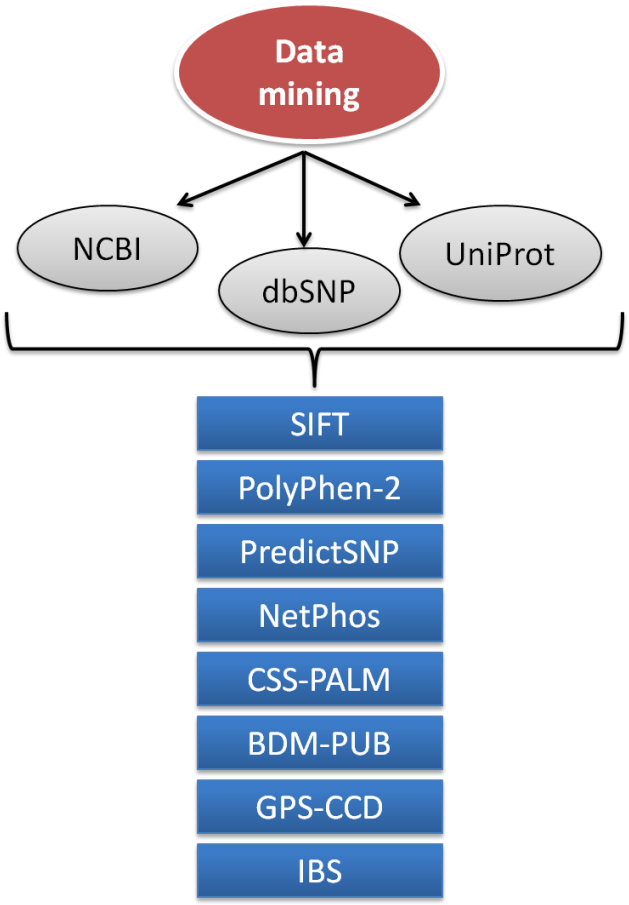

Figure 3.1: Schematic representation of the computational methodology for predicting nsSNPs and PTM sites in ATG5. SIFT, PolyPhen-2, and PredictSNP tools were used for the prediction of nsSNPs, various post-translational modification sites were predicted using NetPhos for phosphorylation sites, CSS-PALM for palmitoylation sites and BDM-PUB for ubiquitination sites and GPS-CCD for calpain-cleavage sites. IBS tool was used to annotate the predicted nsSNPs and PTM sites over protein domain structure.

3.1.5. Impact of damaging nsSNPs of ATG5 on protein structure and function

Diagrammatic representation of computational tools used to study the impact of predicted nsSNPs on ATG5 protein structure and function is shown in Figure 3.2. This methodology was followed to analyze the effect of nsSNPs on the ATG5 sequence as well as structural level. For this, we carried out pathogenicity analysis, conservation analysis and stability analysis of deleterious nsSNPs. Further, mutant structures were generated for computation of total energy, and prediction of binding affinity change upon mutations (Figure 3.2).

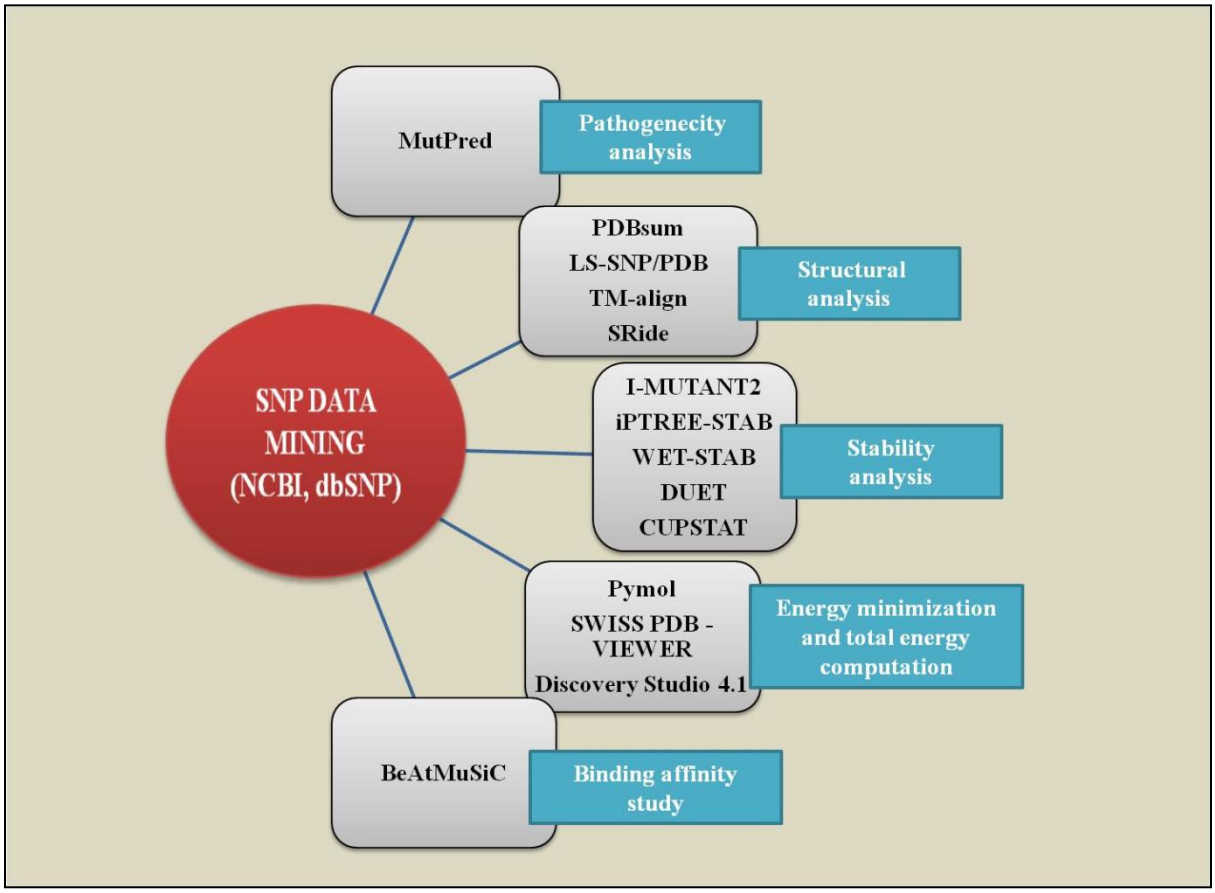

Figure 3.2: Diagrammatic representation of computational tools used to study the impact of predicted nsSNPs on ATG5 protein structure and function [209].

3.1.5.1. Pathogenicity analysis

The nsSNPs leads to amino acid substitutions that may affect protein structure and function and may ultimately lead to a disease [210]. The nsSNPs that may cause alterations in the active sites of the protein like the catalytic sites or PTM sites could be detrimental to the function of the protein. Therefore, such mutations could lead to the development of a disease [211]. In our study, we used MutPred tool pathogenicity analysis of the predicted nsSNPs of

ATG5 (rs34793250 M129V, rs77859116 I65V and rs115576116 A95D) of ATG5. For this, we provided ATG5 protein sequence and the nsSNPs as input to the tool [211].

3.1.5.2. Conservation analysis

For carrying out the conservation analysis of amino acids in ATG5 we used the ConSurf server [212]. We submitted the ATG5 protein structure (PDB ID: 4GDK, chain B) to the ConSurf server, which carried out multiple sequence alignment (MSA) using MAFFT. By using CSBLAST as the homolog search algorithm, 150 ATG5 protein homologs retrieved from the UNIREF90 database. An empirical Bayesian method was utilized by the tool to calculate the conservation scores which depicts the relative degree of evolutionary conservation at each site of the protein sequence. The score ranged from a scale of 1-9, where, the lowest score 1 shows the most variable positions while a score that of 5 indicates the average conservation and a score of 9 specify the most conserved positions.

3.1.5.3. Stability analysis

Protein stability analysis was carried out in order to investigate the impact of the deleterious nsSNPs on ATG5 protein structure. Both ATG5 protein sequence, as well as its structure, was provided as an input to various computational tools to compute protein stability at sequence and structure level. Protein sequence analysis was done using I-Mutant2.0, iPTREE-STAB and WET-STAB tools, while, I-MUTANT2, DUET, and CUPSTAT (Cologne University Protein Stability Analysis Tool) tools analyzed the protein structure for predicting protein stability. I-Mutant2.0 is a web server based on the support vector machine (SVM) feature and ProTherm database for calculating the change in protein stability due to point mutations [213]. The tool can analyze both the protein sequence as well as protein structure to compute the change in free energy (ΔΔG).Accordingly, the ΔΔG value less than 0 signifies decrease in protein stability, while a ΔΔG value more than 0 signifies an increase in the protein stability. iPTREE-STAB is an interpretable decision tree-based method in which the discrimination and prediction are mainly based on decision tree coupled with, adaptive boosting algorithm, and classification and regression tree (r)[214]. WETSTAB is a web server based on a weighted decision table method (WET) for computation of the change in protein stability on the basis of thermal denaturation, due to single-site mutations [215].

DUET is another *in silico* tool that calculates protein stability change due to point mutations in proteins by analyzing the corresponding protein structure. The two utilizes two integrated methods, mCSM and SDM, based on SVM [216]. Last but not the least, CUPSAT tool compute the change in protein stability due to single-site mutations on the basis of factors that include torsion angle distribution and amino acid-atom potentials that analyze the amino acid location at mutation site [217].

Furthermore, SRide server was used to identify the stabilizing residues in the protein structure responsible for maintaining protein stability [218]. Both the native and the mutant structures were analyzed to detect such stabilizing residues. Various parameters like the stabilization center (SC), long-range order (LRO), surrounding hydrophobicity (Hp), and conservation scores are analyzed in order to predict the stabilizing residues in the protein [219]

3.1.5.4. Structural analysis

The predicted nsSNPs were located and annotated over the corresponding protein structure by using a web-based tool, LS-SNP/PDB. This server interprets the nsSNPs on the basis of various parameters that include evolutionary conservation and electrostatic potential, etc [220]. The nsSNPs were annotated over the protein structure for ATG5, PDB ID: 4GDK according to the LS-SNP/PDB server. The native structure (4GDK), as well as the mutant structure, was generated using PyMol and SWISS-PdbViewer [221]. The mutant structure for individual mutation was named as 4GDKx (with I65V mutation), 4GDKy (for A95D mutation), and 4GDKz (for M129V mutation). Similarly structures carrying double mutations were named as 4GDKxy (for I65V + A95D mutations), 4GDKyz (for A95D + M129V mutations), and 4GDKxz (for I65V + M129V mutations). Also, we generated a mutant structure with all the three mutations (I65V + A95D + M129V mutations). Further, we used PDBsum server for overall structural analysis of the x native structure and mutant structures [222].

In addition, TM align tool was used for the calculation of RMSD value and Tm score and for comparing the native and mutant structures for structural similarities [223].

3.1.5.5. Total energy computation

Energy minimization and total energy were computed for the native (4GDK) as well the mutant structures (mentioned above) using CHARMM force field (Discovery Studio 3.5). The final potential energy (CHARMM energy), Van der Waals energy, electrostatic energy, and final RMS gradient was calculated to study the impact of nsSNPs when present individually or in combinations. The combination of mutations was studied in order to analyze their effect on protein structure as these could prove more damaging to protein structure and function than the single mutations [224, 225]

3.1.5.6. Binding affinity analysis

In order to analyze the impact of nsSNPs on the binding affinity of ATG5 protein with its interacting partners ATG12 and ATG16L1, we utilized an *in the silico* tool, BeAtMuSic (http://babylone. ulb.ac.be/beatmusic). This tool predicts binding affinity change due to the presence of mutations that affect protein-protein interactions and overall protein stability. The protein native structure, as well as the mutations, were provided as input to the tool. The calculations were made on the basis of statistical potentials developed from already known protein structures [226].

3.2. Genotyping of predicted nsSNPs and selected non-coding SNPs of ATG5 in HBV infected patients and healthy individuals

3.2.1. Collection of Blood samples from Hepatitis B patients and healthy individuals (control)

Our study group included 550 HBV infected patients and 250 healthy individuals from North India. The blood samples were collected from Post Graduate Institute (PGI), Chandigarh. Duly signed consent forms were obtained from all the subjects before sampling and the study was conducted with the approval of Institutional Ethics Committee. All the patients' samples were tested for various biochemical liver function test (LFT) for enzymes alanine aminotransferase (ALT) or called serum glutamic pyruvic transaminase (SGPT) and aspartate aminotransferase (AST), also called as serum glutamic oxaloacetic transaminase (SGOT) and Bilirubin. The patients were categorized into four groups (Asymptomatic, Acute, Chronic and Liver cirrhosis) based on clinical diagnosis report. All patients were above 18 years of age and others who were co-infected with viruses like HIV or HCV and pregnant women were

excluded from the study group. The same criterions were followed for the healthy donors (control group) except that they all were HBV negative. The clinical and demographic characteristics of HBV patients and healthy control group are shown in Table 3.1.

Table 3.1: Clinical and demographic features of HBV infected patients and healthy control group**.**

Characteristerics	**HBV patients (550)**	**Control group (250)**
Gender		
Male (count)	330	120
Female (count)	220	130
Age (Average)	18-75 (37.02)	22-69 (32.6)
Serological and biochemical parameters		
HBsAg	+	-
HCV	-	-
HIV	-	-
AST(IU/L)	100±79	28±10
ALT(IU/L)	105±89	30±12
Clinical diagnosis (Infection stage)		
Asymptomatic carrier	140	-
Acute hepatitis	114	-
Chronic hepatitis	256	-
Liver cirrhosis	40	-

HBV: Hepatitis B virus; HBsAg: Hepatitis B surface antigen; HCV: Hepatitis C virus; HIV: Human immunodeficiency virus; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase (Mean±SD)

3.2.2. Sample processing

Peripheral blood (5ml) was collected from the HBV patients and healthy individuals in vacutainer tubes with and without anticoagulant (EDTA). For isolation of serum, blood was collected in a tube without anticoagulant. The tube containing blood was allowed to coagulate

for 2-3 hours and then centrifugation was carried out at 2000 rpm for 15 minutes. The serum was isolated without disturbing the settled coagulated blood at the bottom of the tube. Genomic DNA was isolated from the blood samples, mixed with anticoagulant, using a standard salting out the procedure with little modifications as described below [227]. The composition of buffers and reagents used for isolation of DNA is given in **APPENDIX B.**

Protocol for DNA isolation

- To 300µl of a blood sample, RBC lysis buffer is added (3x the volume of blood) for the lysis of red blood cells (RBCs). The tube is then kept on the rocker at room temperature (RT) for 25-30 minutes to permit uninterrupted shaking.
- The tube is then centrifuged at 13000 rpm for 1 minute to get a cream white pellet of white blood cells (WBCs).
- The supernatant obtained is then discarded and the WBC pellet is re-suspended in 300µl of TE buffer (pH 8.0). A vortex machine could be used for the mixing of the pellet. After that, 20µl of 10% SDS solution (final concentration (Fc) = 0.62%) is added. The solution is then incubated at 56°C for 30 minutes in a dry/water bath.
- Then, 150µl of 7.5M ammonium acetate (Fc=2.4M) is added to the solution and mixed vigorously for about 1 minute using a vortex machine for the precipitation of proteins.
- The mixture is then centrifuged at 13000 rpm, at RT, for 15 minutes to obtain the pellet of the precipitated proteins.
- The clear supernatant is then transferred to a fresh tube. To this, chilled absolute ethanol is added (2x the volume of supernatant). The tube is immediately shaken to permit genomic DNA precipitation.
- The tube is then centrifuged at 13000 rpm for 10 minutes to obtain DNA pellet.
- The pellet is washed with 150µl of 70% ethanol by centrifugation. The final pellet is then air-dried at RT for 10-20 minutes.
- The dried DNA pellet is dissolved in the 40µl of TE buffer (pH 7.3) and kept at 65°C for 10 minutes. The DNA is stored at 20°C until further use.

3.2.3. Primer designing

Apart from nsSNPs, we also selected three other important non-coding SNPs of ATG5 i.e. rs2245214, rs12212740 and rs510432 to look for their association with HBV infection susceptibility. These SNPs of ATG5 have already been shown to be associated with Asthma [45, 46], Thyroid carcinoma [48] and Paget disease of bone [32]. Specific amplification primers were designed for detection of the selected SNPs. Such primers were able to amplify the target sequence containing restriction site at the SNP position. For nsSNP rs115576116 G/T natural restriction site was present at the SNP position, therefore, the natural PCR-RFLP analysis was carried out for genotyping in the study population. However, for SNPs, rs34793250, rs77859116, rs2245214, rs12212740, and rs510432, no restriction site was naturally found at the SNP position. Therefore, we manually designed forward primer to introduce an artificial restriction site by adding a mismatched base adjacent to the SNP site which could be then genotyped by Artificial-RFLP (A-RFLP). The general representation of strategy for designing primers for A-RFLP is shown in Figure 3.3.

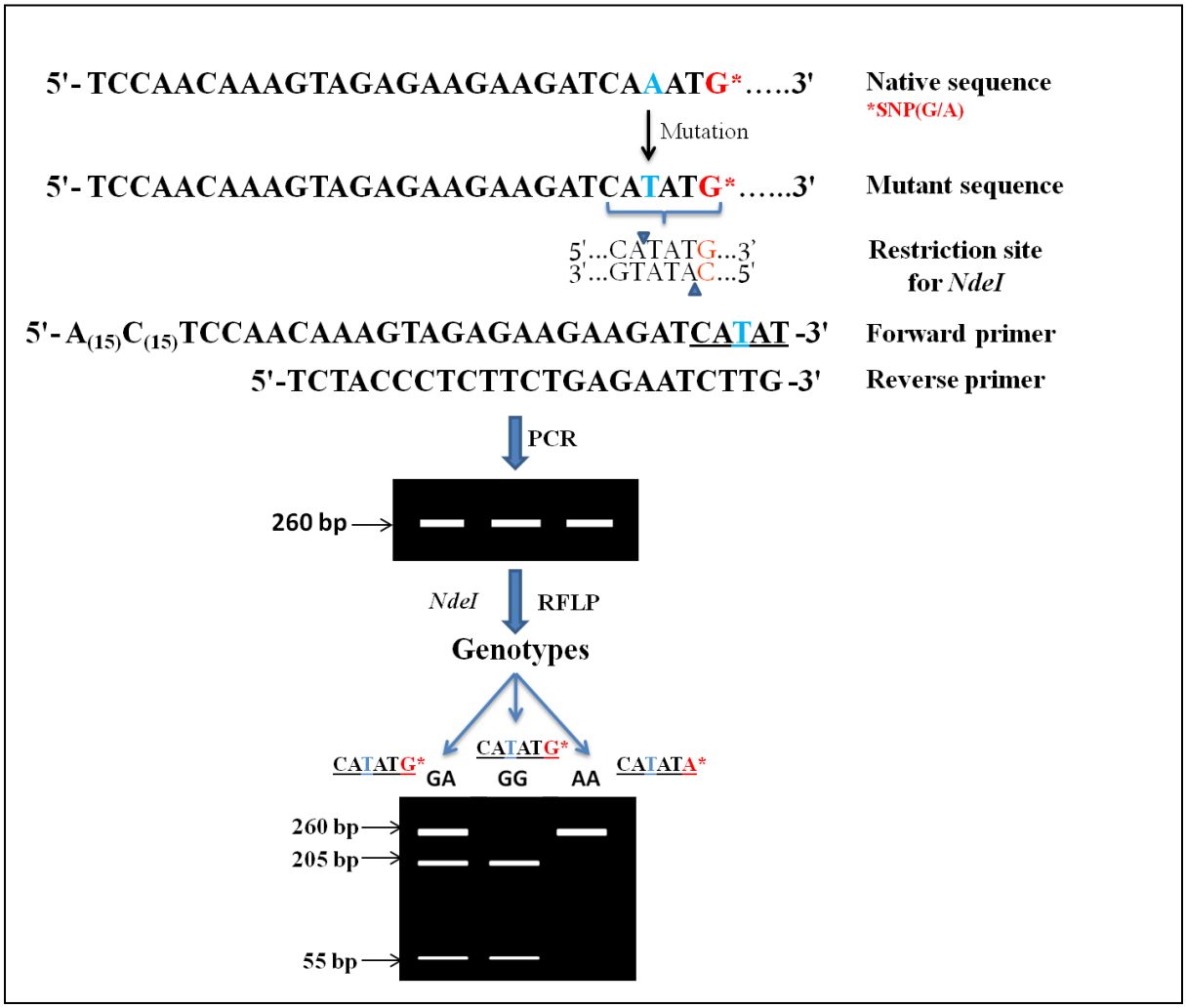

Figure 3.3: The strategy followed for designing primers to incorporate restriction site for the enzyme *NdeI* and to genotype rs510432 A/G SNP by PCR-AFLP method.

3.2.4. Optimization of PCR conditions

Gradient PCR was performed for each primer set to choose the most optimal annealing temperature and primer concentration to amplify the target sequence containing the selected SNP. Total 25µl PCR reaction was prepared that included, 12.5 µl PCR Green Master Mix (GoTaq® Green Master Mix, Promega, Madison, US), and primer (0.25µM) and DNA template (20ng). The reaction was carried out at different annealing temperatures using the veriflex option in the thermal cycler/PCR machine (Applied Biosystems, Foster City, CA).

3.2.5. Optimization of RFLP conditions

For RFLP analysis, PCR products were digested with 1-3 units of the respective restriction enzymes and were incubated at 37 °C, from 2 hours to overnight for selecting the most optimal enzyme units and digestion conditions. Further, the digested products were analyzed by agarose gel electrophoresis. For this, the products were run on 2-3% agarose gel containing ethidium bromide (EtBr) for visualization of various bands for the detection of different genotypes of respective SNPs.

3.2.6. Genotyping of selected SNPs by PCR-RFLP followed by Agarose gel electrophoresis

First of all, PCR was performed with the optimized conditions and the PCR products were run on 1-2% agarose gel containing EtBr to check amplification of the target sequence containing the selected SNP. Then, RFLP was performed and again the digested products were run on 2.5-3% agarose gel containing EtBr for detecting different genotypes of respective SNPs in each sample. The banding pattern after PCR-RFLP was analyzed in all 800 samples, individually.

3.2.7. Haplotype and linkage disequilibrium (LD) analysis

Haplotype and LD analysis for the four selected SNPs of ATG5 (rs34793250, rs2245214, rs12212740 and rs510432) in both cases and controls was performed using SHEsisPlus (online version) [228, 229]. The web-based tool uses expectation maximization (EM) algorithm for the haplotype reconstruction and frequency estimation [230]. The lowest frequency threshold (LFT) of haplotype analysis was set at 0.03 (default value). Every single haplotype with a frequency less than this value was discarded in the association analysis. For

LD analysis, pair-wise D' and R^2 were calculated for the given loci. P-value of 0.05 was set as the threshold for significance level.

3.2.8. Statistical analysis

Genotype data was compiled using Microsoft Excel software (Microsoft Corporation) and the allelic and genotypic frequencies for all the SNPs were calculated. Odds ratios (OR) and 95% confidence intervals (95% CI) were computed to evaluate the risk associated with alleles and genotypes. MEDCALC software (https://www.medcalc.org/) was used to calculate an odds ratio, class interval and *p-value* for significance level. Power of the study was calculated using Sampsize calculator (http://sampsize.sourceforge.net/). Virological and biochemical data is represented as Mean±Standard deviation (SD). One-way analysis of variance (ANOVA) followed by Dunnett's multiple comparisons test was carried out by using the GraphPad Prism version 7.00 (GraphPad Software, La Jolla California USA). P-value less than 0.05 was considered to be statistically significant.

Chapter 4

RESULTS

Objective 1: To predict deleterious nsSNPs in the autophagy-related gene ATG5

4.1. Prediction of deleterious nsSNPs

Sorting Intolerant from Tolerant (SIFT), Polymorphism Phenotype (PolyPhen2), and PredictSNP tools were used to predict the deleterious nsSNPs in ATG5. On the basis of the predicted scores, the nsSNPs could be tolerated or damaging to the protein function. We have identified 3 nsSNPs rs34793250 M129V, rs77859116 I65V and rs115576116 A95D among which two nsSNPs rs34793250 (M129V) and rs115576116 (A95D) were found to deleterious to the ATG5 protein function according to the other *in silico* tools (Table 4.1). We have selected these nsSNPs to analyze their impact on ATG5 protein structure and function by various other computational tools (Objective 2) and also selected them for genotyping in Hepatitis B individuals to find out their link with the infection susceptibility.

Table 4.1: The non-synonymous Single Nucleotide Polymorphisms (nsSNPs) of ATG5 having deleterious effects

SNP INFORMATION		SIFT		PolyPhen-2		PredictSNP			
						PolyPhen-1	nsSNPAnalyzer	PhD-SNP	MAPP
SNP ID	Amino acid change	Score	Predicted impact	Probability score	Predicted impact	Prediction	Prediction	Prediction	Prediction
rs34793250	M129V	1.00	TOLERATED	0.000	BENIGN	**Deleterious**	**Disease**	Neutral	**Deleterious**
rs77859116	I65V	0.21	TOLERATED	0.000	BENIGN	Neutral	Neutral	Neutral	Neutral
rs115576116	A95D	0.32	TOLERATED	0.000	BENIGN	Neutral	Neutral	**Deleterious**	Neutral

4.2 Linkage disequilibrium analysis predicts significantly linked genetic variants and haplotypes

We have used a computational tool to detect genetic variants in ATG5 that are strongly linked with each other. For the haplotype analysis, we selected the SNPs positioned in the chromosome number 6 and also those which were genotyped in the CEU population were selected (Figure 4.1). Three different blocks were found in the same loci with a strong correlation between each other (0.96 between Block 1 and 2, 0.95 between block 2 and 3). Also, 26 markers were predicted in the first block which was of 38 kb in size. The second block was 58kb in size and contained 39 markers. The third block had 16 markers and was of 15kb in size (Figure 4.1b). The linkage disequilibrium revealed other important SNPs which are tabulated and shown in **Table 8.1 APPENDIX A.**

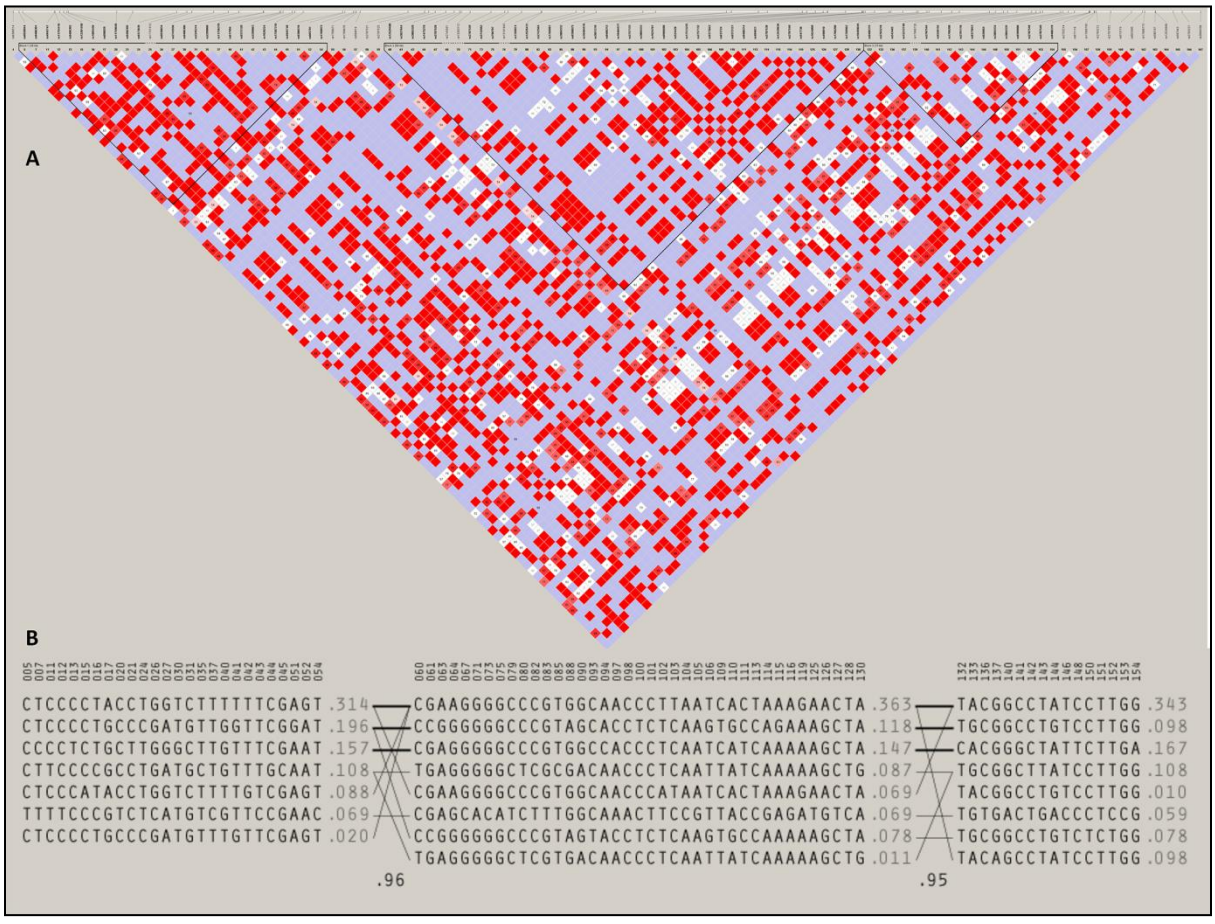

Figure 4.1: Linkage disequilibrium (LD) analysis of ATG5. (A) The Haploblock diagram of the Linkage disequilibrium plot showing 3 blocks having 26, 39 and 16 markers respectively. (B) The diagram showing the correlation of 0.96 between block (1) and (2) and correlation of 0.95 between block (2) and (3) of the genotyped SNPs of ATG5.

4.3 Several PTM sites predicted in ATG5 which were annotated on the protein structure.

11 phosphorylation sites were found at the serine (S), threonine (T) or tyrosine (Y) residues in ATG5 protein sequence through a NetPhos algorithm (Table 4.2).

Table 4.2: Different phosphorylation sites investigated in ATG5

S. No.	Name	Position	Context sequence	Score	Prediction
1	ATG5	66	QEDISEIWF	0.982	*S*
2	ATG5	106	VHFKSFPEK	0.946	*S*
3	ATG5	225	EVCPSAIDP	0.990	*S*
4	ATG5	259	SEHLSYPDN	0.963	*S*
5	ATG5	28	QDEITEREA	0.957	*T*
6	ATG5	75	EYEGTPLKW	0.682	*T*
7	ATG5	192	RIYQTTTER	0.827	*T*
8	ATG5	193	IYQTTTERP	0.597	*T*
9	ATG5	194	YQTTTERPF	0.981	*T*
10	ATG5	36	AEPYYLLLP	0.907	*Y*
11	ATG5	260	EHLSYPDNF	0.506	*Y*

The prediction score ≥ 0.5 was considered as phosphorylated

2 palmitoylation sites were predicted in ATG5 as shown in Table 4.3.

Table 4.3: Various palmitoylation sites predicted in ATG5

S. No.	ID	Position	Peptide
1	ATG5	19	WFGRIPTCFTLYQDE
2	ATG5	128	IEAHFMSCMKEADAL

Overall 7 ubiquitination sites were identified in ATG5 by using BDMPUB web server that utilizes a Bayesian Discriminant Method for prediction (Table 4.4).

Table 4.4: Various ubiquitination sites predicted in ATG5

S. No.	Name	Peptide	Position	Score
1	ATG5	***MTDDKDVLRDVW	5	0.93
2	ATG5	TLVTDKVKKHFQKVM	53	0.82
3	ATG5	KVKKHFQKVMRQEDI	58	1.23
4	ATG5	DLLHCPSKDAIEAHF	118	0.59
5	ATG5	AHFMSCMKEADALKH	130	0.83
6	ATG5	MKEADALKHKSQVIN	136	1.00
7	ATG5	EADALKHKSQVINEM	138	1.45

We found 4 Calpain-cleavage sites in ATG5 which were identified using GPS-CCD 1.0 tool (Table 4.5). Furthermore, we made a schematic representation of protein domain structure and annotated the predicted nsSNPs and all the PTM sites. Illustrator for Biological Sequences (IBS) tool was used to create the schematic protein structure and annotation of various regulatory signatures of ATG5 [208] as shown in Figure 4.2.

Table 4.5: Various Calpain-cleavage sites identified in ATG5

S. No.	Name	Peptide	Position	Score
1	ATG5	MKEADALK | HKSQVIN	136	0.6615
2	ATG5	PFRIYOTT | TERPFIQ	193	0.6846
3	ATG5	PDNFLHIS | IPQPTD	268	0.7808
4	ATG5	PDNFLHISI | IPQPTD*	269	0.6500

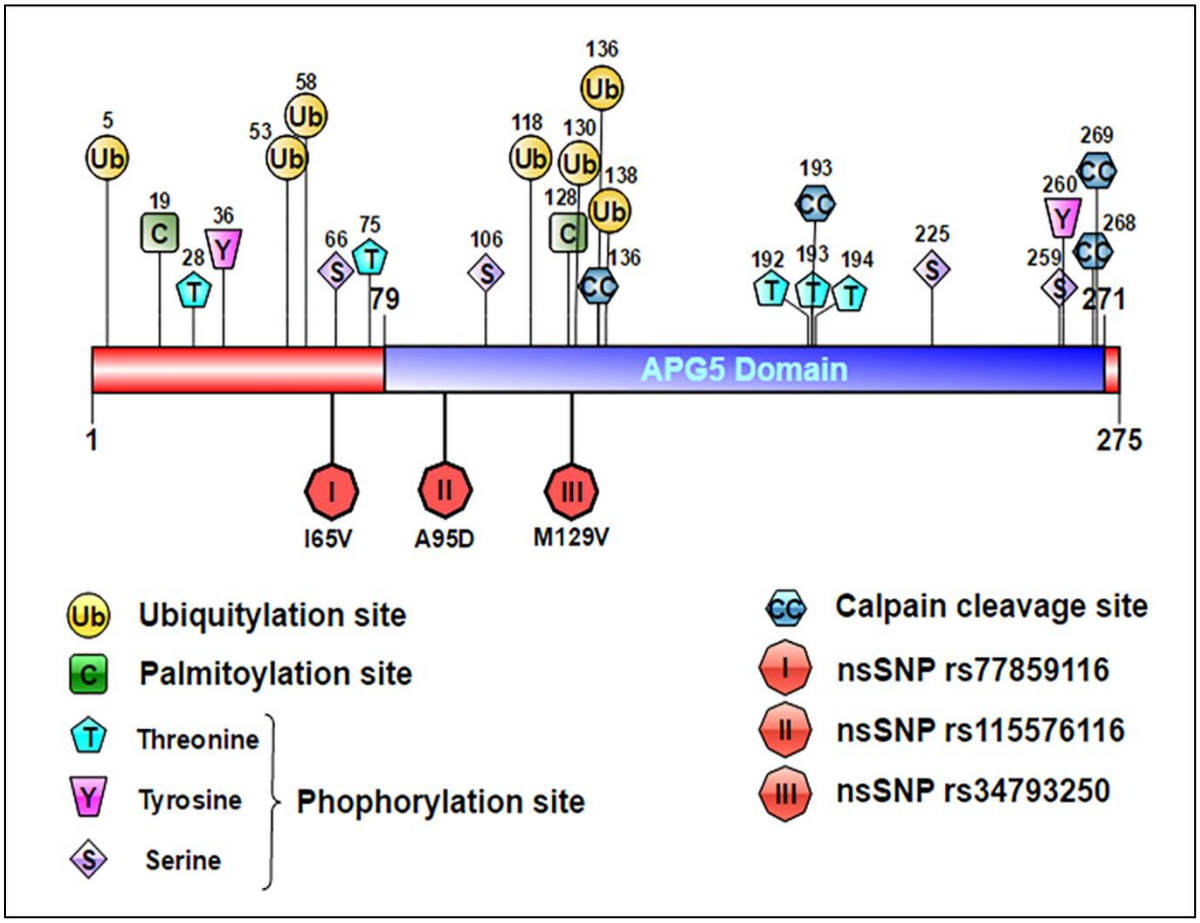

Figure 4.2: The various functional sites predicted in the ATG5 protein. Schematic diagram showing the estimated location of the predicted PTM sites (phosphorylation, ubiquitylation, palmitoylation, and Calpain-cleavage sites) and the nsSNPs in ATG5 [209].

Objective 2: To investigate the impact of predicted nsSNPs on the ATG5 protein at sequence as well as structural levels.

4.4. Conservation profile of the nsSNPs of ATG5.

In order to calculate the conservation rate of each amino acid position in the ATG5 protein, we have used the ConSurf server. According to the results, the predicted nsSNPs positions are at the variable sites in the protein. The normalized conservation score for the amino acids ILE65, ALA95 and MET129 were found as 1.063, 1.307 and 0.241, respectively. The overall conservation analysis of ATG5 protein chain shows it's various functionally important residues including the highest and lowest conserved residues as shown in Figure 4.3.The normalized conservation scores, color and residue variety for all the amino acid residues of ATG5 protein is shown in **Table 8.2 APPENDIX A.**

.

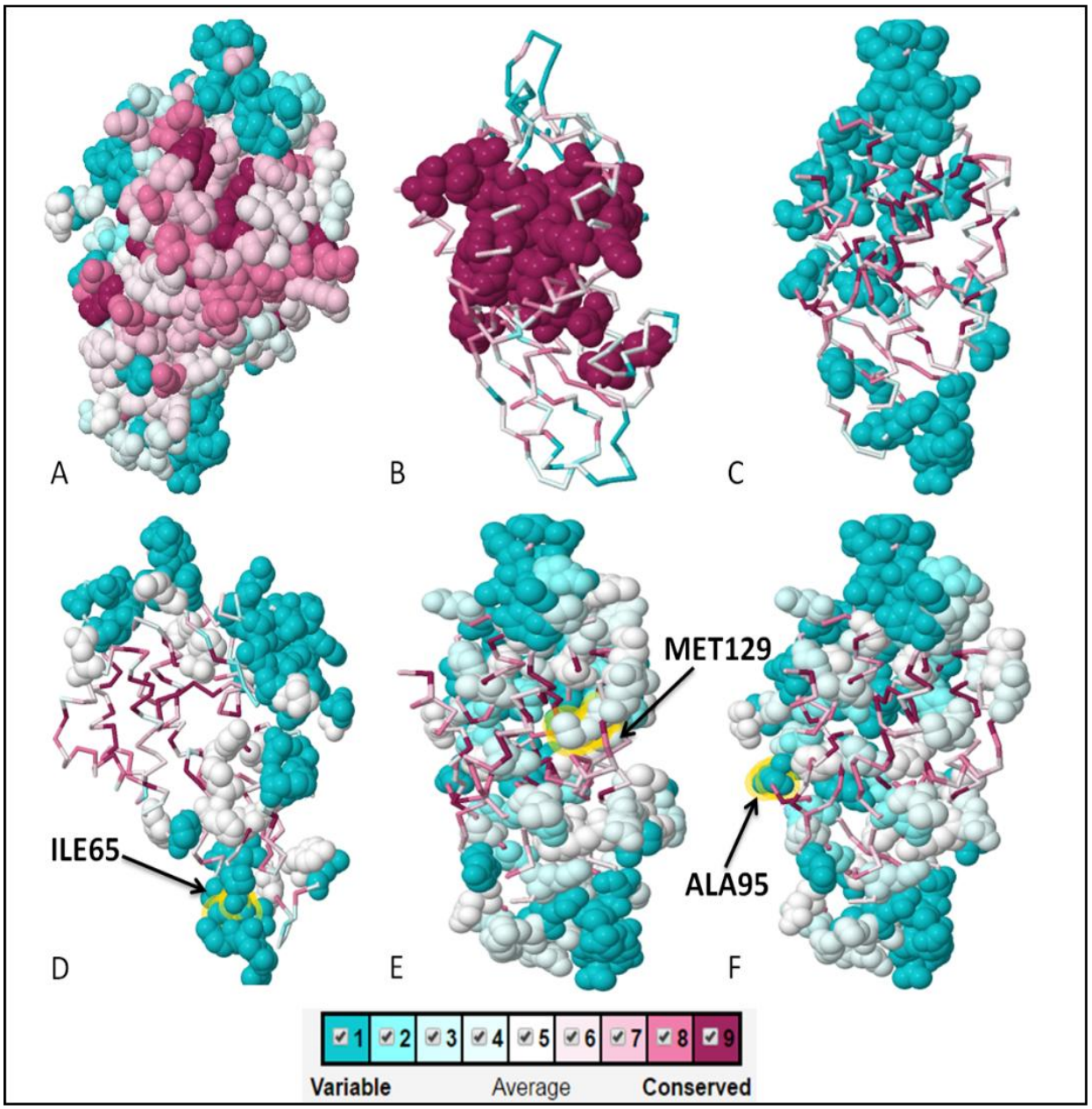

Figure 4.3: The conservation analysis of ATG5 performed by the ConSurf server. (A) shows overall conserved as well variable residues in ATG5 (B) represents the highly conserved residues (C). represents the highly variable residues (D) location of I65 amino acid (E) location M129 amino acid and (F) represents the location of A95 amino acid (The color scale represents the conservation scores [9 - conserved, 1 - variable]) [209].

4.5 nsSNPs destabilize the ATG12–ATG5/ATG16L1 protein complex.

The free energy change value or sign (ΔΔG) is calculated by subtracting the unfolding Gibbs free energy value of the native protein (kcal/mol) from the unfolding Gibbs free energy value of the mutated protein. According to the ternary classification system, support vector machine (SVM) in I-MUTANT2, a ΔΔG value less than 0.5 signify a decrease in the protein stability while a ΔΔG value> 0.5 signify increased stability [231]. The ΔΔG value was observed to be negative for all the three nsSNPs predicted by different tools as shown in Table 4.6. The negative values indicate lower stability and hence the mutations M129V, I65V, and A95D were found to decrease the protein stability as predicted from the protein sequence compared to other mutations (M129V and I65V). Similarly, the protein stability was carried for all the three nsSNPs of ATG5 at structural level using three tools (I-Mutant 2, DUET and CUPSTAT). The results show decrease in protein structure stability upon mutations as indicated by the predicted Gibbs free energy values shown in Table 4.7

Table 4.6: Prediction of change in Gibbs free energy due to the mutations from the ATG5 protein sequence using three computational tools, I-MUTANT2, iPTREE-STAB, and WET-STAB

ATG5 nsSNP	I-MUTANT2 ΔΔG Value (kcal/mol)	iPTREE-STAB ΔΔG Value (kcal/mol)	WET-STAB ΔΔG Value (kcal/mol)	PREDICTION
I65V	-1.53	-1.3200	-1.005	**Destabilizing**
A95D	-0.01	-1.1814	-0.312	**Destabilizing**
M129V	-1.05	-1.6632	-0.312	**Destabilizing**

ΔΔG <0: Decrease Stability; ΔΔG >0: Increase Stability

Table 4.7: Prediction of change in Gibbs free energy due to the mutations from the ATG5 protein structure using three computational tools, I-MUTANT2, DUET and CUPSTAT

ATG5 Mutations (4GDK structure)	I-MUTANT2		DUET				CUPSTAT		PREDICTION
	ΔΔG Value (kcal/mol)	Relative solvent accessibility (%)	mCSM ΔΔG Value (kcal/mol)	SDM ΔΔG Value (kcal/mol)	DUET COMBINED ΔΔG Value (kcal/mol)	Relative solvent accessibility (%)	ΔΔG Value (kcal/mol)	Relative solvent accessibility (%)	
I65V	-0.50	21.1	-1.024	-0.02	-1.005	18.60	-1.31	21.62	**Destabilizing**
A95D	-1.10	20.9	-1.064	-1.44	-1.046	17.10	-2.37	24.5	**Destabilizing**
M129V	-1.16	0.5	-1.826	-1.6	-1.865	0.20	-0.69	0.5	**Destabilizing**

ΔΔG <0: Decrease Stability; ΔΔG >0: Increase Stability

SRide server was utilized for the identification of various stabilizing residues in the ATG5 native and the mutant structures. A total of 15 amino acids were predicted as stabilizing residues in the ATG5 protein. However, there were variations in a number of stabilizing residues in the different mutant structures. Therefore, we have generated the mutant structures that carried individual and the combination of mutations and analyzed the effect of these mutations in the mutated protein structure. Table 4.8 shows the predicted stabilizing residues in ATG5 native as well as in mutant structures.

Table 4.8: Prediction of stabilizing residues in ATG5 native and mutant structures by using the SRide server

Modeled structures	Mutation	No. of Stabilizing residues	List of stabilizing residues	Effect of Mutation
Native structure	-	15	ILE (16), PRO (17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), LEU(135), TYR(184), HIS(266), ILE(267), SER(268), ILE(269).	-
Mutant structure	I65V	15	ILE(16), PRO(17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), LEU(135), TYR(184), HIS(266), ILE(267), SER(268), ILE(269)	No effect
	A95D	14	ILE(16), PRO(17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), LEU(135), TYR(184), HIS(266), ILE(267), ILE(269)	Loss of amino acid SER (268)
	M129V	15	ILE(16), PRO(17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), TYR(184), PHE(264), HIS(266), ILE(267), SER(268), ILE(269)	Amino acid LEU (135) replaced with PHE (264)
	I65V, A95D	14	ILE(16), PRO(17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), LEU(135), TYR(184), HIS(266), ILE(267), ILE(269)	Loss of amino acid SER (268)
	A95D, M129V	15	ILE(16), PRO(17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), TYR(184), PHE(264), HIS(266), ILE(267), SER(268), ILE(269)	Amino acid LEU (135) replaced with PHE (264)
	I65V, M129V	15	ILE(16), PRO(17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), TYR(184), PHE(264), HIS(266), ILE(267), SER(268), ILE(269)	Amino acid LEU (135) replaced with PHE (264)
	I65V, A95D, M129V	15	ILE(16), PRO(17), THR(18), PRO(40), SER(43), PRO(82), ILE(83), VAL(102), HIS(103), TYR(184), PHE(264), HIS(266), ILE(267), SER(268), ILE(269)	Amino acid LEU (135) replaced with PHE (264)

4.6 Annotation of predicted mutations on the ATG5 protein structure

According to the LS-SNP/PDB database, all the predicted nsSNPs (rs34793250, rs77859116, and rs115576116) are located at positions M129V, I65V, and A95D, respectively. These nsSNPs were annotated on the available PDB structure 4GDK as predicted by the database (Figure 4.4)

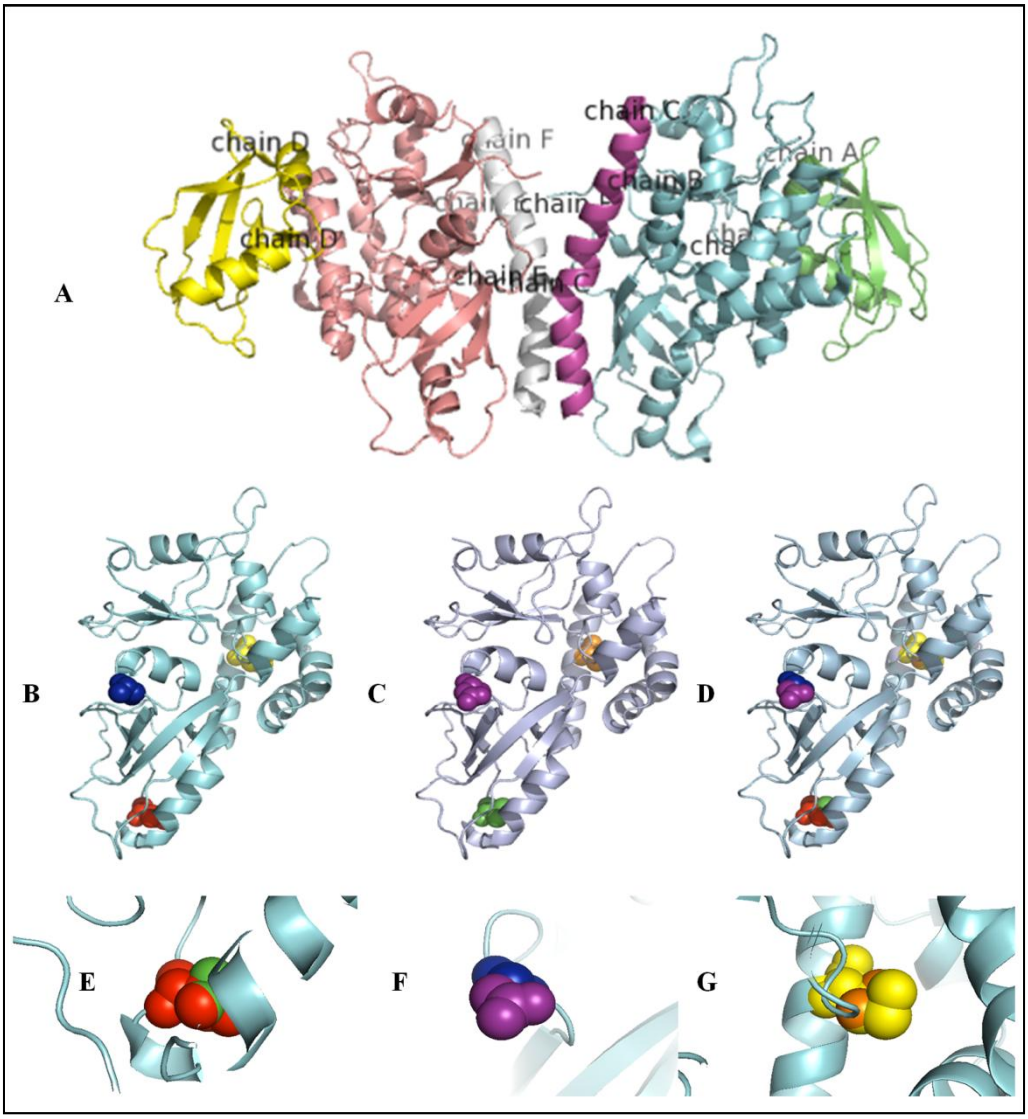

Figure 4.4: Representation of predicted mutations on ATG5 protein structure. (A) 4GDK structure containing different protein chains, chain A & D (ATG12), B & E (ATG5), chain C & E (ATG16L1), (B) ATG5 native structure (chain B) with wild-type residues (I65 [red], M129 [yellow] & A95 ([blue]) (C) ATG5 mutant structure 65V, 129V and 95D. (Green - VAL, purple - ASP, orange - VAL), (D) Native and mutant structures superimposed (E), (F) & (G) shows I65V, A95D, M129V mutations respectively [209].

4.7 nsSNPs increased total energy of the ATG12–ATG5/ATG16L1 complex.

The mutant structures were generated that carried individual mutations and combination of mutations. Both the native (wild type-without any mutation) and mutant structures (7 structures) were subjected to energy minimizations and then total energy was computed. CHARMM force field parameter was used for the computation of the same by using Discovery Studio 3.5. CHARMM energy/potential energy (Kcal/mol), Van der Waals energy (Kcal/mol), electrostatic energy (Kcal/mol), final RMS gradient was computed for native and mutant structures (Table 4.9). It was observed that the mutations (M129V, A95D, and I65V) increased the potential energy of the native structure by 3025.7 Kcal/mol and hence could destabilize the protein complex. The CHARMM energy of the native structure was observed to be −52,421.2 (Kcal/mol) and for the mutant structure, it was found to be −49,395.5 (Kcal/mol). The mutation M129V was observed to increase the potential energy of the complex structure to the maximum that is −49,132.2 Kcal/ mol when compared to other mutations (I65V and A95D) (Figure 4.5).

Table 4.9: Total energy (CHARMM energy) of the native and mutant structures of ATG5 after energy minimizations (The native structure and the mutation with the maximum destabilizing effect have been shown in bold)

Structure 4GDK	Force field	Potential energy/CHARMM energy (Kcal/mol)	Vander Waals energy (Kcal/mole)	Electrostatic energy (Kcal/mole)	Final RMS gradient (Kcal/mole X A^0)
NATIVE	**CHARMM**	**-52421.2**	**-5618.77**	**-54038.7**	**1.36291**
I65V	CHARMM	-49477.7	-5494.97	-50950.5	1.137
A95D	CHARMM	-49432	-5491.57	-50884.4	1.36632
M129V	**CHARMM**	**-49132.2**	**-5511.07**	**-50542.9**	**1.15829**
I65V, A95D	CHARMM	-49330.5	-5492.55	-50740.7	1.17877
A95D, M129V	CHARMM	-49426.3	-5505.35	-50892.1	1.26558
M129V, I65V	CHARMM	-49286.5	-5491.13	-50829.7	1.32941
I65V, A95D, M129V	CHARMM	-49395.5	-5511.04	-50735.2	1.17018

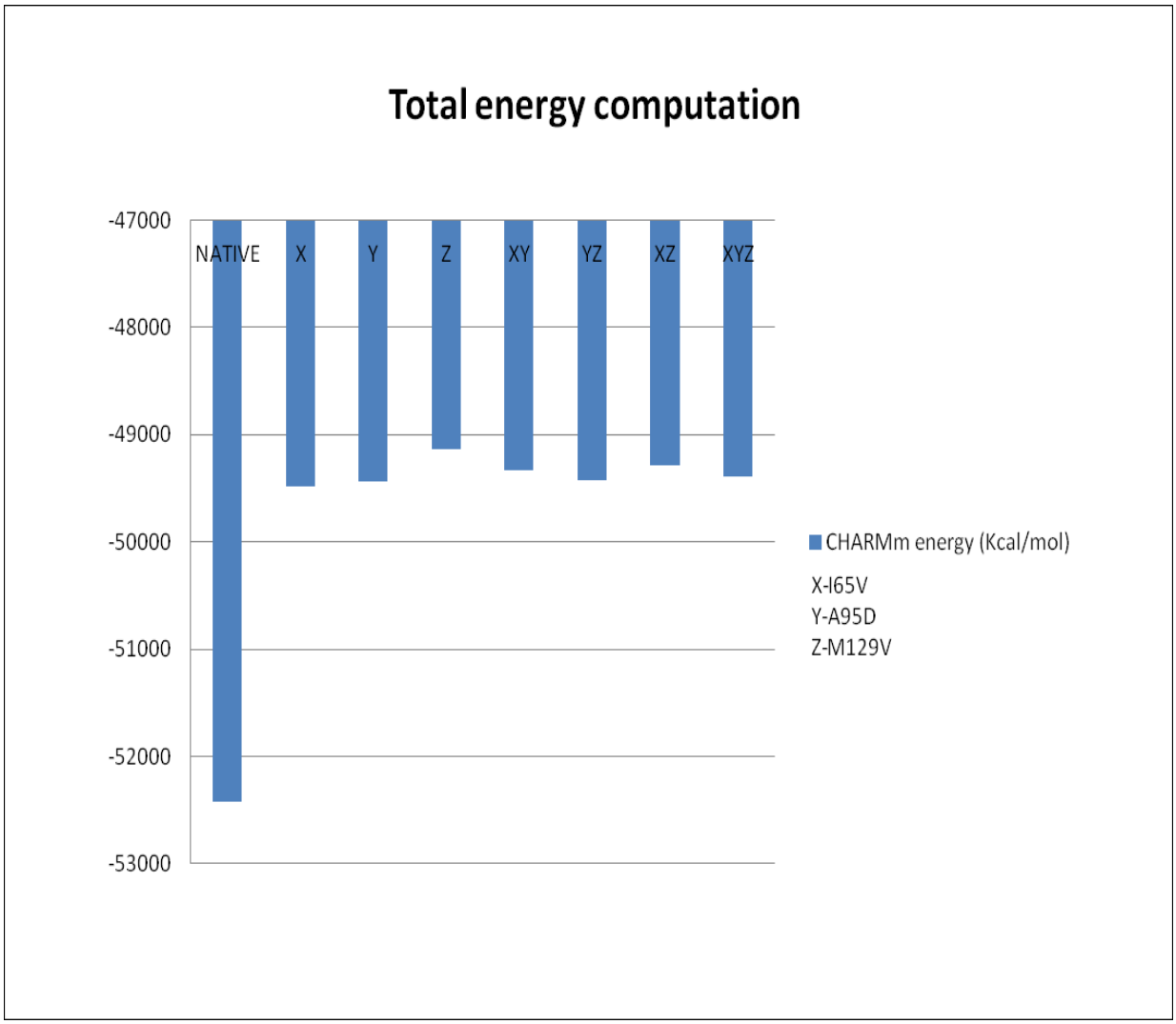

Figure 4.5: Graphical representation of total energy of native and the mutant structures. The mutant structures were generated mutations individually(single mutation) as well as in combinations (double and triple mutations (all three)) [209].

4.8 nsSNPs decreased binding affinity of the components of the ATG12–ATG5/ATG16L1 complex.

BeAtMuSiC program was used to study the effect of the predicted nsSNPs on the interactions of ATG5 with ATG12 and ATG16L. According to the results, all the nsSNPs were observed to decrease the binding affinity of ATG5 with ATG12 and ATG16L1 (Table 4.10). However, one of the mutations A95D was observed to decrease the binding energy value to the maximum and was predicted to be present at the interacting interface of ATG5 and ATG16L1 which may have a strong impact on their interaction and hence over its stability.

Table 4.10: The binding energy change observed upon mutations predicted by BeAtMusiC tool

4GDK	BeAtMuSiC				
	ΔΔGBind (kcal/mol)	Solvent accessibility in partner(s) (%)	Solvent accessibility (in complex) (%)	Interface	Prediction
Chain AD (ATG12) + Chain BE(ATG5)					
I65V	0.31	22.1 %	22.1 %	No	Decreases binding affinity
A95D	0.37	71.55 %	71.55 %	No	Decreases binding affinity
M129V	0.45	0.49 %	0.49 %	No	Decreases binding affinity
Chain BE (ATG5) + Chain CF(ATG16L1)					
I65V	0.31	22.1 %	22.1 %	No	Decreases binding affinity
A95D	0.84	71.55 %	21.59 %	Yes	Decreases binding affinity
M129V	0.43	0.49 %	0.49 %	No	Decreases binding affinity

Objective 3: To genotype the predicted deleterious nsSNPs as well as the selected non-coding SNPs in hepatitis B virus-infected patients and healthy individuals.

4.9 Genotyping nsSNP rs34793250 T/C [M129V]

The dbSNP was used to retrieve the information of the nsSNP rs34793250 to design specific primers that contained a restriction site for the enzyme *Rsa*I at the position of the SNP. The primers pair was checked for their specificity using the BLAST tool [232] and IDT OligoAnalyzer tool [233]. The Table 4.11 shows the primer sequences, amplified product size and expected fragment sizes after RFLP.

Table 4.11: The primer sequences and expected banding pattern for nsSNP rs34793250 T/C [M129V]. The mutant base in the forward primer is shown in the lowercase letter.

ATG5 SNP	Type	Primer sequences 5'– 3'	Amplified product size (bp)	Restriction enzyme	Genotypes (bp)
rs34793250 T/C [M129V]	Missense variant	Forward primer $A_{(15)}C_{(15)}$ TGTTTTAAAGCATCAGCTTCTTTCg Reverse primer TGGGTTATTTCAGTGCTAAGAGA	254	*Rsa*I	TT-198, 56 TC-254, 198, 56 CC-254

4.9.1 Optimized PCR and RFLP conditions

PCR was carried using the specific primers for genotyping the nsSNP rs34793250. For selecting the optimum annealing temperature and optimal primer concentration, gradient PCR was carried out initially. For this, the PCR was carried out in different tubes kept at gradient annealing temperature i: e from 53°C to 57°C. Primer concentration range was also used (5-15 pmol/µl). Amplified PCR products were run on 1.5% agarose gel containing EtBr. We obtained a single 254 bp band with the highest intensity at 57°C with which was regarded as the optimal annealing temperature. Also, 10 pmol/µl primer concentration was found to be the optimum concentration of forward and reverse primers to amplify the target sequence containing the nsSNP as shown in Figure 4.6.

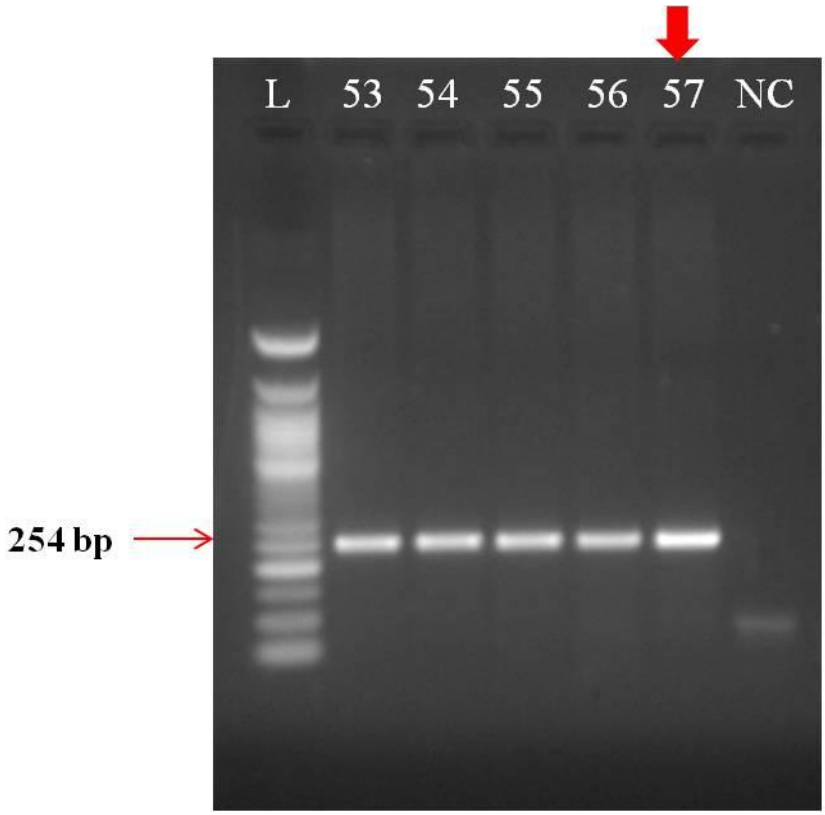

Figure 4.6: Agarose gel image for the amplified products obtained after gradient PCR for nsSNP rs34793250 [M129V]. Lane 1 to lane 6 represents the gradient temperature in increasing order from 53°C to 57°C. L represents 50bp DNA ladder and NC stands for the negative control (without DNA sample) used in the reaction. We obtained a bright 254bp single band without a primer-dimer formation at 57°C with 10 pmol/µl primer concentration.

4.9.2 PCR-RFLP results

The optimized PCR conditions were used to amplify DNA segments from 550 HBV samples and 250 controls samples in order to genotype the nsSNP (rs34793250). The optimized PCR-RFLP conditions are shown in Table 4.12. Thereafter, the PCR products were digested with *Rsa*I (3 units) by incubating at 37°C, overnight. We observed the homozygous wild-type genotype TT (198bp and 56bp bands) and heterozygous genotype TC (254bp, 198bp, and 56bp bands) in different HBV as well as healthy control samples. The representative agarose gel picture is shown in Figure 4.7.

Table 4.12: The optimized PCR conditions for nsSNP rs34793250 T/C [M129V]

PCR Stage	Temperature	Duration
Initial denaturation	95°C	2 minutes
Denaturation	95°C	45seconds
Annealing	57°C	1 minute
Extension	72°C	45 seconds
Final extension	72°C	5 minutes

35 cycles

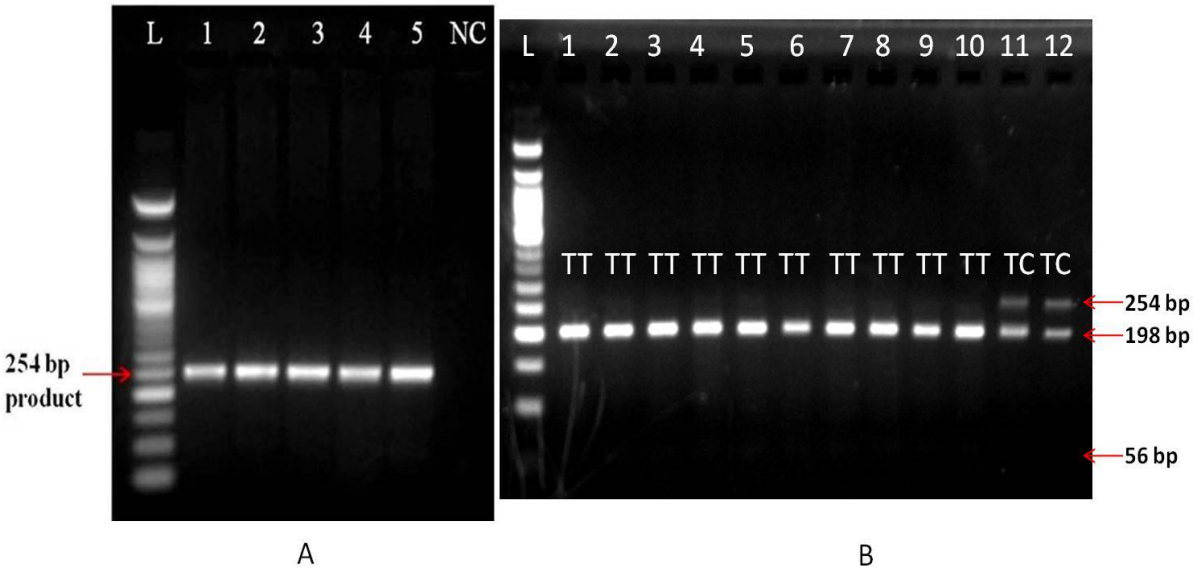

Figure 4.7: **The representative PCR-RFLP agarose gel image for genotyping nsSNP rs34793250 T/C [M129V] in HBV infected and healthy control samples.** (A). PCR amplification gel picture showing 254 bp product obtained in HBV infected samples (Lane 1 to 5). NC represents the negative control used in the reaction (L=50bp DNA ladder) (B). RFLP analysis of the HBV infected samples (Lane 1 to 12) (L=50bp DNA ladder).

4.9.3 Allelic and Genotypic frequencies

Genotyping data was compiled and compared for calculating the allelic and genotyping frequencies in 550 HBV infected and 250 healthy control samples. The nsSNP rs34793250 (T>C) was observed to be associated with HBV susceptibility. The C allele was found to be significantly associated with HBV risk (OR=3.35, 95% CI=1.31-8.59, p=0.01). The allelic and the genotypic frequencies obtained are shown in Table 4.13. We estimated the power of the study using Sampsize calculator by computing odds ratio and minor allele frequency for the given sample size which was observed to be 62%. Moreover; we found a significant association of C allele in the asymptomatic, acute and chronic stages of HBV infection. No association of this SNP was found in cirrhosis patients (Table 4.14).

Table 4.13: **Genotypic and allelic frequencies of ATG5 polymorphism rs34793250 (T/C) in HBV infected cases and healthy control samples. (N = total number of samples)**

ATG5 SNP rs ID Genotype and allele	HBV (N=550)	Control (N =250)	Odds ratio [95% CI]	p-value
rs34793250 (T>C)				
TT	514[93]	245[98]	Ref	/
TC	36[7]	5[2]	**3.43** [1.3303 to 8.8538]	**0.01**
CC	0[0]	0[0]		
T	1064 [97]	495[99]	Ref	/
C	36[3]	5[1]	**3.35** [1.3065 to 8.5878]	**0.01**

Significant values ($p<0.05$) are shown in bold.

4.9.4 Comparison of Virological and Biochemical parameters with different genotypes

Viral load in different HBV infection categories was co-related with the genotypes of HBV patients. No significant difference was found in mean viral load between the host genotypes for this SNP. Although not significant, the low viral load was found in patients with homozygous mutant genotype CC as compared to homozygous wild-type genotype TT, in all categories of HBV infection (Table 4.15). Additionally, statistical analysis was carried out to study the relation between the patient genotypes and biochemical parameters namely, AST, ALT (Table 4.15). The patients (asymptomatic and chronic stage) with heterozygous genotype (TC) were found to have high mean ALT level as compared to patients with homozygous wild-type genotype (TT vs. TC, p= 0.02). The mean AST level was also found to be significantly high in patients (chronic stage) with heterozygous genotype (TC) than with homozygous wild-type genotype (TT vs. TC, p= 0.02). We did not find any significant difference in the mean Bilirubin levels of patients with different genotype (Table 4.15).

Table 4.14: Genotypic and allelic frequencies of ATG5 nsSNP rs34793250 T/C in different HBV infection stages and healthy control samples

HBV infection Categories (no of samples)	Genotype and Allele	Cases (N=550) n/N	Control (N=250) n/N	Odds ratio [95% CI]	p value
ASYMPTOMATIC (140)	TT	129/140	245/250	Ref.	
	TC	11/140	5/250	**4.18** [1.42 to 12.28]	**0.01**
	CC	0/140	0/250		
	T	269 (96.07%)	495 (99%)	Ref.	/
	C	11 (3.93%)	5(1%)	**4.05** [1.39 to 11.77]	**0.01**
ACUTE (114)	TT	106/114	245/250	Ref.	
	TC	8/114	5/250	**3.70** [1.18 to 11.57]	**0.02**
	CC	0/114	0/250		
	T	220 (96.49%)	495 (99%)	Ref.	/
	C	8 (3.51%)	5 (1%)	**3.6** [1.16 to 11.13]	**0.03**
CHRONIC (256)	TT	241/256	245/250	Ref.	
	TC	15/256	5/250	**3.05** [1.09 to 8.52]	**0.03**
	CC	0/256	0/250		
	T	497 (97.07%)	495 (99%)	Ref.	/
	C	15 (2.93%)	5 (1%)	**2.99** [1.08 to 8.28]	**0.04**
CIRRHOSIS (40)	TT	38/40	245/250	Ref.	
	TC	2/40	5/250	2.58 [0.48 to 13.77]	0.27
	CC	0/40	0/250		
	T	78 (97.5%)	495 (99%)	Ref.	/
	C	2 (2.5%)	5 (1%)	2.53 [0.48 to 13.31]	0.27

p value <0.05 was considered significant (shown in bold)

Table 4.15: Analysis of the virological and biochemical parameters in HBV infected patients with nsSNP rs34793250 (T/C) genotype.

HBV infection category (no of samples)	Genotypes (no of samples)	Viral load (copies/ml) [Mean±SD]	p value	AST (IU/L) [Mean±SD]	p value	ALT (IU/L) [Mean±SD]	p value	Bilirubin (mg/dL) [Mean±SD]	p value
Asymptomatic (140)	TT (129) TC (11) CC(0)	130±111 77± 97 0	0.13*	60±45 48±20 0	0.38*	46±51 99±190 0	**0.02***	0.56±0.25 0.62±0.26	0.36
Acute (114)	TT (106) TC (8) CC(0)	1624±1459 834±429 0	0.13*	69±51 63±32 0	0.74*	87±59 71±36 0	0.45*	0.92±0.42 0.68±0.19	0.11
Chronic (256)	TT (n=241) TC (n=15) CC(n=0)	2151191±3710419 983546±1364270 0	0.23*	102±61 156±281 0	**0.02***	104±72 166±319 0	**0.02***	1.1±0.50 0.94±0.52	0.13
Cirrhosis (40)	TT (38) TC (2) CC(0)	1890032059 ±10152697922 45676579±0.71 0	0.80*	168±75.3 118±54 0	0.36*	154±85 138±26 0	0.79*	0.92±0.34 0.76±0.11	0.16

A ‘p value’ <0.05 was considered to be statistically significant. Significant values are shown in **bold.**

4.10 Genotyping of nsSNP rs77859116 T/C [I65V]

The dbSNP was used to retrieve the information of the nsSNP rs77859116. We have designed specific primers such that they amplify a region that contains desired restriction site for the enzymes *Ase*I at the SNP position. The primers pair was further checked for their specificity using the BLAST tool [232] and IDT OligoAnalyzer tool [233]. The Table 4.16 shows the primer sequences, amplified product size and expected fragment sizes after RFLP.

Table 4.16: The primer sequences and expected banding pattern for nsSNP rs77859116 T/C [I65V]. The mutant base in the forward primer is shown in the lowercase letter.

ATG5 SNP	**Type**	**Primer sequences 5'– 3'**	**Amplified product size (bp)**	**Restriction enzyme**	**Genotypes (bp)**
rs77859116 T/C [I65V]	Missense variant	Forward primer $C_{(15)}A_{(15)}$CCTTCATATTCAAACCATATCTCAtTA Reverse primer TATTCAAAATAATGCAAAGAACACGG	259	*Ase*I	TT-204, 55 TC-255, 204, 55 CC-259

4.10.1 Optimized PCR-RFLP conditions

PCR was carried using the designed primers for genotyping the nsSNP rs77859116. For selecting the optimum annealing temperature for the primer set, gradient PCR was carried out initially. For this, the PCR was carried out in 13 different reaction tubes kept at gradient annealing temperatures, from 48°C to 68°C. Amplified PCR products were analyzed on 1.5% agarose gel with ethidium bromide. We obtained a single 259 bp band with maximum intensity at 56°C which was regarded as the optimum annealing temperature. Also, 10 pmol/µl primer concentration was found to be the optimum concentration of forward and reverse primers to amplify the target sequence containing the nsSNP as shown in Figure 4.8

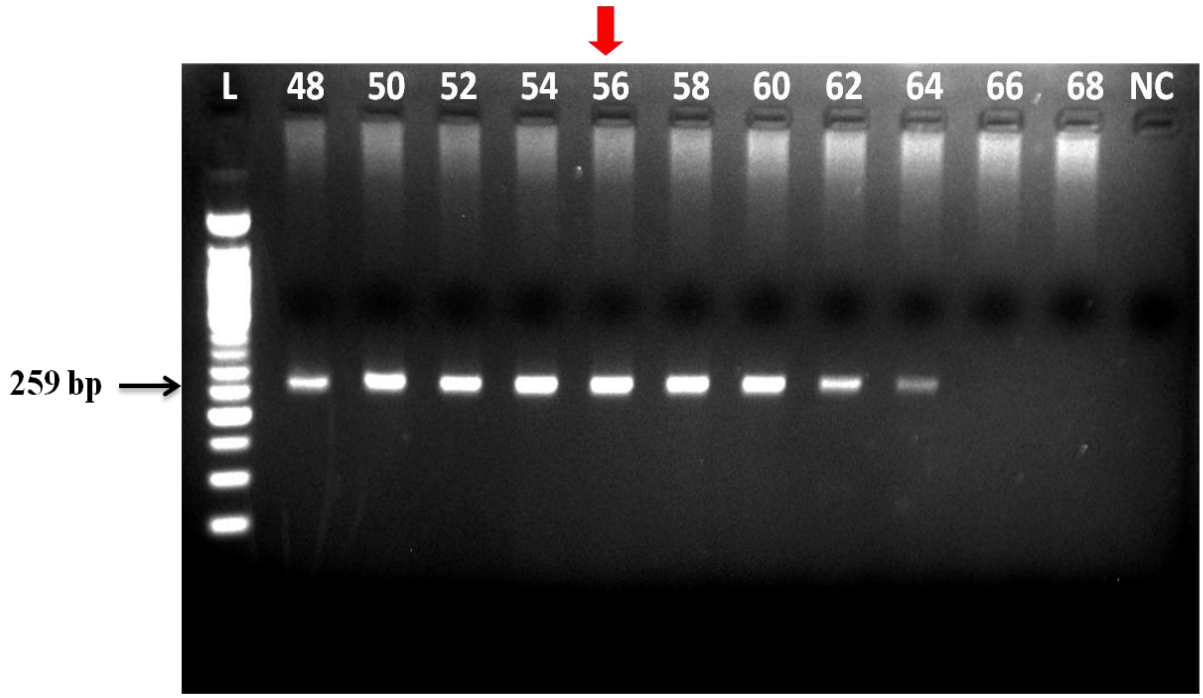

Figure 4.8: Agarose gel image for the amplified products obtained after gradient PCR for the nsSNP rs77859116. Lane 2 to lane 12 represents the gradient temperature in increasing order from 48°C to 68°C, shown on the top of the gel. L represents 50bp DNA ladder and NC in lane 13, is the negative control (No DNA template) used in the reaction. We obtained a bright 259bp single band without a primer-dimer formation at 56°C and hence it was selected as an optimum annealing temperature for the PCR.

4.10.2 PCR-RFLP results

The optimized PCR conditions were used to amplify DNA segments from 550 HBV patient samples and 250 healthy control samples in order to genotype the nsSNP. The optimized PCR conditions are shown in Table 4.17. The PCR products were digested with *Ase*I restriction enzyme (2 units, 37°C overnight incubation). We observed only the homozygous wild-type genotype TT (204bp and 55 bp bands) in HBV infected as well as healthy control samples. The representative agarose gel picture is shown in Figure 4.9.

Table 4.17: The optimized PCR conditions for genotyping nsSNP rs77859116 T/C [I65V]

PCR Stage	Temperature	Duration
Initial denaturation	95°C	2 minutes
Denaturation	95°C	30 seconds
Annealing	56°C	1 minute
Extension	72°C	40 seconds
Final extension	72°C	5 minutes

35 cycles

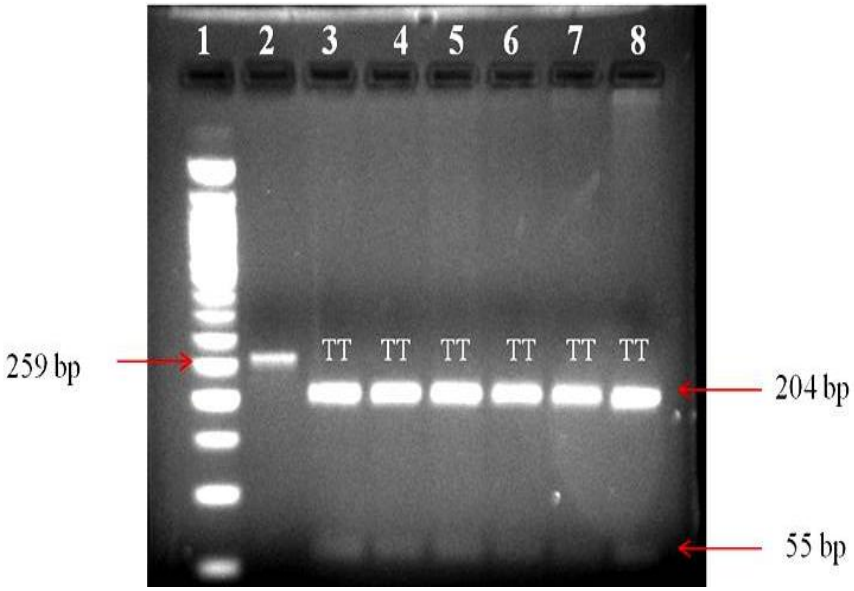

Figure 4.9: The representative PCR-RFLP agarose gel picture for genotyping nsSNP rs77859116 T/C [I65V] in HBV infected as well as healthy control samples. Lane 1 represents a 50bp DNA ladder; Lane 2 contains the undigested 259 bp PCR product (for reference). Lane 3-8 contains the *Ase*I digested PCR products from HBV infected samples.

4.10.3 Allelic and Genotypic frequencies

Genotyping data was compiled and compared for calculating the allelic and genotypic frequencies. We observed only homozygous wild-type genotype (TT) in all the HBV infected samples, therefore, we have analyzed only 150 HBV infected and 150 healthy control samples. The allelic and genotypic frequencies are shown in the Table 4.18.

Table 4.18: Genotypic and allelic frequencies of ATG5 polymorphism rs77859116 (T/C) in HBV cases and healthy control samples (N = total number of samples)**.**

ATG5 SNP rs ID Genotype and allele	**HBV (N=550)**	**Control (N =250)**	**Odds ratio [95% CI]**	**p-value**
rs77859116 (T>C)				
TT	150/150[100]	150/150 [100]		
TC	0/150[0]	0/150[0]	Not estimable	Not estimable
CC	0/150[0]	0/150[0]		
T	300[100]	300[100]	Not estimable	Not estimable
C	0 [0]	0[0]		

4.11 Genotyping of nsSNP rs115576116 G/T [A95D]

The dbSNP was used to retrieve the information of the nsSNP rs77856116. Fortunately, there was a restriction site for enzyme *Bgl*II naturally present at the position of the nsSNP. Therefore, we selected natural primer pair from the PRIMER BLAST software. The primer pair was further checked for their specificity using the BLAST tool [232] and IDT OligoAnalyzer tool [233]. The Table 4.19 shows the primer sequences, amplified product size and expected banding pattern for different genotypes after RFLP.

Table 4.19: The primer sequences, amplified product size, restriction enzyme and expected banding pattern for nsSNP rs115576116 G/T [A95D]

ATG5 SNP	**Type**	**Primer sequences 5'– 3'**	**Amplified product size (bp)**	**Restriction enzyme**	**Genotypes (bp)**
rs115576116 G/T [A95D]	Missense variant	Forward primer AATTGGGGGAGGGGAACAGA Reverse primer AAACCACCTGATGCCTGGAAA	966	*Bgl*II	GG-966 GT-966, 729, 237 TT-729, 237

4.11.1 Optimized PCR-RFLP conditions

PCR was carried using the designed primers for genotyping the nsSNP rs77859116. For selecting the optimum annealing temperature for the primer set, gradient PCR was carried out initially. For this, the PCR was carried out in 6 different reaction tubes kept at gradient annealing temperatures, from 60°C to 70°C. Amplified PCR products were analyzed on 1.5% agarose gel with ethidium bromide. We obtained a single 966 bp band with the highest intensity at 64°C and was regarded as the optimum annealing temperature for amplification of the region containing the nsSNP as shown in Figure 4.10.

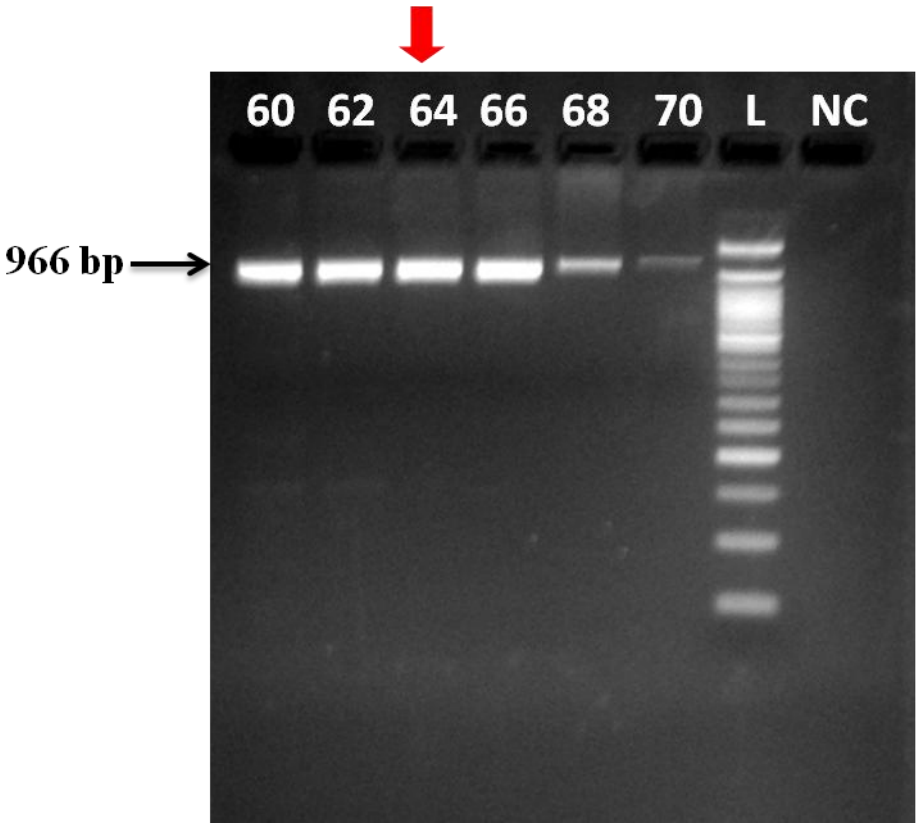

Figure 4.10: Agarose gel image for the amplified products obtained after gradient PCR for the nsSNP rs115576116. Lane 1 to lane 6 represents the gradient temperature in increasing order from 60°C to 70°C. L represents 50bp DNA ladder and NC in lane 8, is the negative control (no DNA template) used in the reaction. We obtained a bright 966bp single band without a primer-dimer formation at 64°C and hence it was selected as an optimum annealing temperature for the PCR.

4.11.2 PCR-RFLP results

The optimized PCR conditions were used to amplify DNA segments from HBV infected samples and healthy control samples. The optimized PCR conditions are shown in Table 4.20. The PCR products were digested with

BglII restriction enzyme (1 unit, 37°C overnight incubation). We observed only the homozygous wild-type genotype GG (966 bp) in all the samples screened. The representative agarose gel picture is shown in Figure 4.11

Table 4.20: **The optimized PCR conditions for genotyping nsSNP rs115576116 G/T [A95D]**

PCR Stage	Temperature	Duration
Initial denaturation	95°C	2 minutes
Denaturation	95°C	30seconds
Annealing	64°C	30s
Extension	72°C	30 seconds
Final extension	72°C	5 minutes

35 cycles

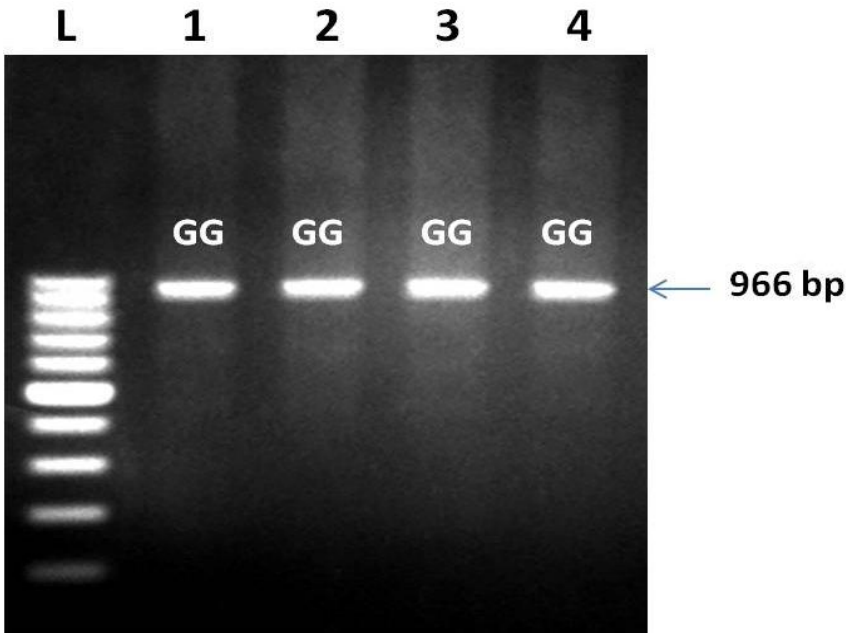

Figure 4.11: The representative PCR-RFLP agarose gel picture for genotyping nsSNP rs115576116 G/T [A95D] in HBV infected as well as healthy control samples. L represents a 100bp DNA ladder; Lane 2 to lane 4 represents the *Bgl*II digested PCR products from HBV infected samples.

4.11.3 Allelic and Genotypic frequencies

Genotyping data was compiled and compared for calculating the allelic and genotypic frequencies. We observed only homozygous wild-type genotype (GG) in all the HBV infected as well healthy control samples. Therefore we have analyzed only 150 HBV infected and 150 healthy control samples. The allelic and genotypic frequencies are shown in the Table 4.21.

Table 4.21: Genotypic and allelic frequencies of ATG5 polymorphism rs115576116 G/T [A95D] in HBV infected cases and healthy control samples. (N = total number of samples)

ATG5 SNP **Genotype and allele**	**HBV** **(N=150)**	**Control** **(N =150)**	**Odds ratio** **[95% CI]**	**p-value**
rs115576116 (G>T)				
GG	150/150[100]	150/150[100]	Not estimable	Not estimable
GT	0/150[0]	0/150[0]		
TT	0/150[0]	0/150[0]		
G	300[100]	300[100]	Not estimable	Not estimable
T	0 [0]	0 [0]		

4.12 Genotyping of non-coding SNP rs2245214 C/G

The dbSNP was used to retrieve the information of the noncoding SNP rs2245214 C/G to design specific primers that contained a restriction site for the enzyme *Sal*I at the position of the SNP. The primers pair was checked for their specificity using the BLAST tool [232] and IDT oligo analyzer. Table 4.21 shows the primer sequences, amplified product size and expected banding pattern for different genotypes after RFLP analysis.

Table 4.22: The primer sequences amplified product size, restriction enzyme and expected banding pattern for non-coding SNP rs2245214 C/G. The mutated base in the forward primer is shown in lower case letter.

ATG5 SNP	Type	Primer sequences 5'− 3'	Amplified product size (bp)	Restriction enzyme	Genotypes (bp)
rs2245214 C/G	Intronic variant	Forward primer $C_{(12)}A_{(7)}$CACCCCCA GCTTTTAAATTTTTATATGAACTTATCAGTcGA Reverse primer GCTCCTAGCCCAGGTTTAGA	518	*Sal*I	CC-518 CG-518, 462, 56 GG-462, 56

4.12.1 Optimized PCR-RFLP conditions

PCR was carried using the designed primers for genotyping the SNP rs2245214 C/G. For selecting the optimum annealing temperature for the primer set, gradient PCR was carried out initially. For this, the PCR was carried out in 6 different reaction tubes kept at gradient annealing temperatures, from 54°C to 64°C. Amplified PCR products were analyzed on 1.5% agarose gel with EtBr. We obtained a single 518 bp band at 60°C which was selected as the optimum annealing temperature for amplification of the region containing the SNP as shown in Figure 4.12

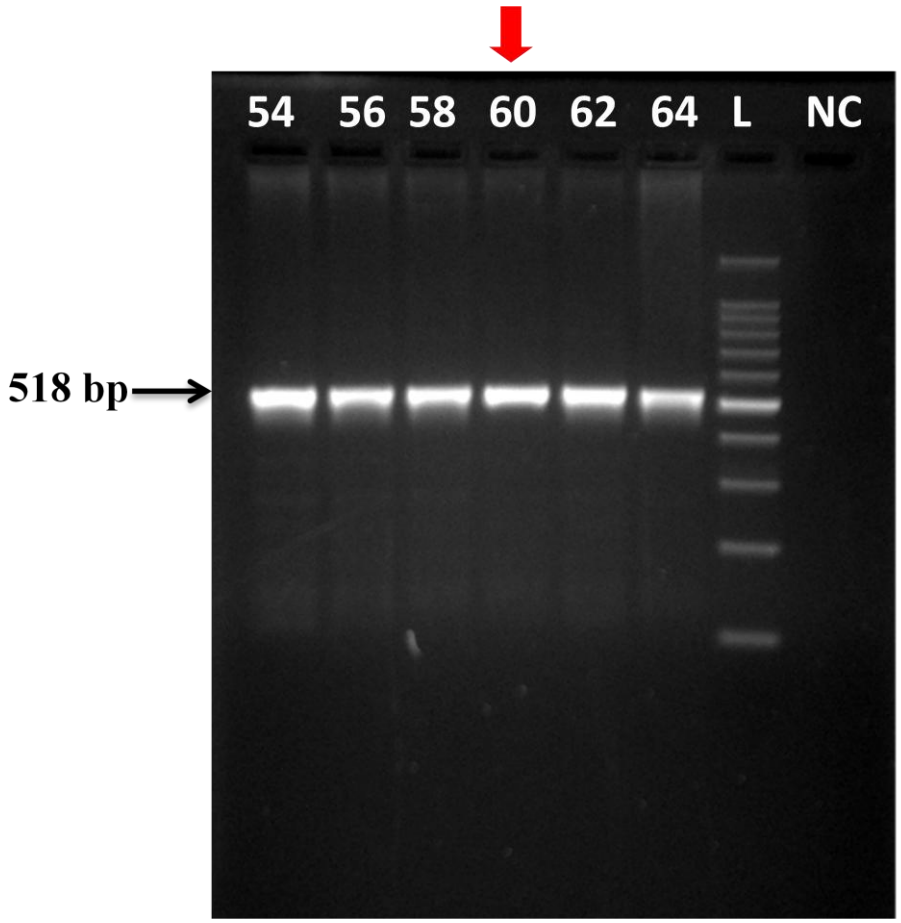

Figure 4.12: Agarose gel image for the amplified products obtained after gradient PCR for the non-coding SNP rs2245214 C/G. Lane 1 to lane 6 represents the gradient temperature in increasing order from 54°C to 64°C. L represents 100 bp DNA ladder and NC in lane 8, is the negative control (without DNA template) used in the reaction. We obtained a bright 518 bp single band without a primer-dimer formation at 60°C and hence it was considered to the optimum annealing temperature for the PCR.

4.12.2 PCR-RFLP results

The optimized PCR conditions and optimal primer concentration were used to amplify DNA segments from HBV infected samples and healthy control samples. The optimized PCR conditions are shown in Table 4.23. The PCR products were digested with *Sal*I restriction enzyme (2 units, 37°C overnight incubation). We observed all the three genotypes; Homozygous wild-type CC (518 bp), heterozygous genotype CG (518 bp, 462 bp, and 56 bp) and homozygous mutant genotype GG (462bp and 56bp) in different HBV infected as well as healthy control samples as shown in Figure 4.13.

Table 4.23: The optimized PCR conditions for genotyping SNP rs2245214 C/G

PCR Stage	Temperature	Duration
Initial denaturation	95°C	2 minutes
Denaturation	95°C	30 seconds
Annealing	60°C	30 seconds
Extension	72°C	30 seconds
Final extension	72°C	5 minutes

35 cycles

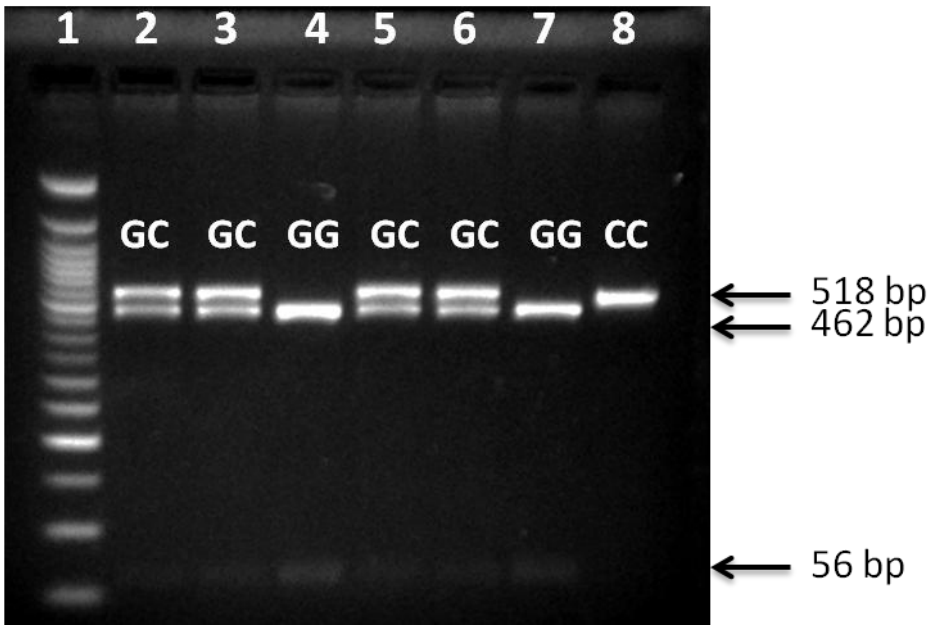

Figure 4.13: The representative PCR-RFLP agarose gel picture for genotyping SNP rs2245214 C/G in HBV infected as well as healthy control samples. Lane 1 represents a 50 bp DNA ladder; Lane 2 to lane 8 represents the *Sal*I digested PCR products from HBV infected samples.

4.12.3 Allelic and Genotypic frequencies

The non-coding SNP rs2245214 (C/G) was also found to be associated with HBV infection susceptibility. The G allele was found to be significantly associated with HBV risk as compared to control samples (OR=1.30, 95%CI=1.05-1.63, p=0.02; estimated power 50%) (Table 4.24). Also, we observed significant association in the homozygous (CC vs. GG, OR=1.93, 95%CI=1.14- 3.27, p=0.02) and recessive genotype models (CC+CG vs. GG, OR= 1.71, 95%CI=1.04-2.82, p=0.03) (Table 4.24). Moreover, the SNP was found to be significantly associated in the chronic (C vs. G, OR=1.36, 95%CI=1.05 to 1.75, p=0.02; CG+CC vs. GG, OR=2.03, 95%CI=1.175 to 3.520, p=0.01; CC vs. GG, OR=2.20, 95%CI=1.22 to 3.95, p=0.01) and cirrhosis stage (C vs. G, OR=1.66, 95%CI=1.03 to 2.68, p=0.03; CC vs. GG, OR=3.10, 95%CI=1.08 to 8.87, p=0.04) of HBV infection while there was no association found in the asymptomatic and acute stages of infection (Table 4.25)

Table 4.24: Genotypic and allelic frequencies of ATG5 polymorphism rs2245214 (C/G) in HBV infected cases and healthy control samples. (N = total number of samples)

ATG5 SNP rs ID	**HBV**	**Control**	**Odds ratio**	**p-value**
Genotype and allele	**(N=550)**	**(N =250)**	**[95% CI]**	
CC	197[36]	107[43]	Ref.	/
CG	275[50]	121[48]	1.23 [0.8982 to 1.6966]	0.19
GG	78[14]	22[9]	**1.93** [1.1352 to 3.2666]	**0.02**
CG+GG	353[64]	143[57]	1.34 [0.9881 to 1.8193]	0.06
CG+CC	472[86]	228[91]	Ref.	/
GG	78[14]	22[9]	**1.71** [1.0400 to 2.8204]	**0.03**
C	669[61]	335[67]	Ref.	/
G	431[39]	165[33]	**1.30** [1.0473 to 1.6336]	**0.02**

p value <0.05 was considered significant (shown in bold)

4.12.4 Comparison of virological and biochemical parameters with different genotypes

We observed a wide variation in average viral load in patients with different genotype in all categories of HBV infection. In asymptomatic stage, mean viral load was found to be low, while in acute, chronic and cirrhosis stages, we found higher mean viral load in patients with homozygous mutant genotype (GG). Moreover, we found a significant difference between the mean viral load of patients with CC and GG genotype. However, we did not observe any significant difference in mean AST, ALT and Bilirubin levels of patients with respect to their genotypes (Table 4.26).

Table 4.25: Genotypic and allelic frequencies of ATG5 nsSNP rs2245214 (C/G) in different HBV infection stages.

HBV infection categories (no of samples)	Genotype and Allele	Cases (N=550)	Control (N=250)	Odds ratio [95% CI]	p-value
I ASYMPTOMATIC (140)	CC	52/140	107/250	Ref.	/
	CG	74/140	121/250	1.26 [0.8106 to 1.9536]	0.31
	GG	14/140	22/250	1.31 [0.6201 to 2.7652]	
	CG+GG	88	143	1.27 [0.8282 to 1.9361]	0.48
	CC+CG	126	228	Ref.	0.28
	GG	14	22	1.15 [0.5692 to 2.3294]	0.69
	C	178/280	335/500	Ref.	/
	G	102/280	165/500	1.16 [0.8562 to 1.5808]	0.33
II ACUTE (114)	CC	41/114	107/250	Ref.	/
	CG	58/114	121/250	1.25 [0.7764 to 2.0157]	0.36
	GG	15/114	22/250	1.78 [0.8417 to 3.7615]	0.13
	CG+GG	73	143	1.33 [0.8433 to 2.1046]	0.22
	CC+CG	99	228	Ref.	/
	GG	15	22	1.57 [0.7818 to 3.1539]	0.20
	C	140/228	335/500	Ref.	/
	G	88//228	165/500	1.28 [0.9218 to 1.7669]	0.14
III CHRONIC (256)	CC	93/256	107/250	Ref.	/
	CG	121/256	121/250	1.15 [0.7907 to 1.6742]	0.46
	GG	42/256	22/250	**2.20** [1.2226 to 3.9463]	**0.01**
	CG+GG	163	143	1.31 [0.9175 to 1.8746]	0.14
	CC+CG	214	228	Ref.	/
	GG	42	22	**2.03** [1.1752 to 3.5202]	**0.01**
	C	307/512	335/500	Ref.	/
	G	205/512	165/500	**1.36** [1.0486 to 1.7529]	**0.02**
IV CIRRHOSIS (40)	CC	11/40	107/250	Ref	/
	CG	22/40	121/250	1.77 [0.8196 to 3.8163]	0.15
	GG	7/40	22/250	**3.10** [1.0800 to 8.8701]	**0.04**
	CG+GG	29	143	1.97 [0.9432 to 4.1258]	0.07
	CC+CG	33	228	Ref.	/
	GG	7	22	2.20 [0.8712 to 5.5471]	0.09
	C	44/80	335/500	Ref.	/
	G	36/80	165/500	**1.66** [1.0296 to 2.6800]	**0.03**

p value <0.05 was considered significant. Significant values are shown in bold

Table 4.26: Analysis of the biochemical parameters (AST, ALT, and Bilirubin) in HBV infected patients by rs2245214 (C/G) genotype

HBV infection category (N=no of samples)	Genotypes (n=no of samples)	Viral load (copies/ml) [Mean±SD]	p value	AST (IU/L) [Mean±SD]	p value	ALT (IU/L) [Mean±SD]	p value	Bilirubin (mg/dL) [Mean±SD]	P value
Asymptomatic (N=140)	CC (n=52)	123±117	0.66*	61±52	0.72*	55±73	0.66*	0.54±0.3	0.37*
	CG (n=74)	132±108	0.71**	58±41	0.44**	49±78	0.34**	0.58±0.2	0.91**
	GG (n=14)	110±106	0.78#	50±20	0.61#	36±15	0.51#	0.53±0.2	0.35#
	CG+GG (n=88)	128±107	0.56##	57±39	0.42##	47±72	0.46##	0.58±0.2	0.76##
	CG+CC (n=126)	128±112	0.97†	60±46	0.92†	51±76	0.92†	0.56±0.3	0.85†
Acute (N=114)	CC (n=41)	1242±1242	0.10*	73±76	0.71*	96±82	0.28*	0.82±0.3	0.09*
	CG (n=58)	1718±1517	0.10**	39±26	0.35**	83±35	0.29**	0.97±0.5	0.84**
	GG (n=15)	1891±1449	0.07#	54±26	0.48#	72±41	0.18#	0.84±0.4	0.10#
	CG+GG (n=73)	1753±1495	0.35##	66±27	0.23##	81±36	0.31##	0.94±0.4	0.53##
	CG+CC (n=99)	1520±1422	0.32†	71±53	0.67†	88±59	0.49†	0.91±0.4	0.41†
Chronic (N=256)	CC (n=93)	1926294±3083860	0.73*	116±114	0.23*	117.1±141	0.24*	1.04±.04	0.21*
	CG (n=121)	2104039± 4050473	0.41**	100±68	0.27**	100±62	0.53**	1.13±0.6	0.80**
	GG (n=42)	2430059 ±3673920	0.58#	95±34	0.15#	103±52	0.21#	1.06±0.5	0.32#
	CG+GG(n=163)	2192717± 3939691	0.51##	99±61	0.41##	101±59	0.76##	1.11±0.6	0.72##
	CG+CC (n=214)	2023733± 3636201	0.95†	107±92	0.50†	108±105	0.65†	1.09±0.5	0.77†
Cirrhosis (N=40)	CC (n=11)	169603285±188534786	**0.009***	147±49	0.22*	120±55	0.11*	0.91±0.4	0.81*
	CG (n=22)	3357082456±3796674214	0.71**	185±94	0.77**	172±97	0.37**	0.94±0.3	0.87**
	GG (n=7)	138511072±14932304	**<0.001#**	153±21	0.28#	149±79	0.13#	0.88±0.3	0.93#
	CG+GG (n=29)	2522638023±1881569975	**<0.001##**	176±82	0.57##	166±91	0.25##	0.92±0.3	0.69##
	CG+CC (n=33)	2226041460±1089146366	**0.005**†**	171±82	0.71†	120±55	0.06†	0.93±0.3	0.99†

* Student's t test (CC vs. CG), **(CC vs. GG), # (CC+CG+GG), ## (CG+CC vs. GG), † one-way ANOVA - (b/w all pairs), p value <0.05 was considered significant, ALT: Alanine aminotransferase; AST: Aspartate aminotransferase.

4.13 Genotyping of SNP rs12212740 G/A

The dbSNP was used to retrieve the information of the non-coding SNP rs12212740 G/A to design specific primers that contained a restriction site for the enzyme *Pvu*II at the position of the SNP. The primers pair was checked for their specificity using the BLAST tool [232] and IDT OligoAnalyzer tool [233]. The Table 4.27 shows the primer sequences, amplified product size and expected fragment sizes after RFLP.

Table 4.27: The primer sequences and expected banding pattern after RFLP for SNP rs12212740 G/A. The mutant base in the forward primer is shown in the lowercase letter.

ATG5 SNP	**Type**	**Primer sequences 5'− 3'**	**Amplified product size (bp)**	**Restriction enzyme**	**Genotypes (bp)**
rs12212740 G/A	Intronic variant	Forward primer $C_{(15)}$ $AAGA_{(6)}$ CCGCCG ATTTTTAAACAAAATGTTAGCAGcT Reverse primer AACCTTGGGCAAGTTATTTGA	557	*Pvu*II	GG-504, 53 GA-557, 504, 53 AA-557

4.13.1 Optimized PCR-RFLP conditions

PCR was carried using the designed primers for genotyping the SNP rs12212740 G/A. For selecting the optimum annealing temperature for the primer set, gradient PCR was carried out initially. For this, the PCR was carried out in 11 different reaction tubes kept at gradient annealing temperatures, from 48°C to 68°C. Amplified PCR products were analyzed on 1.5% agarose gel with EtBr. We obtained a single 557 bp band at 52°C which was selected as the optimum annealing temperature for amplification of the region containing the SNP as shown in Figure 4.14.

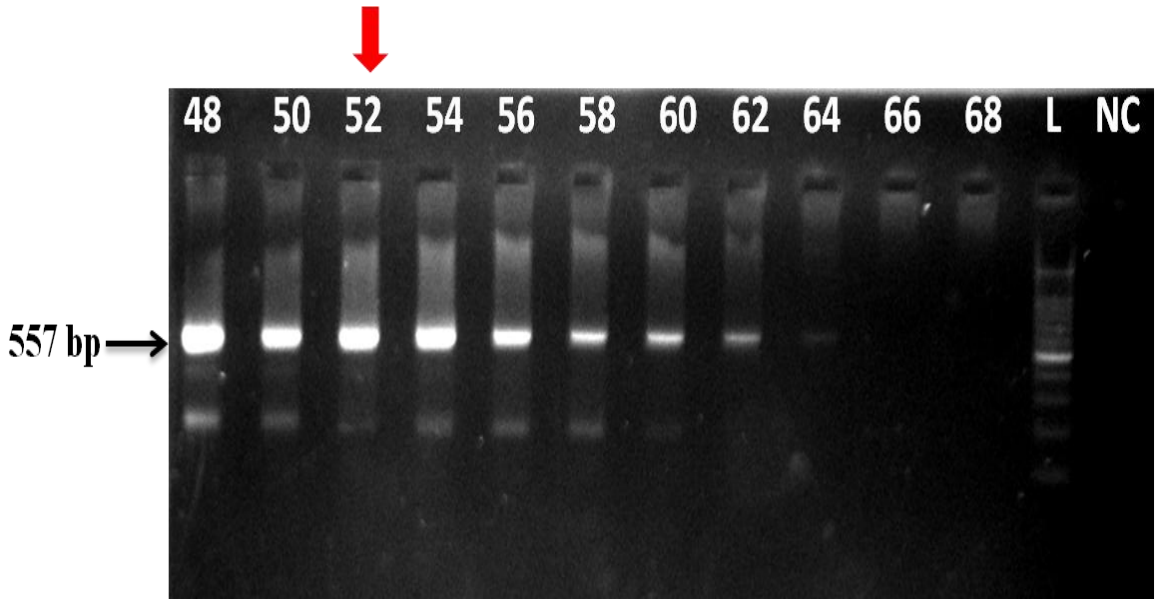

Figure 4.14: Agarose gel image for the amplified products obtained after gradient PCR for non-coding SNP rs12212740 G/A. Lane 1 to lane 11 represents the gradient temperature in increasing order from 48°C to 68°C. L represents 100 bp DNA ladder and NC in lane 13, is the negative control (without DNA template) used in the reaction. We obtained a bright 557 bp with slightest primer dimer at 52°C and hence it was considered to the optimum annealing temperature for the PCR.

4.13.2 PCR-RFLP results

The optimized PCR conditions were used to amplify DNA segments from HBV infected samples and healthy control samples. The optimized PCR conditions are shown in Table 4.28. The PCR products were digested with *Pvu*II restriction enzyme (3 units, 37°C overnight incubation). We observed all the three genotypes; Homozygous wild-type GG (504 bp and 53 bp), heterozygous genotype GA (557 bp, 504 bp and 53 bp) and homozygous mutant genotype AA (557 bp) in different HBV infected as well as healthy control samples. The optimized agarose gel picture after RFLP is shown in Figure 4.15.

Table 4.28: The optimized PCR conditions for genotyping SNP rs12212740 G/A

PCR Stage	Temperature	Duration
Initial denaturation	95°C	2 minutes
Denaturation	95°C	45seconds
Annealing	52°C	1 minute
Extension	72°C	45 seconds
Final extension	72°C	5 minutes

30 cycles

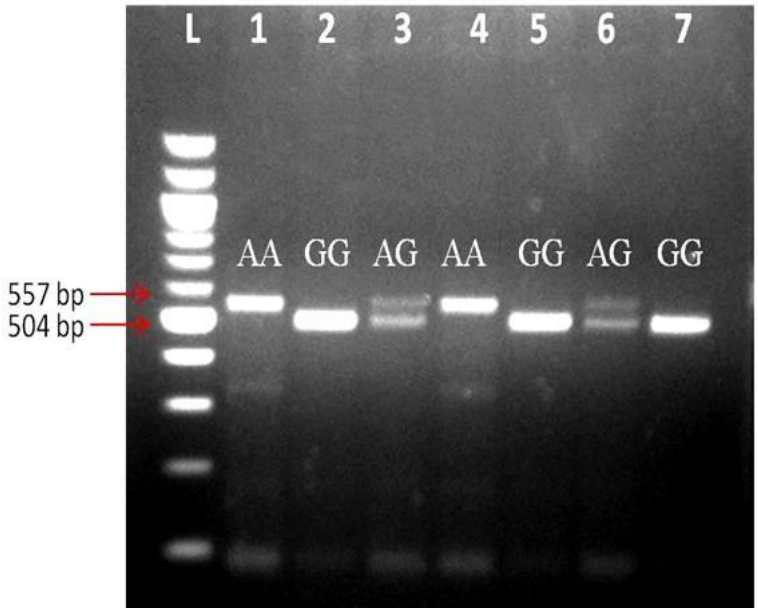

Figure 4.15: The representative PCR-RFLP agarose gel picture for genotyping SNP rs12212740 G/A in HBV infected as well as healthy control samples. L represents a 100 bp DNA ladder; Lane marked as 1 to lane 7 represents the *Pvu*II digested PCR products from HBV infected samples.

4.13.3 Allelic and Genotypic frequencies

We did not observe any significant association for rs12212740 (G/A) SNPs with HBV infection susceptibility. Allelic and Genotypic frequencies are shown in Table 4.29. Moreover, no significant association was found in different categories of HBV infection (Table 4.30).

Table 4.29: Genotypic and allelic frequencies of ATG5 polymorphism rs12212740 (G/A) in HBV cases and healthy control samples (N = total number of samples).

ATG5 SNP rs ID Genotype and allele	HBV (N=550)	Control (N =250)	Odds ratio [95% CI]	p-value
rs12212740 (G>A)				
GG	500[90.91]	232[92.8]	Ref.	\
GA	48[8.73]	17[6.8]	1.31 [0.7374 to 2.3275]	0.36
AA	2[0.36]	1[0.4]	0.92 [0.0827 to 10.1642]	0.95
GA+AA	50[9.09]	18[7.2]	1.29 [0.7356 to 2.2582]	0.38
GA+GG	548[99.64]	249[99.6]	Ref.	\
AA	2[0.36]	1[0.4]	0.91 [0.0820 to 10.0693]	0.94
G	1054[96]	481 [96]	Ref.	\
A	46 [4]	19 [4]	1.10 [0.6405 to 1.9060]	0.72

p-value < 0.05 was considered significant

4.13.4 Comparison of virological and biochemical parameters with different genotypes

We observed a significant difference in the mean viral load of patients with the wild-type (GG) and mutant genotype (AA) of the patient (acute stage). The average viral load was observed to be higher in the patient with homozygous mutant genotype (GG vs. AA, 1802± 1614 [n=105] vs. 604±126 [n=8], p=0.03; GG vs. GA+AA, 1802±1614 [n=105] vs. 600±118 [n=9], p=0.04). Moreover, the patients with AA genotype in the acute stage of infection had higher mean ALT (GG vs. GA, 66±34 [n=105] vs. 107±142 [n=8], p= 0.023) and mean AST levels (GG vs. GA, 84±44[n=105] vs. 128±147[n=8], p= 0.035).We also found significantly higher mean Bilirubin level in cirrhosis patients with GA genotype (GG vs. GA, 171±78 [n=36] vs. 142±14 [n=4], p=0.001) (Table 4.31).

Table 4.30: Genotype frequency distribution of ATG5 polymorphism rs12212740 (G/A) in different HBV infection stages

HBV infection categories	Genotype and Allele	Cases (N=550)	Control (N=250)	Odds ratio [95% CI]	p-value
		n/N	n/N		
	GG	125/140	232/250	Ref.	/
	GA	14/140	17/250	1.53 [0.7292 to 3.2037]	0.26
I	AA	1/140	1/250	1.86 [0.1151 to 29.9282	0.66
ASYMPTOMATIC	AA+GA	15	18	1.55 [0.7537 to 3.1740]	0.23
(140)	GA+GG	139	249	Ref.	/
	AA	1	1	1.79 [0.1112 to 28.8640]	0.68
	G	264	481	Ref.	/
	A	16	19	1.53[0.7759 to 3.0339]	0.22
	GG	105/114	232/250	Ref.	/
	GA	8/114	17/250	1.04 [0.4350 to 2.4853]	0.93
	AA	1/114	1/250	2.21 [0.1369 to 35.6663]	0.58
II	AA+GA	9	18	1.10 [0.4804 to 2.5405]	0.81
ACUTE	GA+GG	218	249	Ref.	/
(114)	AA	1	1	2.20 [0.1366 to 35.5459]	0.58
	G	105	481	Ref.	/
	A	9	19	1.16 [0.5311 to 2.5391]	0.71
	GG	234/256	232/250	Ref.	/
	GA	22/256	17/250	1.28 [0.6642 to 2.4786]	0.45
	AA	0/256	1/250	0.33 [0.0134 to 8.1550]	0.50
III	AA+GA	22	18	1.21 [0.6333 to 2.3186]	0.56
CHRONIC	GA+GG	256	249	Ref.	/
(256)	AA	0	1	0.32 [0.0131 to 7.9975]	0.49
	G	490	481	Ref.	/
	A	22	19	1.14 [0.6074 to 2.1269]	0.69
	GG	36/40	232/250	Ref.	/
	GA	4/40	17/250	1.52 [0.4828 to 4.7621]	0.48
	AA	0/40	1/250	2.12 [0.0849 to 53.1224]	0.64
IV	AA+GA	4	18	1.43 [0.4585 to 4.4729]	0.54
CIRRHOSIS	GA+GG	40	249	Ref.	/
(40)	AA	0	1	2.05 [0.0822 to 51.2845]	0.66
	G	76	481	Ref.	/
	A	4	19	1.33[0.4413 to 4.0230]	0.61

p value <0.05 was considered. Significant values are shown in bold.

Table: 4.31: Analysis of the biochemical parameters (AST, ALT, and Bilirubin) in HBV infected patients with ATG5 polymorphism rs12212740 (G/A) genotype.

HBV infection category (N=no of samples)	Genotypes (n=no of samples)	Viral load (copies/ml) [Mean±SD]	p value	AST (IU/L) [Mean±SD]	p value	ALT (IU/L) [Mean±SD]	p value	Bilirubin (mg/dL) [Mean±SD]	p value
Asymptomatic (N=140)	GG (n=125)	131 ± 113	0.10*	59±46	0.81*	49±74	0.56*	0.56±0.3	0.63*
	GA (n=14)	86 ± 75	0.15#	56±26	0.81#	61±57	0.65#	0.60±0.2	0.71#
	AA (n=1)	28	0.21†	51±0	0.99†	24±0	0.91†	0.40±0	0.94†
	GA+AA (n=15)	82 ± 74		56±25		58±55		0.59±0.2	
	GA+GG(n=139)	127 ± 111		59±44		50±72		0.56±0.3	
Acute (N=114)	GG (n=105)	1802 ± 1614	**0.03***	66±34	**0.02***	84±44	**0.04***	0.90±0.4	0.63*
	GA (n=8)	604 ± 126	**0.04#**	107±142	0.06#	128±147	0.19#	0.83±0.3	0.61#
	AA (n=1)	568	**0.03**†	24±0	0.09†	23±0	0.13†	0.90±0	0.92†
	GA+AA (n=9)	600 ± 118		98±136		116±142		0.83±0.3	
	GA+GG(n=113)	1717 ± 1585		69±50		87±57		0.90±0.4	
Chronic (N=256)	GG (n=234)	2138212 ± 3761332	0.42*	105±88	>0.999*	109±102	0.39*	1.09±0.5	>0.999*
	GA (n=22)	1487651 ± 1748828		105±37		90±44		1.09±0.5	
	AA (n=0)	0		0		0		0	
Cirrhosis (N=40)	GG (n=36)	2016656546 ± 10594197937	0.75*	171±78	0.47*	152±89	0.90*	0.86±0.3	0.001*
	GA (n=4)	352635053 ± 47398398		142±14		158±20		1.40±0.1	
	AA (n=0)	0		0		0		0	

* Student's t test (GG vs. GA), **(GG vs. CC), # (GG+GA+AA), ## (GA+GG vs. AA),† one-way ANOVA (b/w all pairs), p value <0.05 was considered significant (shown in bold), ALT: Alanine aminotransferase; AST: Aspartate aminotransferase.

4.14 Genotyping of SNP rs510432 A/G

The dbSNP was used to retrieve the information of the promoter SNP rs510432 A/G to design specific primers that contained a restriction site for the enzyme *Nde*I at the position of the SNP. The primers pair was checked for their specificity using the BLAST tool [232] and IDT oligoanalyzer [233]. Table 4.32 shows the primer sequences, size of PCR product and different fragments after RFLP.

Table 4.32: The primer sequences amplified product size, restriction enzyme and expected banding pattern for SNP rs510432 A/G.

ATG5 SNP	**Type**	**Primer sequences 5'– 3'**	**Amplified product size (bp)**	**Restriction enzyme**	**Genotypes (bp)**
rs510432 A/G	Promoter variant	Forward primer $A_{(15)}C_{(15)}$TCCAACAAAGTAGAGAAGAAGATCAtAT Reverse primer TCTACCCTCTTCTGAGAATCTTG	260	*Nde*I	GG-205, 55 GA-260, 205, 55 AA-260

4.14.1 Optimized PCR-RFLP conditions

PCR was carried using the designed primers for genotyping the SNP rs510432 A/G. For selecting the optimum annealing temperature and primer concentration, gradient PCR was carried out initially. For this, the PCR was carried out in 6 different reaction tubes kept at gradient annealing temperatures, from 50°C to 60°C. Amplified PCR products were analyzed on 1.5% agarose gel with EtBr. Although we obtained bright 260 bp bands at every temperature, we selected 56°C as the optimum annealing temperature for amplification of the region containing the SNP as shown in Figure 4.16.

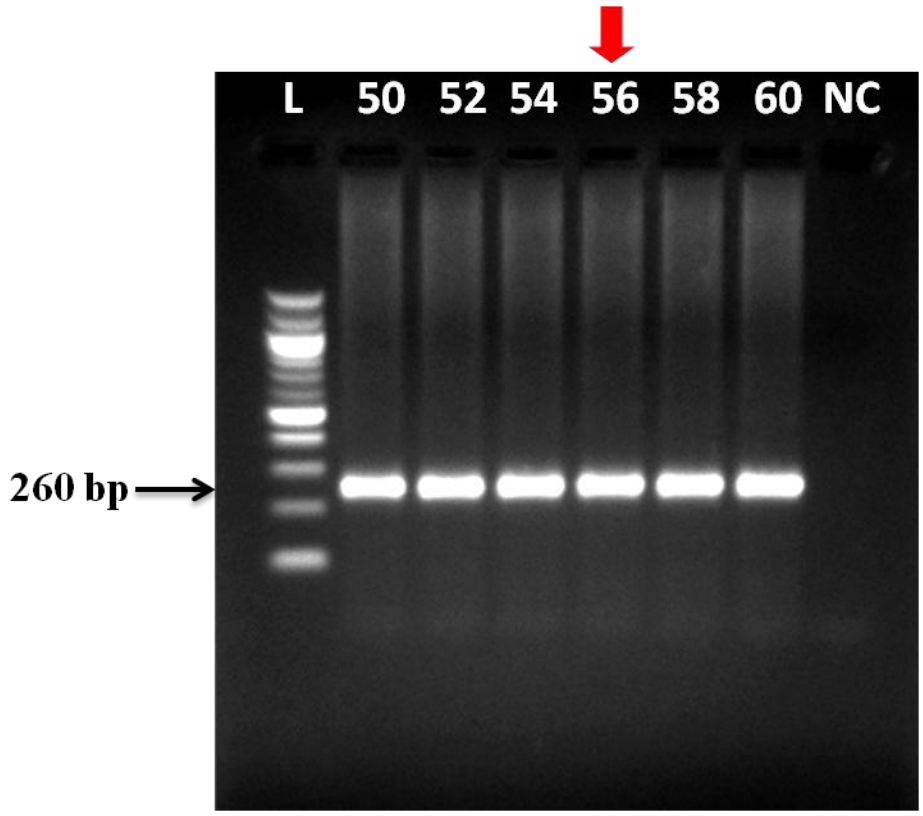

Figure 4.16: Agarose gel image for the amplified products obtained after gradient PCR for the promoter SNP rs510432 A/G. L represents 100 bp DNA ladder, Lane 2 to lane 6 represents the gradient temperature in increasing order from 50°C to 60°C. NC in lane 8, is the negative control (without DNA template) used in the reaction. We obtained a bright 260 bp without primer dimer formation at 56°C and hence it was considered to the optimum annealing temperature for the PCR.

4.14.2 PCR-RFLP results

The optimized PCR conditions were used to amplify DNA segments from HB samples and healthy control samples. The optimized PCR conditions are shown in Table 4.33. The PCR products were digested with *Nde*I restriction enzyme. We observed all the three genotypes; Homozygous wild-type GG (205 bp and 55 bp), heterozygous genotype GA (260 bp, 205 bp and 55 bp) and homozygous mutant genotype AA (260 bp) in different HBV infected as well as healthy control samples. The representative agarose gel picture is shown in Figure 4.17

Table 4.33: The optimized PCR conditions for genotyping SNP rs510432 A/G.

PCR Stage	Temperature	Duration
Initial denaturation	95°C	2 minutes
Denaturation	95°C	40 seconds
Annealing	56°C	1 minute
Extension	72°C	40 seconds
Final extension	72°C	5 minutes

30 cycles

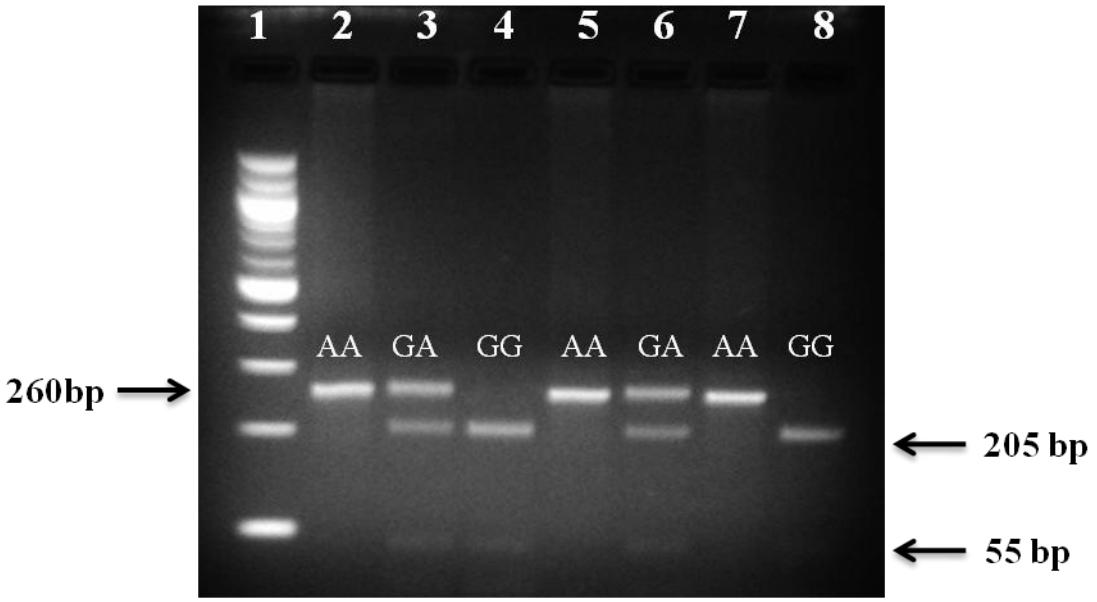

Figure 4.17: The representative PCR-RFLP agarose gel picture for genotyping SNP rs510432 A/G in HBV infected as well as healthy control samples. L represents a 100 bp DNA ladder; Lane 2 to Lane 8 represents the *Nde*I digested PCR products from HBV infected samples.

4.14.3 Allelic and Genotypic frequencies

No significant association was observed for the promoter SNP rs510432 (G>A), in the allelic model of polymorphism. However, the genotypic frequencies reveal its association with HBV susceptibility in the heterozygous (GG vs. GA, OR=1.66, 95%CI=1.11- 2.487, p=0.01; estimated power 22%) and dominant model (GG vs. GA+AA, OR=1.62, 95%CI=1.10- 2.40), p=0.02) of polymorphism (Table 4.34). This promoter SNP was also found to be significantly associated in asymptomatic (heterozygous model, OR=2, 95%CI=1.10 to 3.65,p=0.02;dominant model, OR=1.8, 95%CI=.025 to 3.36, p=0.04) and in chronic stage GG vs. GA, OR=1.81, 95%CI=1.11 to 2.96,p=0.02; GG vs. GA+AA, OR=1.79, 95%CI=.1.19 to 2.90, p= 0.02) of HBV infection while no association was found in the acute and cirrhosis patients (Table 4.35).

Table 4.34: Genotypic and allelic frequencies of ATG5 polymorphism rs510432 (G>A) in HBV infected cases and healthy control samples. (N = total number of samples)

ATG5 SNP rs ID Genotype and allele	**HBV (N=550)**	**Control (N =250)**	**Odds ratio [95% CI]**	**p-value**
rs510432 (G>A)				
GG	75[13.64]	51[20.4]	Ref.	/
GA	388[70.54]	159[63.6]	**1.66** [1.1115 to 2.4774]	**0.013**
AA	87[15.82]	40[16]	1.48[0.8824 to 2.4790]	0.136
GA+AA	475[86.36]	199[79.6]	**1.62** [1.0964 to 2.4030]	**0.015**
GA+GG	463[84.18]	210[84]	Ref.	/
AA	87[15.82]	40[16]	0.99 [0.6557 to 1.4841]	0.948
G	538[49]	261[52]	Ref.	/
A	562[51]	239 [48]	1.14[0.9232 to 1.4096]	0.222

Significant values ($p < 0.05$) are shown in **bold.**

4.14.4 Comparison of virological and biochemical parameters with different genotypes

We have found a significant difference in the mean viral load of patients with homozygous wild-type genotype (GG) and homozygous mutant genotype (AA) in the asymptomatic category (GG vs. AA, 80.5±71 [n=17] vs. 161±77 [n=17], p=0.004). However, we did not observe a similar trend in other categories of HBV infection, where the higher viral load was found in the wild-type genotype (GG) of patients as compared to the mutant genotype (AA). Also, we did not observe any significant difference in AST, ALT and Bilirubin levels in patients with respect to their genotypes for this SNP (Table 4.36).

Table 4.35: Genotypic and allelic frequencies of ATG5 SNP rs510432 (G/A) in different HBV infection stages and healthy control.

HBV infection categories (no of samples)	Genotype and Allele	Cases (N=550)	Control (N=250)	Odds ratio [95% CI]	p-value
		n/N	n/N [%]		
	GG	17/140	51/250	Ref.	/
	GA	106/140	159/250	**2.00** [1.0961 to 3.6494]	**0.024**
I	AA	17/140	40/250	1.26 [0.5789 to 2.8079]	0.546
ASYMPTOMATIC	AA+GA	123	199	**1.85** [1.0246 to 3.3557]	**0.040**
(140)	GA+GG	123	210	Ref.	/
	AA	17	40	0.73 [0.3944 to 1.3348]	0.302
	G	140/280	261/500	Ref.	/
	A	140/280	239/500	1.09 [0.8149 to 1.4634]	0.555
	GG	18/114	51/250	Ref.	/
	GA	75/114	159/25lll0	1.34 [0.7310 to 2.4436]	0.346
	AA	21/114	40/250	1.49 [0.7002 to 3.1600]	0.302
II	AA+GA	96	199	1.37 [0.7577 to 2.4657]	0.299
ACUTE	GA+GG	93	210	Ref.	/
(114)	. AA	21	40	1.19 [0.6625 to 2.1213]	0.566
	G	111/228	261/500	Ref.	/
	A	117/228	239/500	1.15 [0.8414 to 1.5748]	0.379
	GG	32/256	51/250	Ref.	/
	GA	181/256	159/250	1.81 [1.1107to 2.9635]	**0.02**
	AA	43/256	40/250	1.71 [0.9244 to 3.1755]	0.09
III	AA+GA	224	199	1.79 [1.1084 to 2.9035]	**0.02**
CHRONIC	GA+GG	213	210	Ref.	/
(256)	. AA	43	40	1.06 [0.6618 to 1.6972]	0.80
	G	245/512	261/500	Ref.	/
	A	267/512	239/500	1.19 [0.9299 to 1.5230]	0.17
	GG	7/40	51/250	Ref.	/
	GA	26/40	159/250	1.19 [0.4882 to 2.9076]	0.70
	AA	7/40	40/250	1.28 [0.4133 to 3.9331]	0.67
IV	AA+GA	33	199	1.20 [0.5053 to 2.8887]	0.67
CIRRHOSIS	GA+GG	33	210	Ref.	/
(40)	. AA	7	40	1.11 [0.4606to 2.6929]	0.81
	G	40/80	261/500	Ref.	/
	A	40/80	239/500	1.09 [0.6811 to 1.7509]	0.71

Significant values ($p<0.05$) are shown in **bold**.

Table 4.36: Analysis of the biochemical parameters (AST, ALT, and Bilirubin) in HBV infected patients with ATG5 polymorphism rs510432 (G/A) genotype.

HBV infection category (N=no of samples)	Genotypes (n=no of samples)	Viral load (copies/ml) [Mean±SD]	p value	AST (IU/L) [Mean±SD]	p value	ALT (IU/L) [Mean±SD]	p value	Bilirubin (mg/dL) [Mean±SD]	p value
Asymptomatic (N=140)	GG (n=17)	81 ± 71	0.11*	60±28	0.93*	52±50	0.87*	0.63±0.2	0.09*
	GA (n=106)	128 ±118	**0.004****	59±44	0.35**	49±74	0.44**	0.54±0.2	>0.999**
	AA (n=17)	161 ±77	0.07#	52±21	0.85#	39±46	0.91#	0.63±0.3	0.36#
	GA+AA (n=96)	132 ±114	0.16##	58±42	0.50##	50±71	0.62##	0.56±0.3	0.23##
	GA+GG (n=93)	121±114	0.30†	59±42	0.97†	48±71	0.98†	0.56±0.2	0.46†
Acute (N=114)	GG (n=18)	1401±1491	0.63*	56±30	0.19*	64±34	0.06*	0.90±0.4	0.92*
	GA (n=75)	1578±1390	0.60**	73±53	0.61**	93±63	0.23**	0.89±0.4	0.79**
	AA (n=21)	1658±1557	0.60#	63±50	0.25#	80±45	0.08#	0.94±0.5	>0.999#
	GA+AA (n=96)	1595±1419	0.75##	71±53	0.56##	90±60	0.57##	0.90±0.4	0.62##
	GA+GG (n=93)	1547±1401	0.98†	70±50	0.73†	88±60	0.40†	0.89±0.4	0.99†
Chronic (N=256)	GG (n=32)	3341376±8903074	0.27*	100±84	0.62*	122±141	0.60*	1.11±0.5	0.84*
	GA (n=181)	2284619±3955857	0.80**	109±97	>0.99**	110±113	0.39**	1.09±0.5	0.57**
	AA (n=43)	2858354±7940397	0.38#	100±37	0.68#	102±48	0.50#	1.05±0.4	0.75#
	GA+AA (n=224)	2399366±4987396	0.67##	107±89	0.64##	108±103	0.58##	1.08±0.5	0.62##
	GA+GG (n=213)	2455472±5066631	0.85†	107±95	0.97†	112±117	0.94†	1.09±0.5	0.97†
Cirrhosis (N=40)	GG (n=7)	321699032±458449347	0.99*	132±15	0.24*	145±33	0.97*	1.02±0.3	0.53*
	GA (n=26)	275928476±12610352832	0.88**	169±81	0.15**	146±78	0.77**	0.94±0.3	0.36**
	AA (n=7)	288714934±337461164	0.65#	182±84	0.18#	159±120	0.93#	0.84±0.4	0.423#
	GA+AA (n=33)	2265170800±11276081185	0.99##	174±81	0.53##	148±88	0.72##	0.92±0.3	0.37##
	GA+GG (n=33)	2271767620±11275629199	0.94†	162±74	0.71†	146±80	0.998†	0.96±0.3	0.83†

* Student's t test (GG vs. GA), **(GG vs. AA), # (GG+GA+AA), ## (GA+GG vs. AA), † one-way ANOVA (b/w all pairs), ALT: Alanine aminotransferase; AST: Aspartate aminotransferase. $p<0.05$ was considered significant.

4.15 Haplotype and linkage disequilibrium analysis

As we have the complete genotype data of only four SNPs out of the total six SNPs of ATG5, therefore, we performed haplotype analysis of only four SNPs (rs34793250, rs2245214, rs12212740 and rs510432). Five haplotypes were predicted in our study population out of which three haplotypes TCGA (OR=1.74, p <0.05), TGGG (OR=1.67, p <0.05) and TGGA (OR=3.97, p value <0.05) were found to be significantly associated with HBV infection susceptibility as shown in Table 4.37. Linkage disequilibrium (LD) analysis showed a strong linkage between the nsSNP rs34793250 and the SNP rs12212740 (D'=0.99) The LD plot obtained for the four SNPs is shown in Figure 4.18.

Table 4.37: Haplotype analysis of 4 SNPs of *ATG5* gene genotyped in our population cases and controls (significant p-value is shown in bold)

Haplotype	Case(freq)	Control(freq)	OR [95% CI]	p-value
TCGA	372(0.338)	128(0.353)	**1.74 [1.378~2.197]**	**<0.05**
TGGG	283(0.257)	97(0.267)	**1.67 [1.289~2.156]**	**<0.05**
TCGG	218(0.198)	94(0.259)	1.24 [0.946~1.613]	0.127
TGGA	140(0.127)	20(0.055)	**3.97 [2.454~6.411]**	**<0.05**
TCAA	42(0.038)	14(0.038)	1.56 [0.844~2.88]	0.195

Significant values (p<0.05) are shown in **bold.**

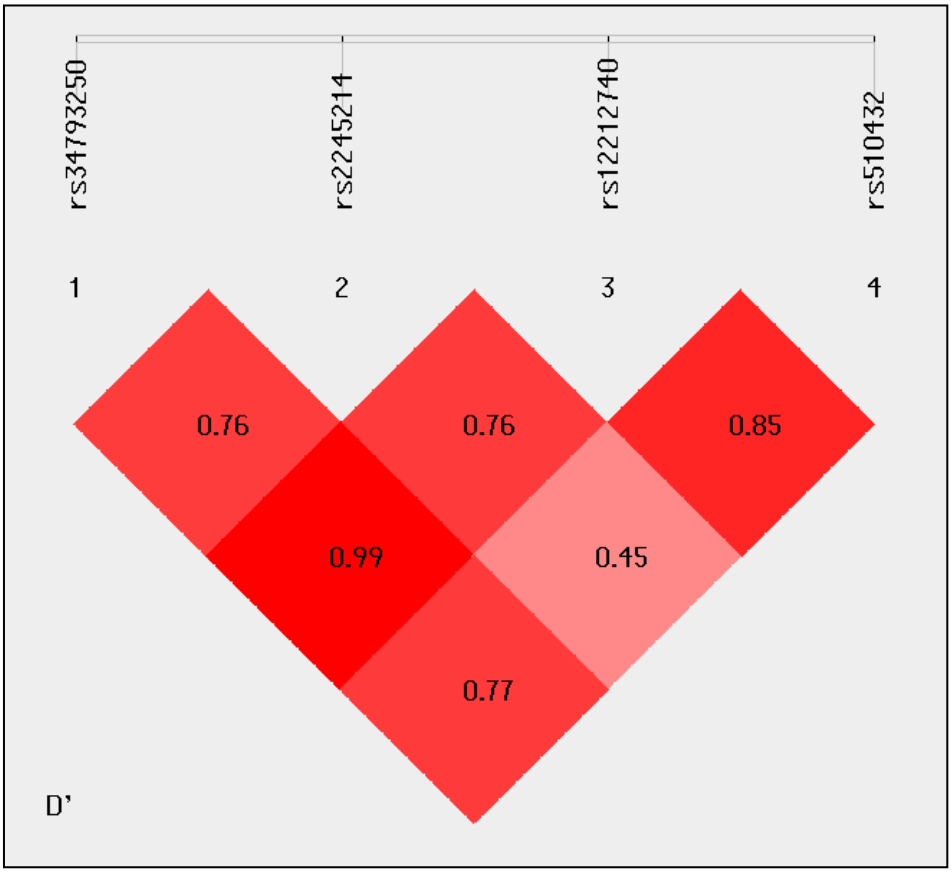

Figure 4.18: Linkage disequilibrium among 4 SNPs (rs34793250, rs2245214, rs12212740 and rs510432) of *ATG5* gene. In this LD plot, the nsSNP rs34793250 and the non-coding SNP rs12212740 has shown significant correlation (D'=0.99) of being in linkage disequilibrium.

Chapter 5

DISCUSSION

Discussion

5.1 *In silico* analysis

5.1.1 Prediction of deleterious nsSNPs, other significant genetic variants, and haplotypes that exist in linkage disequilibrium

nsSNPs is single nucleotide variation in DNA sequence that leads to change in the amino acid of the encoded protein. The substituted amino acid may prove deleterious to protein function by altering the protein structure and stability [234]. Subsequently, these polymorphisms could be damaging to the protein structure and function. However, not all nsSNPs are detrimental and lead to disease development. There are a few nsSNPs which lead to structural changes in the protein but do not disrupt the overall protein function. Such nsSNPs are referred as tolerant [235]. We have used SIFT tool in order to predict nsSNPs in ATG5 that could have a deleterious impact on ATG5 function. Three nsSNPs predicted ATG5, rs34793250 [M129V] (T/C), rs77859116 [I65V] (T/C) and rs115576116 [A95D] (G/T) were predicted to have tolerated impact by SIFT and PolyPhen-2. However, the same nsSNPs were found to be deleterious by PredictSNP. Therefore, these nsSNPs were selected to further analyze their effect on the protein structure and function in our objective 2 and were genotyped to find out their association with HBV infection susceptibility.

LD analysis holds great significance in human genetics as it provides information about the past evolutionary events, the biology of recombination and provides a base for mapping genes in humans and other species [177]. Moreover, LD uncovers the correlation among neighboring alleles and reveals the cluster of genetic variations that are inherited together [236]. The groups of such genetic variants called “haplotypes” are extremely valuable for the genotype-phenotype investigation. Thus, in our study, we predicted linked genetic variants or haplotypes of ATG5 that are in linkage disequilibrium (Figure 4.1)**.** We have obtained, three distinct alternative blocks of different sizes (38 kb, 58 kb and 15 kb). Also, a significant association was observed among the blocks. We have obtained a total of 26 markers in the block 1, 39 markers in the block 2 and 16 markers in the block 3. The information revealed by these LD patterns is implicated in association mapping to find out their link with the disease susceptibility. The genetic variants that are observed to have a higher frequency in diseased individuals as compared to the healthy individuals may increase an individual’s

susceptibility to that particular disease. Hence, such SNPs could be identified for risk analysis in that population.

5.1.2 Putative functional sites in ATG5

Posttranslational modifications (PTMs) are believed to have an indispensable role in regulating a variety of cellular events and have been shown to influence protein-protein interactions [237]. Moreover, any alterations in PTMs may result in modulation of regulatory pathway ultimately leading to manifestation in form of a disease [238]. In this context, amino acid substitutions (nsSNPs) may lead to PTM alterations by disrupting the PTM sites and may result in disease development [239]. In a study, a mutation in cyclin D1 has been shown to disrupt the phosphorylation that leads to nuclear accumulation contributing to genesis and progression of neoplastic growth [240]. Similarly, there are various other studies that analyze and link PTM associated mutations with disease development [241-243]. Therefore, we have analyzed PTM sites in ATG5 protein annotated over protein domain structure along with predicted nsSNPs (Figure 4.2). In this context, we have obtained 11 phosphorylation sites in ATG5 through the NetPhos algorithm. Phosphorylation has been already demonstrated to influence ATG5 function. According to a recent study, phosphorylation of ATG5 is mediated by p38 leads to the inhibition of autophagy [153]. Also, we have obtained 2 palmitoylation sites in ATG5 through CSS-PALM server. Palmitoylation has been shown to regulate proteins trafficking, compartmentalization, and membrane interactions and various cellular events including apoptosis and cellular signaling [244, 245]. By using BDM-PUB web-based tool, we obtained 7 ubiquitination sites in ATG5. Ubiquitination is a highly dynamic and enzymatically catalyzed posttranslational modification that influences various processes, like cell cycle, protein trafficking, and even the immune responses [246]. On the other hand, deregulation of ubiquitination has major consequences and is involved in cancer, various neurodegenerative diseases, metabolic syndrome and other infections [247]

Furthermore, we have predicted a total of 4 Calpain–cleavage sites in ATG5 with GPS-CCD 1.0 tool. Calpain-mediated cleavage has been demonstrated to play a significant role in several cellular events like cell death/apoptosis [144], cytoskeleton remodeling [248], and the cell cycle [249]. Calpain activation and cleavage have also been shown to influence the function of ATG5. In this context, ATG5 has been shown to switch autophagy into apoptosis process after its Calpain cleavage [144].

5.1.3 nsSNPs were found to destabilize the ATG12–ATG5/ATG16L1 protein complex

ATG12–ATG5/ATG16L1 complex is indispensable for the process of autophagy. ATG12–ATG5/ATG16L1 has an essential role in autophagosome formation. A recent study showed a mutation in ATG5, E122D inhibits the conjugation between ATG5 and ATG12. Moreover, ATG5 expression and autophagy flux were also found to be reduced [49]. The study emphasizes the need to identify such disease-causing variants in order to understand the pathogenic mechanism of various human diseases. Therefore we performed sequence and structure-based analysis of the predicted nsSNPs (rs34793250, rs77859116, and rs115576116). The results indicated the degree of destabilization caused, based on decreased Gibbs free energy change values obtained from different tools. Both at the sequence and structural level, we found that all three nsSNPs decreased the stability of the ATG5 protein as predicted from the change in Gibbs free energy values (Table 4.6 & 4.7).

Moreover, these nsSNPs resulted in loss or replacement of stabilizing residues in the protein complex. Though all the predicted nsSNPs increased the energy of the complex, the mutation M129V was observed to increase the energy to the maximum. Furthermore, we predicted the change in binding affinity of ATG5 towards ATG12 and ATG16L1 proteins. It was observed that all the mutations lead to an overall decrease in binding affinity of ATG5 with ATG12 and ATG16L1. On the whole, all the results from different parameters indicated that these mutations in ATG5 are likely to have a potentially damaging impact over the protein as well as its complex structure and could play important role in various diseases.

5.2 Genotyping of predicted nsSNPs and noncoding SNPs in HBV infected patients

The selected nsSNPs rs34793250 [M129V] (T/C) and noncoding SNPs were genotyped in 550 HBV infected patients and 200 healthy control samples to find out their role in HBV infection susceptibility. The target region was successfully amplified through PCR and the genotyping was done by RFLP followed by agarose gel electrophoresis. Statistical analysis was done to find out significant differences in the frequencies of the alleles and the genotypes obtained for selected SNPs in HBV infected as well as healthy control samples. Moreover, virological (viral load) and biochemical parameters (AST, ALT, and Bilirubin) were compared with different genotypes obtained for each SNP.

5.2.1 nsSNPs rs34793250 [M129V] (T/C), and noncoding SNPs rs2245214 (C/G) and rs510432 (G/A) may increase individuals' risk for HBV infection

ATG5 rs34793250 (T/C) [M129V]

The nsSNP rs34793250 (T/C) of ATG5 was found to be the most destabilizing mutation as it increased the total energy of the ATG5-ATG12-ATG16 complex structure to the maximum. Furthermore, this nsSNP was observed to be significantly associated with HBV infection susceptibility. The mutant allele C was found to be significantly linked with HBV risk. Genotype models of polymorphism were also analyzed to compare genotypic frequencies. A significant association was found in the Dominant and heterozygous models of polymorphism. In order to evaluate the role of the mutant allele in the progression of HBV infection, we compared the observed allelic and genotypic frequencies with different HBV infection categories. Significant association of C allele was observed in the asymptomatic, acute and chronic stages of HBV infection. Whereas, no association of this SNP was found in cirrhosis patients (Table 4.14). However, the odds ratio did not seem to increase with the increase in the severity of HBV infection, indicating no significant role played by the nsSNP in HB progression. Furthermore, viral load in different HBV infection categories was co-related with the genotypes of HB patients. Also, there was no significant difference observed in the mean viral load between the host genotypes for this SNP. Although not significant, the low viral load was found in patients with homozygous mutant genotype CC as compared to homozygous wild-type genotype TT, in all categories of HBV infection (Table 4.15). Additionally, statistical analysis was carried out to study the relation between the patient genotypes and biochemical parameters namely, AST, ALT, and Bilirubin levels (Table 4.15). The mean AST level was found to be significantly high in patients (chronic stage) with heterozygous genotype (TC) than with homozygous wild-type genotype. The same pattern was observed for ALT levels where the patients (asymptomatic and chronic stage) with heterozygous genotype (TC) were found to have high mean ALT level when compared to patients with homozygous wild-type genotype. However, no significant variation was observed in Bilirubin levels in relation to the genotype of patients**.**

ATG5 rs2245214 (C/G)

The noncoding SNP rs2245214 (C/G), is located within an intron, has already been studied in tuberculosis [250] [251], systemic lupus erythematosus [252], non-medullary thyroid cancer

[48] and Paget disease of bone [32]. The SNP was shown to be significantly linked with increased risk of developing non-medullary thyroid cancer and Paget disease of bone. However, in a recent study, no significant association was found for rs2245214 SNP with the distribution and susceptibility of Head and neck squamous cell carcinoma (HNSCC) in Spanish patients [253].

Also, in another study, this SNP was studied for its influence on trained immunity induced by Bacillus Calmette-Guerin (BCG) vaccine, both in vitro and in vivo. [254]. Furthermore, this Intronic polymorphism results in a change of recognition site for a SRp40 protein which is involved in mRNA splicing [32].Therefore, we have selected this SNP to study its association with HBV infection susceptibility. According to our results, we found this SNP rs2245214 to be associated with HBV susceptibility. The G allele was observed to be significantly related to HBV risk as compared to control samples. Also, we observed a significant association in the homozygous and recessive genotype models of polymorphism. The allelic and genotypic frequencies were also calculated in different HBV infection stages. We found the SNP was significantly associated with the chronic and cirrhosis stages of HBV infection. There was wide variation in average viral load in patients with different genotype in all categories of HBV infection. In asymptomatic stage, mean viral load was found to be low, while in acute, chronic and cirrhosis stages, the mean viral load in patients with homozygous mutant genotype (GG) was found to be higher. However, we did not observe any significant difference in mean viral load, AST, ALT, and Bilirubin levels of patients with respect to their genotypes when compared with controls.

ATG5 rs510432 (G/A)

The genetic variant rs510432 (G/A) in ATG5 promoter region, found 335bp upstream of the transcription start site (TSS), is an important polymorphism that has been shown to be significantly linked with childhood asthma [186]. The disease-associated allele was also found to enhance the ATG5 promoter activity. Moreover, STAT1 and C-Fos transcription factors were also identified in the study which has been already studied for their role in asthma. Furthermore, in the same study, ATG5 mRNA expression was also demonstrated to be elevated in the nasal epithelial cells of acute asthmatics [186]. In a recent study, the minor genotype of this promoter SNP rs510432 was linked with an increased risk of atopic diseases in children, emphasizing its role in the development of atopic march in children [255]. In

another study [256], the SNP was demonstrated to be linked with increased melanoma stage and a borderline association was also observed with the presence of non brisk tumor infiltrating lymphocytes (TIL) in patients. The study highlights the potential of the ATG gene polymorphisms to be identified as markers of melanoma risk, progression, and/or therapeutic targets [256]. On the other hand, a very recent study demonstrates the association of the variant genotype (AA) with decreased risk of Coal Workers' Pneumoconiosis (CWP) in a Chinese Population [257]. Another study found out the significant association of the promoter SNP, with a positive response to adalimumab (ADA) therapy in Crohn's disease patients [258]. Recently, the promoter polymorphism was also demonstrated to be associated with predisposition to sepsis progression and mortality in a Chinese Han population. Moreover, the study showed the decrease in ATG5 transcription levels in presence of risk alleles (A) of polymorphism rs510432 by modulating its promoter activity [259]. In our population, however, no significant association was found for this promoter SNP in the allelic model of polymorphism. However, the genotypic frequencies reveal its association with increased HBV infection risk in the heterozygous and dominant models of polymorphism. This promoter SNP was also found to be significantly associated in asymptomatic and in the chronic stage of HBV infection. As we did not observe any increase in odds ratio/risk factor in the chronic stage when compared with asymptomatic stage, we believe that this SNP does not play role in progression of the infection. Moreover, we did not find any significant association in the acute and cirrhosis patients. Also, the mean viral load in patients with homozygous wild-type genotype (GG) in the asymptomatic category ($p=0.003$) was found to be significantly higher. However, we did not observe a similar trend in other categories of HBV infection. Also, AST, ALT and Bilirubin levels in patients were comparable with respect to the genotypes for this SNP in cases and controls.

Therefore, these ATG5 SNPs (nsSNPs rs34793250 [M129V] (T/C), and non-coding SNPs rs2245214 (C/G) and rs510432 (G/A)) could be further explored and investigated in other population to validate their contribution towards HBV infection progression, outcome or treatment response.

5.2.2 No significant association was observed for the nsSNPs rs77859116 [I65V] (T/C), rs115576116 [A95D] (G/T) and non-coding SNP rs12212740 (G/A) of ATG5 with HBV infection.

ATG5 rs77859116 (T/C) [I65V] and ATG5 rs115576116 (G>T) [A95D]

We did not find any significant relation of the nsSNPs rs77859116 (T>C) and rs115576116 (G>T) of ATG5, with HBV infection susceptibility. Also, for these nsSNPs rs77859116 (T>C) and rs115576116 (G>T) we observed only homozygous wild-type genotype TT and GG, respectively, in 300 samples (150 cases + 150 controls for each SNP).

ATG5 rs12212740 (G/A)

The noncoding SNP rs12212740 of ATG5 is located at intron 3, 7kb downstream and 8kb upstream of Exon 3 and 4 respectively. The SNP has been already been shown to be associated with Asthma [46, 260]. However, the functional consequences of this SNP are still unknown. In our population, we did not observe any significant association of the SNP with HBV infection susceptibility or progression. The mean viral load of patients with the wild type (GG) and mutant genotype (AA) of the patient were compared to find out any significant difference. The average viral load was observed to be higher in the patient with homozygous mutant genotype (AA) in acute stage of infection. Moreover, the patients with AA genotype in the acute stage of infection had higher mean ALT and mean AST levels. However, we found significantly higher mean Bilirubin level in cirrhosis patients with heterozygous genotype (GA).

Therefore, these three SNPs (rs77859116, rs115576116 and rs12212740) could be genotyped in other population groups in order to understand their role in HBV infection and obtain more conclusive results.

5.2.3 Haplotype distribution and linkage disequilibrium among 4 nsSNPs of *ATG5*

ATG5 polymorphisms, as discussed earlier, have studied in variety of diseases. The distribution haplotypes of the promoter polymorphisms of ATG5 –769 T>C rs506027 and –335 G>A rs510432 and an intronic SNP 830 C>T rs573775, have been already studied and shown in Asthmatic patients. The results showed no significant association of these SNPs with Asthma susceptibility. However, GA/AA genotypes in rs510432 were found to be associated with higher sputum neutrophil count in adult asthmatic patients. Moreover, the two promoter SNPs; rs506027 (T>C) and rs510432 (G>A) polymorphisms were also observed to be in complete linkage equilibrium thereby explaining the differential effects of the two haplotypes (TG and CA) on ATG5 promoter activity in both human airway epithelial cells

and human mast cell [187]. The same results were also demonstrated by another study where the sepsis associated polymorphisms haplotype rs506027–rs510432 T-G was shown to enhance the promoter activity [259]. Therefore, haplotype and linkage disequilibrium analysis of polymorphisms within a gene or between two genes becomes essential for understanding the contribution of the linked/unlinked variants in disease pathogenesis. In this context, another study demonstrates the linkage of ATG5 (rs2245214, C>G) and ATG16 (rs2241880, T>C) polymorphisms in PDB patients. The results indicate that an individual carrying C allele of the ATG16L1 polymorphism rs2241880 and G allele of ATG5 rs2245214 polymorphism would be at increased risk of getting PDB [32].

In our population, haplotype and linkage disequilibrium (LD) analysis was carried out for four SNPs of ATG5; rs34793250 [M129V] (T/C), and non-coding SNPs rs2245214 (C/G), rs12212740 (G/A) and rs510432 (G/A) for which we had the complete genotype data for 550 HBV infected cases and 250 controls. Five haplotypes, TCGA, TGGG, TCGG, TGGA, and TCAA, were identified by haplotype analysis, for both the groups (cases + controls). Among these, three haplotypes TCGA, TGGG and TGGA were commonly observed and were found to be associated in HBV infection ($p<0.05$), while, the distribution of TCGG and TCAA haplotypes was not found to be significantly different in the cases and controls ($p>0.05$). The associated haplotypes TCGA, TGGG and TGGA contained combination of wild-type and mutant alleles of the genotyped SNPs and proved to be risk factor of HBV infection. Moreover, the LD plot reveals a block with significant correlation of linkage disequilibrium ($D'=0.99$) between two SNPs (rs34793250-rs12212740). Therefore, the acquisition of these linked alleles (associated haplotypes) of these important SNPs of *ATG5* may increase an individual's susceptibility to HBV infection. Further experiments could be carried out in order to study the effect of identified haplotypes in intervening autophagy process.

Chapter 6

CONCLUSION

CONCLUSION AND FUTURE PROSPECTIVES

The predicted deleterious nsSNPs by *in silico* tools, as well as the other noncoding SNPs of ATG5, were successfully genotyped in HBV patients as well as in healthy controls. We found a significant association of the most destabilizing nsSNP rs34793250 (T/C) [M129V] with HBV infection susceptibility. According to the observed allelic frequency, the C allele may predispose individuals to HBV infection. Moreover, the SNP was found to be significantly associated with different HBV infection stages (asymptomatic, acute and chronic). The TC genotype of this SNP was also found to be significantly related to higher mean AST and ALT levels in patients.SNP rs2245214 (C/G), located in the noncoding region, was also found to be linked with HBV infection susceptibility. Thus, the presence of G allele in the genotype of an individual may increase the risk for HBV acquisition. The third SNP, rs510432 (G/A) located in the promoter region of ATG5, did not show any association with HBV susceptibility in the allelic model. However, a significant association has been observed in the genotypic models (heterozygous and dominant), demonstrating the risk imposed by the mutant allele. Moreover, the patients with mutant AA genotype had significantly higher mean viral load. The SNP rs12212740 (G/A) does not play any role in HBV infection as we did not find any significant association among the different genetic models and stages of infection. However, the patients with genotype GA were found to have higher mean viral load, ALT and Bilirubin levels which again points towards the significance of this SNP in HBV infection.

In summary, the study highlights the important SNPs rs34793250 (T/C), rs2245214 (C/G) and rs510432 (G/A) in ATG5 which may increase individuals' risk of HBV infection. We have already shown the deleterious impact of the nsSNP rs34793250 over ATG5 protein structure and function through an *in silico* approach. The *in silico* methodology, therefore, helped in the selection of an important nsSNP in ATG5 which may increase HBV infection susceptibility in our study population. The other SNP rs2245214 which is significantly associated with HBV infection has been already linked to Paget disease of bone and thyroid carcinoma. Moreover, this polymorphism has been shown to result in loss of recognition site for a SRp40 protein which is involved in mRNA splicing. Last but not the least, the ATG5 promoter SNP rs510432, also plays a substantial role in predisposing individuals towards

HBV infection. Moreover, this SNP constitutes the binding site for two transcription factors STAT1and C-Fos and therefore it may lead to transcriptional aberrations that could lead to disease development. Consequently, these three SNPs of ATG5 have considerable implications and significance in addition, to being involved in HBV infection susceptibility. Therefore, further *in silico* and experimental studies could be carried out to validate the effect of these nsSNPs on the protein stability, PTMs and binding affinities. Moreover, *in silico* and experimental studies could be performed for analyzing the impact of non-coding SNPs on the splicing events and transcription levels/promoter activity. Additionally, the autophagy levels could also be detected in patients having particular genotype to analyze their role in altering the autophagy pathway. Furthermore, replication studies and patient follow up could be done in order to elucidate the role of SNPs in HBV infection susceptibility, progression, outcome or treatment. The overall conclusion with future prospectives of the study has been shown in Figure 6.1. After validation, the SNPs could be utilized as biomarkers for risk analysis or for the prediction of drug response in patients.

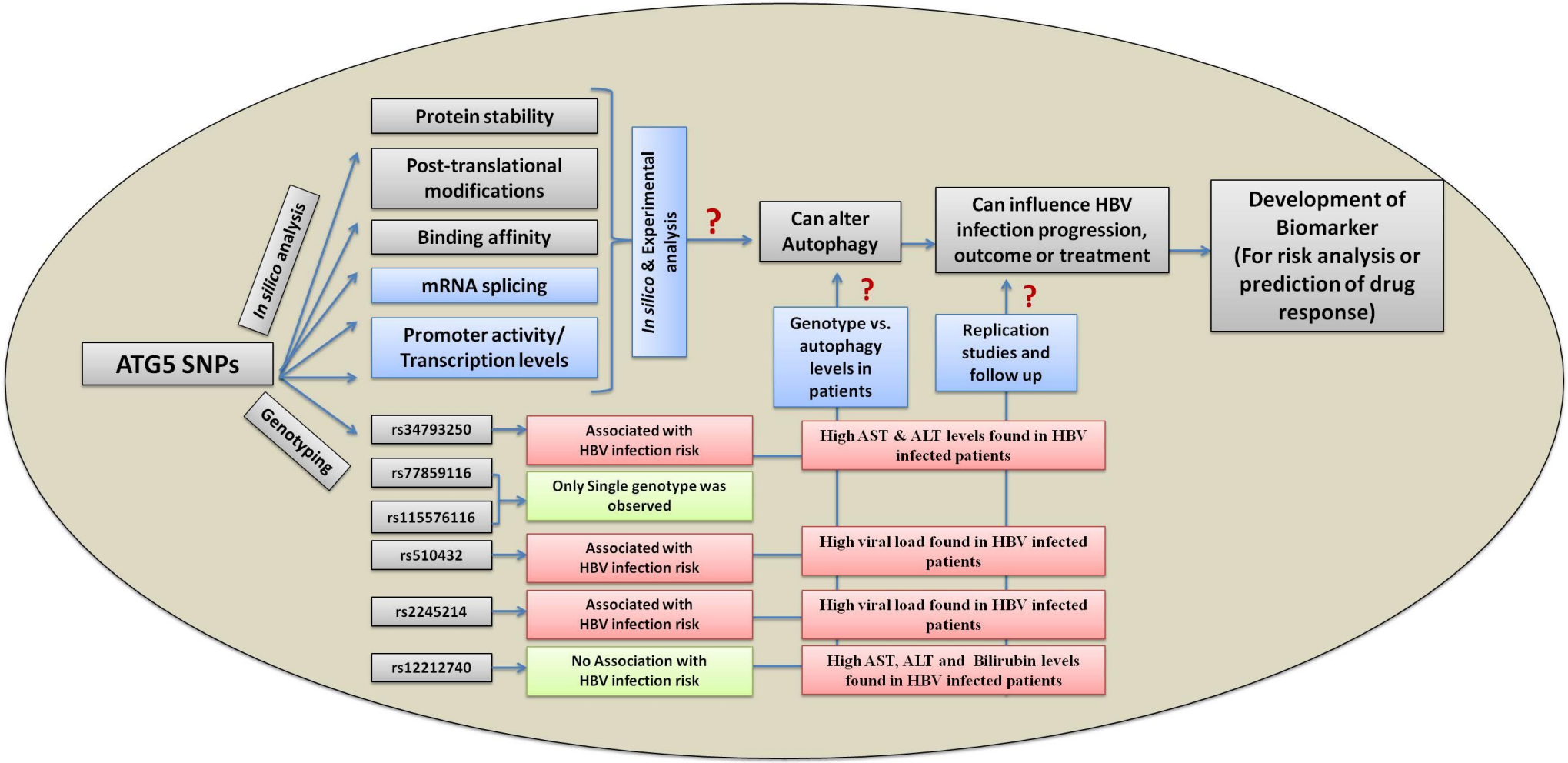

Figure 6.1: The overall conclusion and the future prospectives of the study.

Chapter 7

REFERENCES

References

[1] World Health Organization, "hepatitis B," 2002.

[2] World Health Organization, "hepatitis B," 2017.

[3] P. Puri, "Tackling the Hepatitis B Disease Burden in India," *J Clin Exp Hepatol,* vol. 4, pp. 312-9, Dec 2014.

[4] G. A. Niro, R. Fontana, D. Gioffreda, M. R. Valvano, A. Lacobellis, D. Facciorusso, and A. Andriulli, "Tumor necrosis factor gene polymorphisms and clearance or progression of hepatitis B virus infection," *Liver Int,* vol. 25, pp. 1175-81, Dec 2005.

[5] K. Migita, K. Sawakami-Kobayashi, Y. Maeda, K. Nakao, S. Kondoh, M. Sugiura, R. Kawasumi, O. Segawa, H. Tajima, M. Machida, M. Nakamura, K. Yano, S. Abiru, E. Kawasaki, H. Yatsuhashi, K. Eguchi, and H. Ishibashi, "Interleukin-18 promoter polymorphisms and the disease progression of Hepatitis B virus-related liver disease," *Transl Res,* vol. 153, pp. 91-6, Feb 2009.

[6] W. Li, Y. Jiang, Q. Jin, X. Shi, J. Jin, Y. Gao, Y. Pan, H. Zhang, J. Jiang, and J. Niu, "Expression and gene polymorphisms of interleukin 28B and hepatitis B virus infection in a Chinese Han population," *Liver Int,* vol. 31, pp. 1118-26, Sep 2011.

[7] H. Wu, G. Zhao, F. Qian, K. Liu, J. Xie, H. Zhou, J. Xu, Y. Xu, Y. Han, Q. Xie, and H. Wang, "Association of IL28B polymorphisms with peginterferon treatment response in Chinese Han patients with HBeAg-positive chronic hepatitis B," *Liver Int,* vol. 35, pp. 473-81, Feb 2015.

[8] L. Wang, Z. Q. Zou, and K. Wang, "Clinical Relevance of HLA Gene Variants in HBV Infection," *J Immunol Res,* vol. 2016, p. 9069375, 2016.

[9] D. Ge, J. Fellay, A. J. Thompson, J. S. Simon, K. V. Shianna, T. J. Urban, E. L. Heinzen, P. Qiu, A. H. Bertelsen, A. J. Muir, M. Sulkowski, J. G. McHutchison, and D. B. Goldstein, "Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance," *Nature,* vol. 461, pp. 399-401, Sep 17 2009.

[10] J. A. Holmes, W. Sievert, and A. J. Thompson, "IL28B polymorphism and genetic biomarkers of viral clearance in hepatitis C virus infection," *Biomark Med,* vol. 5, pp. 461-78, Aug 2011.

[11] M. A. Khattab, H. M. Abdelghany, M. M. Ramzy, and R. M. Khairy, "Impact of IL28B gene polymorphisms rs8099917 and rs12980275 on response to pegylated interferon-α/ribavirin therapy in chronic hepatitis C genotype 4 patients," *Journal of biomedical research,* vol. 30, p. 40, 2016.

[12] X. Wu, Z. Xin, X. Zhu, L. Pan, Z. Li, H. Li, and Y. Liu, "Evaluation of susceptibility locus for response to interferon-alpha based therapy in chronic hepatitis B patients in Chinese," *Antiviral Res,* vol. 93, pp. 297-300, Feb 2012.

[13] M. J. Sonneveld, V. W. Wong, A. M. Woltman, G. L. Wong, Y. Cakaloglu, S. Zeuzem, E. H. Buster, A. G. Uitterlinden, B. E. Hansen, H. L. Chan, and H. L. Janssen, "Polymorphisms near IL28B and serologic response to peginterferon in HBeAg-positive patients with chronic hepatitis B," *Gastroenterology,* vol. 142, pp. 513-520 e1, Mar 2012.

[14] T. J. Liang, "Hepatitis B: the virus and disease," *Hepatology,* vol. 49, pp. S13-21, May 2009.

[15] C. W. Shepard, E. P. Simard, L. Finelli, A. E. Fiore, and B. P. Bell, "Hepatitis B virus infection: epidemiology and vaccination," *Epidemiol Rev,* vol. 28, pp. 112-25, 2006.

[16] W. Martinet, P. Agostinis, B. Vanhoecke, M. Dewaele, and G. R. De Meyer, "Autophagy in disease: a double-edged sword with therapeutic potential," *Clin Sci (Lond),* vol. 116, pp. 697-712, May 2009.

[17] S. B. Kudchodkar and B. Levine, "Viruses and autophagy," *Rev Med Virol,* vol. 19, pp. 359-78, Nov 2009.

[18] H. J. Kim, S. Lee, and J. U. Jung, "When autophagy meets viruses: a double-edged sword with functions in defense and offense," *Semin Immunopathol,* vol. 32, pp. 323-41, Dec 2010.

[19] D. Sir, Y. Tian, W. L. Chen, D. K. Ann, T. S. Yen, and J. H. Ou, "The early autophagic pathway is activated by hepatitis B virus and required for viral DNA replication," *Proc Natl Acad Sci U S A,* vol. 107, pp. 4383-8, Mar 02 2010.

[20] J. Li, Y. Liu, Z. Wang, K. Liu, Y. Wang, J. Liu, H. Ding, and Z. Yuan, "Subversion of cellular autophagy machinery by hepatitis B virus for viral envelopment," *J Virol,* vol. 85, pp. 6319-33, Jul 2011.

[21] Y. Tian, D. Sir, C. F. Kuo, D. K. Ann, and J. H. Ou, "Autophagy required for hepatitis B virus replication in transgenic mice," *J Virol,* vol. 85, pp. 13453-6, Dec 2011.

[22] D. Sir and J. H. Ou, "Autophagy in viral replication and pathogenesis," *Mol Cells,* vol. 29, pp. 1-7, Jan 2010.

[23] Y. I. Li, E. F. Libby, M. J. Lewis, J. Liu, J. J. Shacka, and D. R. Hurst, "Increased autophagic response in a population of metastatic breast cancer cells," *Oncol Lett,* vol. 12, pp. 523-529, Jul 2016.

[24] R. A. Nixon and D. S. Yang, "Autophagy and neuronal cell death in neurological disorders," *Cold Spring Harb Perspect Biol,* vol. 4, 2012.

[25] A. B. Gustafsson and R. A. Gottlieb, "Autophagy in ischemic heart disease," *Circ Res,* vol. 104, pp. 150-8, Jan 30 2009.

[26] M. Taneike, O. Yamaguchi, A. Nakai, S. Hikoso, T. Takeda, I. Mizote, T. Oka, T. Tamai, J. Oyabu, T. Murakawa, K. Nishida, T. Shimizu, M. Hori, I. Komuro, T. S. Takuji Shirasawa, N. Mizushima, and K. Otsu, "Inhibition of autophagy in the heart induces age-related cardiomyopathy," *Autophagy,* vol. 6, pp. 600-6, Jul 2010.

[27] I. Gukovsky, S. J. Pandol, O. A. Mareninova, N. Shalbueva, W. Jia, and A. S. Gukovskaya, "Impaired autophagy and organellar dysfunction in pancreatitis," *J Gastroenterol Hepatol,* vol. 27 Suppl 2, pp. 27-32, Mar 2012.

[28] A. Orvedahl and B. Levine, "Eating the enemy within: autophagy in infectious diseases," *Cell Death Differ,* vol. 16, pp. 57-69, Jan 2009.

[29] Y. C. Lin, P. F. Chang, H. F. Lin, K. Liu, M. H. Chang, and Y. H. Ni, "Variants in the autophagy-related gene IRGM confer susceptibility to non-alcoholic fatty liver disease by modulating lipophagy," *J Hepatol,* vol. 65, pp. 1209-1216, Dec 2016.

[30] L. Li, X. Chen, and H. Gu, "The signaling involving in autophagy machinery in keratinocytes and therapeutic approaches for skin diseases," *Oncotarget,* May 12 2016.

[31] H. Zhuang, K. Ali, S. Ardu, C. Tredwin, and B. Hu, "Autophagy in dental tissues: a double-edged sword," *Cell Death Dis,* vol. 7, p. e2192, 2016.

[32] R. Usategui-Martin, J. Garcia-Aparicio, L. Corral-Gudino, I. Calero-Paniagua, J. Del Pino-Montes, and R. Gonzalez Sarmiento, "Polymorphisms in autophagy genes are associated with paget disease of bone," *PLoS one,* vol. 10, p. e0128984, 2015.

[33] M. Hirano, Y. Nakamura, K. Saigoh, H. Sakamoto, S. Ueno, C. Isono, K. Miyamoto, M. Akamatsu, Y. Mitsui, and S. Kusunoki, "Mutations in the gene encoding p62 in Japanese patients with amyotrophic lateral sclerosis," *Neurology,* vol. 80, pp. 458-63, Jan 29 2013.

[34] D. Oz-Levi, B. Ben-Zeev, E. K. Ruzzo, Y. Hitomi, A. Gelman, K. Pelak, Y. Anikster, H. Reznik-Wolf, I. Bar-Joseph, T. Olender, A. Alkelai, M. Weiss, E. Ben-Asher, D. Ge, K. V. Shianna, Z. Elazar, D. B. Goldstein, E. Pras, and D. Lancet, "Mutation in TECPR2 reveals a role for autophagy in hereditary spastic paraparesis," *Am J Hum Genet,* vol. 91, pp. 1065-72, Dec 7 2012.

[35] N. Mizushima, "Autophagy: process and function," *Genes Dev,* vol. 21, pp. 2861-73, Nov 15 2007.

[36] I. Tanida, "Autophagosome formation and molecular mechanism of autophagy," *Antioxid Redox Signal,* vol. 14, pp. 2201-14, Jun 2011.

[37] W. Li, M. Chen, E. Wang, L. Hu, M. J. Hawkesford, L. Zhong, Z. Chen, Z. Xu, L. Li, Y. Zhou, C. Guo, and Y. Ma, "Genome-wide analysis of autophagy-associated genes in foxtail millet

(Setaria italica L.) and characterization of the function of SiATG8a in conferring tolerance to nitrogen starvation in rice," *BMC Genomics,* vol. 17, p. 797, Oct 12 2016.

[38] Z. Yang and D. J. Klionsky, "Mammalian autophagy: core molecular machinery and signaling regulation," *Curr Opin Cell Biol,* vol. 22, pp. 124-31, Apr 2010.

[39] J. Romanov, M. Walczak, I. Ibiricu, S. Schuchner, E. Ogris, C. Kraft, and S. Martens, "Mechanism and functions of membrane binding by the Atg5-Atg12/Atg16 complex during autophagosome formation," *EMBO J,* vol. 31, pp. 4304-17, Nov 14 2012.

[40] R. Mukherjee, G. V. Shravanthi, P. B. Reddy, P. Rao, and D. Reddy, "Escalated Expression of Autophagy Related Gene Atg5 in Hepatitis B Virus Infection," *J. Microbiol. Res. Rev.,* vol. 2, pp. 68-73, 2014.

[41] J. D. Rioux, R. J. Xavier, K. D. Taylor, M. S. Silverberg, P. Goyette, A. Huett, T. Green, P. Kuballa, M. M. Barmada, L. W. Datta, Y. Y. Shugart, A. M. Griffiths, S. R. Targan, A. F. Ippoliti, E. J. Bernard, L. Mei, D. L. Nicolae, M. Regueiro, L. P. Schumm, A. H. Steinhart, J. I. Rotter, R. H. Duerr, J. H. Cho, M. J. Daly, and S. R. Brant, "Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis," *Nat Genet,* vol. 39, pp. 596-604, May 2007.

[42] C. Capela, A. D. Dossou, R. Silva-Gomes, G. E. Sopoh, M. Makoutode, J. F. Menino, A. G. Fraga, C. Cunha, A. Carvalho, F. Rodrigues, and J. Pedrosa, "Genetic Variation in Autophagy-Related Genes Influences the Risk and Phenotype of Buruli Ulcer," *PLoS Negl Trop Dis,* vol. 10, p. e0004671, Apr 2016.

[43] K. Douroudis, K. Kingo, T. Traks, E. Reimann, K. Raud, R. Ratsep, R. Mossner, H. Silm, E. Vasar, and S. Koks, "Polymorphisms in the ATG16L1 gene are associated with psoriasis vulgaris," *Acta Derm Venereol,* vol. 92, pp. 85-7, Jan 2012.

[44] M. R. Kang, M. S. Kim, J. E. Oh, Y. R. Kim, S. Y. Song, S. S. Kim, C. H. Ahn, N. J. Yoo, and S. H. Lee, "Frameshift mutations of autophagy-related genes ATG2B, ATG5, ATG9B and ATG12 in gastric and colorectal cancers with microsatellite instability," *J Pathol,* vol. 217, pp. 702-6, Apr 2009.

[45] L. J. Martin, J. Gupta, S. S. Jyothula, M. B. Kovacic, J. M. B. Myers, T. L. Patterson, M. B. Ericksen, H. He, A. M. Gibson, and T. M. Baye, "Functional variant in the autophagy-related 5 gene promotor is associated with childhood asthma," *PLoS one,* vol. 7, p. e33454, 2012.

[46] A. H. Poon, F. Chouiali, S. M. Tse, A. A. Litonjua, S. N. Hussain, C. J. Baglole, D. H. Eidelman, R. Olivenstein, J. G. Martin, S. T. Weiss, Q. Hamid, and C. Laprise, "Genetic and histologic evidence for autophagy in asthma pathogenesis," *J Allergy Clin Immunol,* vol. 129, pp. 569-71, Feb 2012.

[47] P. Lopez, E. Alonso-Perez, J. Rodriguez-Carrio, and A. Suarez, "Influence of Atg5 mutation in SLE depends on functional IL-10 genotype," *PLoS one,* vol. 8, p. e78756, 2013.

[48] T. S. Plantinga, E. van de Vosse, A. Huijbers, M. G. Netea, L. A. Joosten, J. W. Smit, and R. T. Netea-Maier, "Role of genetic variants of autophagy genes in susceptibility for non-medullary thyroid cancer and patients outcome," *PLoS one,* vol. 9, p. e94086, 2014.

[49] M. Kim, E. Sandford, D. Gatica, Y. Qiu, X. Liu, Y. Zheng, B. A. Schulman, J. Xu, I. Semple, S. H. Ro, B. Kim, R. N. Mavioglu, A. Tolun, A. Jipa, S. Takats, M. Karpati, J. Z. Li, Z. Yapici, G. Juhasz, J. H. Lee, D. J. Klionsky, and M. Burmeister, "Mutation in ATG5 reduces autophagy and leads to ataxia with developmental delay," *Elife,* vol. 5, Jan 26 2016.

[50] A. J. Frodsham, "Host genetics and the outcome of hepatitis B viral infection," *Transpl Immunol,* vol. 14, pp. 183-6, Aug 2005.

[51] V. Chatzidaki, E. Kouroumalis, and E. Galanakis, "Hepatitis B virus acquisition and pathogenesis in childhood: host genetic determinants," *J Pediatr Gastroenterol Nutr,* vol. 52, pp. 3-8, Jan 2011.

[52] S. Tuncbilek, "Relationship between cytokine gene polymorphisms and chronic hepatitis B virus infection," *World J Gastroenterol,* vol. 20, pp. 6226-35, May 28 2014.

[53] S. Teng, E. Michonova-Alexova, and E. Alexov, "Approaches and resources for prediction of the effects of non-synonymous single nucleotide polymorphism on protein function and interactions," *Curr Pharm Biotechnol,* vol. 9, pp. 123-33, Apr 2008.

[54] K. Ramavartheni and S. Rao, "Identification of structurally and functionally significant deleterious nsSNPs of GSS gene: in silico analysis," *Advances in Bioscience and Biotechnology,* vol. 2010, 2010.

[55] S. Teng, A. K. Srivastava, C. E. Schwartz, E. Alexov, and L. Wang, "Structural assessment of the effects of amino acid substitutions on protein stability and protein protein interaction," *Int J Comput Biol Drug Des,* vol. 3, pp. 334-49, 2010.

[56] J. H. MacLachlan and B. C. Cowie, "Hepatitis B virus epidemiology," *Cold Spring Harb Perspect Med,* vol. 5, p. a021410, May 01 2015.

[57] J. Hou, Z. Liu, and F. Gu, "Epidemiology and Prevention of Hepatitis B Virus Infection," *Int J Med Sci,* vol. 2, pp. 50-57, 2005.

[58] S. Datta, S. Chatterjee, V. Veer, and R. Chakravarty, "Molecular biology of the hepatitis B virus for clinicians," *J Clin Exp Hepatol,* vol. 2, pp. 353-65, Dec 2012.

[59] P. Sandhu, M. Haque, T. Humphries-Bickley, S. Ravi, and J. Song, "Hepatitis B Virus Immunopathology, Model Systems, and Current Therapies," *Front Immunol,* vol. 8, p. 436, 2017.

[60] M. Blanchet and C. Sureau, "Infectivity determinants of the hepatitis B virus pre-S domain are confined to the N-terminal 75 amino acid residues," *J Virol,* vol. 81, pp. 5841-9, Jun 2007.

[61] S. Ezzikouri, M. Ozawa, M. Kohara, N. Elmdaghri, S. Benjelloun, and K. Tsukiyama-Kohara, "Recent insights into hepatitis B virus-host interactions," *J Med Virol,* vol. 86, pp. 925-32, Jun 2014.

[62] R. J. Lamontagne, S. Bagga, and M. J. Bouchard, "Hepatitis B virus molecular biology and pathogenesis," *Hepatoma Res,* vol. 2, pp. 163-186, 2016.

[63] N. P. Nelson, D. J. Jamieson, and T. V. Murphy, "Prevention of Perinatal Hepatitis B Virus Transmission," *J Pediatric Infect Dis Soc,* vol. 3 Suppl 1, pp. S7-S12, Sep 2014.

[64] N. Inaba, R. Ohkawa, A. Matsuura, J. Kudoh, and H. Takamizawa, "Sexual transmission of hepatitis B surface antigen. Infection of husbands by HBsAg carrier-state wives," *Br J Vener Dis,* vol. 55, pp. 366-8, Oct 1979.

[65] E. Franco, B. Bagnato, M. G. Marino, C. Meleleo, L. Serino, and L. Zaratti, "Hepatitis B: Epidemiology and prevention in developing countries," *World J Hepatol,* vol. 4, pp. 74-80, Mar 27 2012.

[66] A. Bertoletti and A. J. Gehring, "The immune response during hepatitis B virus infection," *J Gen Virol,* vol. 87, pp. 1439-49, Jun 2006.

[67] M. R. Thursz, D. Kwiatkowski, C. E. Allsopp, B. M. Greenwood, H. C. Thomas, and A. V. Hill, "Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia," *N Engl J Med,* vol. 332, pp. 1065-9, Apr 20 1995.

[68] C. L. Thio, D. L. Thomas, P. Karacki, X. Gao, D. Marti, R. A. Kaslow, J. J. Goedert, M. Hilgartner, S. A. Strathdee, P. Duggal, S. J. O'Brien, J. Astemborski, and M. Carrington, "Comprehensive analysis of class I and class II HLA antigens and chronic hepatitis B virus infection," *J Virol,* vol. 77, pp. 12083-7, Nov 2003.

[69] K. Matsuura, M. Isogawa, and Y. Tanaka, "Host genetic variants influencing the clinical course of hepatitis B virus infection," *J Med Virol,* vol. 88, pp. 371-9, Mar 2016.

[70] T. Hohler, A. Kruger, G. Gerken, P. M. Schneider, K. H. Meyer zum Buschenefelde, and C. Rittner, "A tumor necrosis factor-alpha (TNF-alpha) promoter polymorphism is associated with chronic hepatitis B infection," *Clin Exp Immunol,* vol. 111, pp. 579-82, Mar 1998.

[71] M. H. Zheng, D. D. Xiao, X. F. Lin, S. J. Wu, M. M. Peng, X. Y. Yu, W. Y. Liu, L. F. Li, K. Q. Shi, Y. C. Fan, and Y. P. Chen, "The tumour necrosis factor-alpha-238A allele increases the risk of chronic HBV infection in European populations," *J Viral Hepat,* vol. 19, pp. e11-7, Feb 2012.

[72] S. A. Gonzalez and E. B. Keeffe, "IL-28B As a Predictor of Sustained Virologic Response in Patients with Chronic Hepatitis C Virus Infection," *Gastroenterol Hepatol (N Y),* vol. 7, pp. 366-73, Jun 2011.

[73] A. F. Stattermayer and P. Ferenci, "Effect of IL28B genotype on hepatitis B and C virus infection," *Curr Opin Virol,* vol. 14, pp. 50-5, Oct 2015.

[74] F. Kurbanov, M. Abdel-Hamid, R. Latanich, J. Astemborski, M. Mohamed, N. M. Mikhail, M. El-Daly, S. El-Kafrawy, D. L. Thomas, and C. L. Thio, "Genetic polymorphism in IL28B is associated with spontaneous clearance of hepatitis C virus genotype 4 infection in an Egyptian cohort," *J Infect Dis,* vol. 204, pp. 1391-4, Nov 2011.

[75] W. K. Seto, D. K. Wong, M. Kopaniszen, P. Proitsi, P. C. Sham, I. F. Hung, J. Fung, C. L. Lai, and M. F. Yuen, "HLA-DP and IL28B polymorphisms: influence of host genome on hepatitis B surface antigen seroclearance in chronic hepatitis B," *Clin Infect Dis,* vol. 56, pp. 1695-703, Jun 2013.

[76] A. A. Al-Qahtani, M. R. Al-Anazi, A. A. Abdo, F. M. Sanai, W. K. Al-Hamoudi, K. A. Alswat, H. I. Al-Ashgar, N. Z. Khalaf, N. A. Viswan, and M. N. Al Ahdal, "Genetic variation in interleukin 28B and correlation with chronic hepatitis B virus infection in Saudi Arabian patients," *Liver Int,* vol. 34, pp. e208-16, Aug 2014.

[77] J. A. Holmes, T. Nguyen, D. Ratnam, N. M. Heerasing, J. V. Tehan, S. Bonanzinga, A. Dev, S. Bell, S. Pianko, R. Chen, K. Visvanathan, R. Hammond, D. Iser, F. Rusli, W. Sievert, P. V. Desmond, D. S. Bowden, and A. J. Thompson, "IL28B genotype is not useful for predicting treatment outcome in Asian chronic hepatitis B patients treated with pegylated interferon-alpha," *J Gastroenterol Hepatol,* vol. 28, pp. 861-6, May 2013.

[78] M. P. Martin, Y. Qi, J. J. Goedert, S. K. Hussain, G. D. Kirk, W. K. Hoots, S. Buchbinder, M. Carrington, and C. L. Thio, "IL28B polymorphism does not determine outcomes of hepatitis B virus or HIV infection," *J Infect Dis,* vol. 202, pp. 1749-53, Dec 01 2010.

[79] O. Kandemir, S. B. Fidanci, N. Demir, A. Gorur, and L. Tamer, "Chronic hepatitis B and IL28B rs12979860 polymorphism: preliminary study," *Mol Biol Rep,* vol. 40, pp. 6189-94, Nov 2013.

[80] A. L. Truelove, T. K. Oleksyk, S. Shrestha, C. L. Thio, J. J. Goedert, S. M. Donfield, G. D. Kirk, D. L. Thomas, S. J. O'Brien, and M. W. Smith, "Evaluation of IL10, IL19 and IL20 gene polymorphisms and chronic hepatitis B infection outcome," *Int J Immunogenet,* vol. 35, pp. 255-64, Jun 2008.

[81] Q. R. Zhu, Y. L. Ge, S. Q. Gu, H. Yu, J. S. Wang, X. H. Gu, L. E. Fei, and Z. Q. Dong, "Relationship between cytokines gene polymorphism and susceptibility to hepatitis B virus intrauterine infection," *Chin Med J (Engl),* vol. 118, pp. 1604-9, Oct 05 2005.

[82] J. Zhou, D. K. Smith, L. Lu, V. K. Poon, F. Ng, D. Q. Chen, J. D. Huang, K. Y. Yuen, K. Y. Cao, and B. J. Zheng, "A non-synonymous single nucleotide polymorphism in IFNAR1 affects susceptibility to chronic hepatitis B virus infection," *J Viral Hepat,* vol. 16, pp. 45-52, Jan 2009.

[83] W. Abbott, E. Gane, I. Winship, S. Munn, and C. Tukuitonga, "Polymorphism in intron 1 of the interferon-gamma gene influences both serum immunoglobulin E levels and the risk for chronic hepatitis B virus infection in Polynesians," *Immunogenetics,* vol. 59, pp. 187-95, Mar 2007.

[84] H. Yu, Q. R. Zhu, S. Q. Gu, and L. E. Fei, "Relationship between IFN-gamma gene polymorphism and susceptibility to intrauterine HBV infection," *World J Gastroenterol,* vol. 12, pp. 2928-31, May 14 2006.

[85] S. H. Ahn, D. Y. Kim, H. Y. Chang, S. P. Hong, J. S. Shin, Y. S. Kim, H. Kim, J. K. Kim, Y. H. Paik, K. S. Lee, C. Y. Chon, Y. M. Moon, and K. H. Han, "Association of genetic variations in CCR5 and its ligand, RANTES with clearance of hepatitis B virus in Korea," *J Med Virol,* vol. 78, pp. 1564-71, Dec 2006.

[86] K. S. Brown, S. D. Ryder, W. L. Irving, R. B. Sim, and T. P. Hickling, "Mannan binding lectin and viral hepatitis," *Immunol Lett,* vol. 108, pp. 34-44, Jan 15 2007.
[87] T. J. Park, J. Y. Chun, J. S. Bae, B. L. Park, H. S. Cheong, H. S. Lee, Y. J. Kim, and H. D. Shin, "CCND2 polymorphisms associated with clearance of HBV infection," *J Hum Genet,* vol. 55, pp. 416-20, Jul 2010.
[88] Y. W. Huang, Y. T. Liao, W. Chen, C. L. Chen, J. T. Hu, C. J. Liu, M. Y. Lai, P. J. Chen, D. S. Chen, S. S. Yang, and J. H. Kao, "Vitamin D receptor gene polymorphisms and distinct clinical phenotypes of hepatitis B carriers in Taiwan," *Genes Immun,* vol. 11, pp. 87-93, Jan 2010.
[89] G. Deng, G. Zhou, Y. Zhai, S. Li, X. Li, Y. Li, R. Zhang, Z. Yao, Y. Shen, B. Qiang, Y. Wang, and F. He, "Association of estrogen receptor alpha polymorphisms with susceptibility to chronic hepatitis B virus infection," *Hepatology,* vol. 40, pp. 318-26, Aug 2004.
[90] J. P. Bronowicki, I. Abdelmouttaleb, L. Peyrin-Biroulet, V. Venard, H. Khiri, N. Chabi, E. K. Amouzou, H. Barraud, P. Halfon, A. Sanni, M. A. Bigard, A. Le Faou, and J. L. Gueant, "Methylenetetrahydrofolate reductase 677 T allele protects against persistent HBV infection in West Africa," *J Hepatol,* vol. 48, pp. 532-9, Apr 2008.
[91] A. Al-Qahtani, M. Al-Ahdal, A. Abdo, F. Sanai, M. Al-Anazi, N. Khalaf, N. A. Viswan, H. Al-Ashgar, H. Al-Humaidan, R. Al-Suwayeh, Z. Hussain, S. Alarifi, M. Al-Okail, and F. N. Almajhdi, "Toll-like receptor 3 polymorphism and its association with hepatitis B virus infection in Saudi Arabian patients," *J Med Virol,* vol. 84, pp. 1353-9, Sep 2012.
[92] J. F. Wu, C. H. Chen, Y. H. Ni, Y. T. Lin, H. L. Chen, H. Y. Hsu, and M. H. Chang, "Toll-like receptor and hepatitis B virus clearance in chronic infected patients: a long-term prospective cohort study in Taiwan," *J Infect Dis,* vol. 206, pp. 662-8, Sep 01 2012.
[93] Z. Hu, Y. Liu, X. Zhai, J. Dai, G. Jin, L. Wang, L. Zhu, Y. Yang, J. Liu, M. Chu, J. Wen, K. Xie, G. Du, Q. Wang, Y. Zhou, M. Cao, L. Liu, Y. He, Y. Wang, G. Zhou, W. Jia, J. Lu, S. Li, H. Yang, Y. Shi, W. Zhou, and H. Shen, "New loci associated with chronic hepatitis B virus infection in Han Chinese," *Nat Genet,* vol. 45, pp. 1499-503, Dec 2013.
[94] E. Wirawan, T. Vanden Berghe, S. Lippens, P. Agostinis, and P. Vandenabeele, "Autophagy: for better or for worse," *Cell Res,* vol. 22, pp. 43-61, Jan 2012.
[95] B. Levine and J. Yuan, "Autophagy in cell death: an innocent convict?," *J Clin Invest,* vol. 115, pp. 2679-88, Oct 2005.
[96] J. D. Mancias and A. C. Kimmelman, "Mechanisms of Selective Autophagy in Normal Physiology and Cancer," *J Mol Biol,* vol. 428, pp. 1659-80, May 08 2016.
[97] S. Ghavami, S. Shojaei, B. Yeganeh, S. R. Ande, J. R. Jangamreddy, M. Mehrpour, J. Christoffersson, W. Chaabane, A. R. Moghadam, H. H. Kashani, M. Hashemi, A. A. Owji, and M. J. Los, "Autophagy and apoptosis dysfunction in neurodegenerative disorders," *Prog Neurobiol,* vol. 112, pp. 24-49, Jan 2014.
[98] Y. Ding and M. E. Choi, "Autophagy in diabetic nephropathy," *J Endocrinol,* vol. 224, pp. R15-30, Jan 2015.
[99] A. I. Chiramel, N. R. Brady, and R. Bartenschlager, "Divergent roles of autophagy in virus infection," *Cells,* vol. 2, pp. 83-104, Jan 25 2013.
[100] K. Sharma, R. Goehe, J. M. Beckta, K. Valerie, and D. A. Gewirtz, "Autophagy and radiosensitization in cancer," *EXCLI J,* vol. 13, pp. 178-91, 2014.
[101] S. Y. Tam, V. W. Wu, and H. K. Law, "Influence of autophagy on the efficacy of radiotherapy," *Radiat Oncol,* vol. 12, p. 57, Mar 21 2017.
[102] C. He and D. J. Klionsky, "Regulation mechanisms and signaling pathways of autophagy," *Annu Rev Genet,* vol. 43, pp. 67-93, 2009.
[103] C. W. Wang and D. J. Klionsky, "The molecular mechanism of autophagy," *Mol Med,* vol. 9, pp. 65-76, Mar-Apr 2003.
[104] D. Sir, D. K. Ann, and J. H. Ou, "Autophagy by hepatitis B virus and for hepatitis B virus," *Autophagy,* vol. 6, pp. 548-9, May 2010.

[105] J. Wang, Y. Shi, and H. Yang, "[Infection with hepatitis B virus enhances basal autophagy]," *Wei Sheng Wu Xue Bao,* vol. 50, pp. 1651-6, Dec 2010.
[106] H. T. Zhang, G. G. Chen, B. G. Hu, Z. Y. Zhang, J. P. Yun, M. L. He, and P. B. Lai, "Hepatitis B virus x protein induces autophagy via activating death-associated protein kinase," *J Viral Hepat,* vol. 21, pp. 642-9, 2014.
[107] M. Hoyer-Hansen and M. Jaattela, "Connecting endoplasmic reticulum stress to autophagy by unfolded protein response and calcium," *Cell Death Differ,* vol. 14, pp. 1576-82, Sep 2007.
[108] J. L. Schneider and A. M. Cuervo, "Liver autophagy: much more than just taking out the trash," *Nat Rev Gastroenterol Hepatol,* vol. 11, pp. 187-200, Mar 2014.
[109] A. Kuma, N. Mizushima, N. Ishihara, and Y. Ohsumi, "Formation of the approximately 350-kDa Apg12-Apg5.Apg16 multimeric complex, mediated by Apg16 oligomerization, is essential for autophagy in yeast," *J Biol Chem,* vol. 277, pp. 18619-25, May 24 2002.
[110] T. Hanada, N. N. Noda, Y. Satomi, Y. Ichimura, Y. Fujioka, T. Takao, F. Inagaki, and Y. Ohsumi, "The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy," *J Biol Chem,* vol. 282, pp. 37298-302, Dec 28 2007.
[111] M. Walczak and S. Martens, "Dissecting the role of the Atg12-Atg5-Atg16 complex during autophagosome formation," *Autophagy,* vol. 9, pp. 424-5, Mar 2013.
[112] A. Kuma, M. Hatano, M. Matsui, A. Yamamoto, H. Nakaya, T. Yoshimori, Y. Ohsumi, T. Tokuhisa, and N. Mizushima, "The role of autophagy during the early neonatal starvation period," *Nature,* vol. 432, pp. 1032-6, Dec 23 2004.
[113] A. Takamura, M. Komatsu, T. Hara, A. Sakamoto, C. Kishi, S. Waguri, Y. Eishi, O. Hino, K. Tanaka, and N. Mizushima, "Autophagy-deficient mice develop multiple liver tumors," *Genes Dev,* vol. 25, pp. 795-800, Apr 15 2011.
[114] R. A. Williams, T. K. Smith, B. Cull, J. C. Mottram, and G. H. Coombs, "ATG5 is essential for ATG8-dependent autophagy and mitochondrial homeostasis in Leishmania major," *PLoS Pathog,* vol. 8, p. e1002695, 2012.
[115] C. Lerena, S. D. Calligaris, and M. I. Colombo, "Autophagy: for better or for worse, in good times or in bad times," *Curr Mol Med,* vol. 8, pp. 92-101, Mar 2008.
[116] V. Menendez-Benito and J. Neefjes, "Autophagy in MHC class II presentation: sampling from within," *Immunity,* vol. 26, pp. 1-3, Jan 2007.
[117] H. K. Lee, L. M. Mattei, B. E. Steinberg, P. Alberts, Y. H. Lee, A. Chervonsky, N. Mizushima, S. Grinstein, and A. Iwasaki, "In vivo requirement for Atg5 in antigen presentation by dendritic cells," *Immunity,* vol. 32, pp. 227-39, Feb 26 2010.
[118] C. Munz, "Autophagy proteins in antigen processing for presentation on MHC molecules," *Immunol Rev,* vol. 272, pp. 17-27, Jul 2016.
[119] B. C. Miller, Z. Zhao, L. M. Stephenson, K. Cadwell, H. H. Pua, H. K. Lee, N. N. Mizushima, A. Iwasaki, Y. W. He, W. Swat, and H. W. t. Virgin, "The autophagy gene ATG5 plays an essential role in B lymphocyte development," *Autophagy,* vol. 4, pp. 309-14, Apr 2008.
[120] L. M. Stephenson, B. C. Miller, A. Ng, J. Eisenberg, Z. Zhao, K. Cadwell, D. B. Graham, N. N. Mizushima, R. Xavier, H. W. Virgin, and W. Swat, "Identification of Atg5-dependent transcriptional changes and increases in mitochondrial mass in Atg5-deficient T lymphocytes," *Autophagy,* vol. 5, pp. 625-35, Jul 2009.
[121] H. H. Pua, J. Guo, M. Komatsu, and Y. W. He, "Autophagy is essential for mitochondrial clearance in mature T lymphocytes," *J Immunol,* vol. 182, pp. 4046-55, Apr 1 2009.
[122] R. Baerga, Y. Zhang, P. H. Chen, S. Goldman, and S. Jin, "Targeted deletion of autophagy-related 5 (atg5) impairs adipogenesis in a cellular model and in mice," *Autophagy,* vol. 5, pp. 1118-30, Nov 2009.
[123] S. Tsukamoto, A. Kuma, M. Murakami, C. Kishi, A. Yamamoto, and N. Mizushima, "Autophagy is essential for preimplantation development of mouse embryos," *Science,* vol. 321, pp. 117-20, Jul 4 2008.

[124] Z. Hu, J. Zhang, and Q. Zhang, "Expression pattern and functions of autophagy-related gene atg5 in zebrafish organogenesis," *Autophagy,* vol. 7, pp. 1514-27, Dec 2011.
[125] Z. Tang, M. G. Lin, T. R. Stowe, S. Chen, M. Zhu, T. Stearns, B. Franco, and Q. Zhong, "Autophagy promotes primary ciliogenesis by removing OFD1 from centriolar satellites," *Nature,* vol. 502, pp. 254-7, Oct 10 2013.
[126] P. Vazquez, A. I. Arroba, F. Cecconi, E. J. de la Rosa, P. Boya, and F. de Pablo, "Atg5 and Ambra1 differentially modulate neurogenesis in neural stem cells," *Autophagy,* vol. 8, pp. 187-99, Feb 1 2012.
[127] T. Hara, K. Nakamura, M. Matsui, A. Yamamoto, Y. Nakahara, R. Suzuki-Migishima, M. Yokoyama, K. Mishima, I. Saito, H. Okano, and N. Mizushima, "Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice," *Nature,* vol. 441, pp. 885-9, Jun 15 2006.
[128] J. Nishiyama, E. Miura, N. Mizushima, M. Watanabe, and M. Yuzaki, "Aberrant membranes and double-membrane structures accumulate in the axons of Atg5-null Purkinje cells before neuronal death," *Autophagy,* vol. 3, pp. 591-6, Nov-Dec 2007.
[129] S. Wang, B. Li, H. Qiao, X. Lv, Q. Liang, Z. Shi, W. Xia, F. Ji, and J. Jiao, "Autophagy-related gene Atg5 is essential for astrocyte differentiation in the developing mouse cortex," *EMBO Rep,* vol. 15, pp. 1053-61, Oct 2014.
[130] M. Chae, G. S. Rhee, I. S. Jang, K. Kim, J. H. Lee, S. Y. Lee, M. Kim, J. Yang, J. Park, and S. H. Lee, "ATG5 expression induced by MDMA (ecstasy), interferes with neuronal differentiation of neuroblastoma cells," *Mol Cells,* vol. 27, pp. 571-5, May 31 2009.
[131] J. O. Pyo, S. M. Yoo, H. H. Ahn, J. Nah, S. H. Hong, T. I. Kam, S. Jung, and Y. K. Jung, "Overexpression of Atg5 in mice activates autophagy and extends lifespan," *Nat Commun,* vol. 4, p. 2300, 2013.
[132] G. V. S. R.M. Mukherjee, P. Balkumar Reddy, P.N. Rao and D.N. Reddy, "Escalated Expression of Autophagy Related Gene Atg5 in Hepatitis B Virus Infection," *Journal of Microbiology Research and Reviews,* vol. 2, pp. 68-73, October, 2014 2014.
[133] D. H. Cho, Y. K. Jo, S. C. Kim, I. J. Park, and J. C. Kim, "Down-regulated expression of ATG5 in colorectal cancer," *Anticancer Res,* vol. 32, pp. 4091-6, Sep 2012.
[134] L. Wang, L. Yao, Y. Z. Zheng, Q. Xu, X. P. Liu, X. Hu, P. Wang, and Z. M. Shao, "Expression of autophagy-related proteins ATG5 and FIP200 predicts favorable disease-free survival in patients with breast cancer," *Biochem Biophys Res Commun,* vol. 458, pp. 816-22, Mar 20 2015.
[135] M. G. Hollomon, N. Gordon, J. M. Santiago-O'Farrill, and E. S. Kleinerman, "Knockdown of autophagy-related protein 5, ATG5, decreases oxidative stress and has an opposing effect on camptothecin-induced cytotoxicity in osteosarcoma cells," *BMC Cancer,* vol. 13, p. 500, 2013.
[136] S. Chen, S. K. Rehman, W. Zhang, A. Wen, L. Yao, and J. Zhang, "Autophagy is a therapeutic target in anticancer drug resistance," *Biochim Biophys Acta,* vol. 1806, pp. 220-9, Dec 2010.
[137] S. Subramani and V. Malhotra, "Non-autophagic roles of autophagy-related proteins," *EMBO Rep,* vol. 14, pp. 143-51, Feb 2013.
[138] S. H. Ying, J. Liu, X. L. Chu, X. Q. Xie, and M. G. Feng, "The autophagy-related genes BbATG1 and BbATG8 have different functions in differentiation, stress resistance and virulence of mycopathogen Beauveria bassiana," *Sci Rep,* vol. 6, p. 26376, 2016.
[139] H. U. Simon and R. Friis, "ATG5: a distinct role in the nucleus," *Autophagy,* vol. 10, pp. 176-7, Jan 2014.
[140] D. Maskey, S. Yousefi, I. Schmid, I. Zlobec, A. Perren, R. Friis, and H. U. Simon, "ATG5 is induced by DNA-damaging agents and promotes mitotic catastrophe independent of autophagy," *Nat Commun,* vol. 4, p. 2130, 2013.

[141] Z. Zhao, B. Fux, M. Goodwin, I. R. Dunay, D. Strong, B. C. Miller, K. Cadwell, M. A. Delgado, M. Ponpuak, K. G. Green, R. E. Schmidt, N. Mizushima, V. Deretic, L. D. Sibley, and H. W. Virgin, "Autophagosome-independent essential function for the autophagy protein Atg5 in cellular immunity to intracellular pathogens," *Cell Host Microbe,* vol. 4, pp. 458-69, Nov 13 2008.

[142] J. M. Kimmey, J. P. Huynh, L. A. Weiss, S. Park, A. Kambal, J. Debnath, H. W. Virgin, and C. L. Stallings, "Unique role for ATG5 in neutrophil-mediated immunopathology during M. tuberculosis infection," *Nature,* vol. 528, pp. 565-9, Dec 24 2015.

[143] N. Jounai, F. Takeshita, K. Kobiyama, A. Sawano, A. Miyawaki, K. Q. Xin, K. J. Ishii, T. Kawai, S. Akira, K. Suzuki, and K. Okuda, "The Atg5 Atg12 conjugate associates with innate antiviral immune responses," *Proc Natl Acad Sci U S A,* vol. 104, pp. 14050-5, Aug 28 2007.

[144] S. Yousefi, R. Perozzo, I. Schmid, A. Ziemiecki, T. Schaffner, L. Scapozza, T. Brunner, and H. U. Simon, "Calpain-mediated cleavage of Atg5 switches autophagy to apoptosis," *Nat Cell Biol,* vol. 8, pp. 1124-32, Oct 2006.

[145] M. Shi, T. Zhang, L. Sun, Y. Luo, D. H. Liu, S. T. Xie, X. Y. Song, G. F. Wang, X. L. Chen, B. C. Zhou, and Y. Z. Zhang, "Calpain, Atg5 and Bak play important roles in the crosstalk between apoptosis and autophagy induced by influx of extracellular calcium," *Apoptosis,* vol. 18, pp. 435-51, Apr 2013.

[146] H. Nakatogawa, "Two ubiquitin-like conjugation systems that mediate membrane formation during autophagy," *Essays Biochem,* vol. 55, pp. 39-50, 2013.

[147] J. H. Kim and H. K. Song, "Swapping of interaction partners with ATG5 for autophagosome maturation," *BMB Rep,* vol. 48, pp. 129-30, Mar 2015.

[148] J. H. Kim, S. B. Hong, J. K. Lee, S. Han, K. H. Roh, K. E. Lee, Y. K. Kim, E. J. Choi, and H. K. Song, "Insights into autophagosome maturation revealed by the structures of ATG5 with its interacting partners," *Autophagy,* vol. 11, pp. 75-87, 2015.

[149] Y. Qiu, K. Hofmann, J. E. Coats, B. A. Schulman, and S. E. Kaiser, "Binding to E1 and E3 is mutually exclusive for the human autophagy E2 Atg3," *Protein Sci,* vol. 22, pp. 1691-7, Dec 2013.

[150] S. Erbil, O. Oral, G. Mitou, C. Kig, E. Durmaz-Timucin, E. Guven-Maiorov, F. Gulacti, G. Gokce, J. Dengjel, O. U. Sezerman, and D. Gozuacik, "RACK1 Is an Interaction Partner of ATG5 and a Novel Regulator of Autophagy," *J Biol Chem,* vol. 291, pp. 16753-65, Aug 5 2016.

[151] Z. H. Chen, J. F. Cao, J. S. Zhou, H. Liu, L. Q. Che, K. Mizumura, W. Li, A. M. Choi, and H. H. Shen, "Interaction of caveolin-1 with ATG12-ATG5 system suppresses autophagy in lung epithelial cells," *Am J Physiol Lung Cell Mol Physiol,* vol. 306, pp. L1016-25, Jun 1 2014.

[152] Z. He, H. Liu, M. Agostini, S. Yousefi, A. Perren, M. P. Tschan, T. W. Mak, G. Melino, and H. U. Simon, "p73 regulates autophagy and hepatocellular lipid metabolism through a transcriptional activation of the ATG5 gene," *Cell Death Differ,* vol. 20, pp. 1415-24, Oct 2013.

[153] E. Keil, R. Hocker, M. Schuster, F. Essmann, N. Ueffing, B. Hoffman, D. A. Liebermann, K. Pfeffer, K. Schulze-Osthoff, and I. Schmitz, "Phosphorylation of Atg5 by the Gadd45beta-MEKK4-p38 pathway inhibits autophagy," *Cell Death Differ,* vol. 20, pp. 321-32, Feb 2013.

[154] A. Vij, Randhawa, R., Parkash, J. and Changotra, H., , "Investigating regulatory signatures of human autophagy related gene 5 (ATG5) through functional in silico analysis," *Meta Gene,* vol. 9, pp. 237–248, 2016.

[155] J. O. Pyo, M. H. Jang, Y. K. Kwon, H. J. Lee, J. I. Jun, H. N. Woo, D. H. Cho, B. Choi, H. Lee, J. H. Kim, N. Mizushima, Y. Oshumi, and Y. K. Jung, "Essential roles of Atg5 and FADD in autophagic cell death: dissection of autophagic cell death into vacuole formation and cell death," *J Biol Chem,* vol. 280, pp. 20722-9, May 27 2005.

[156] T. Doring and R. Prange, "Rab33B and its autophagic Atg5/12/16L1 effector assist in hepatitis B virus naked capsid formation and release," *Cell Microbiol,* vol. 17, pp. 747-64, May 2015.

[157] C. Guevin, D. Manna, C. Belanger, K. V. Konan, P. Mak, and P. Labonte, "Autophagy protein ATG5 interacts transiently with the hepatitis C virus RNA polymerase (NS5B) early during infection," *Virology,* vol. 405, pp. 1-7, Sep 15 2010.
[158] S. Hwang, N. S. Maloney, M. W. Bruinsma, G. Goel, E. Duan, L. Zhang, B. Shrestha, M. S. Diamond, A. Dani, S. V. Sosnovtsev, K. Y. Green, C. Lopez-Otin, R. J. Xavier, L. B. Thackray, and H. W. Virgin, "Nondegradative role of Atg5-Atg12/ Atg16L1 autophagy protein complex in antiviral activity of interferon gamma," *Cell Host Microbe,* vol. 11, pp. 397-409, Apr 19 2012.
[159] F. Takeshita, K. Kobiyama, A. Miyawaki, N. Jounai, and K. Okuda, "The non-canonical role of Atg family members as suppressors of innate antiviral immune signaling," *Autophagy,* vol. 4, pp. 67-9, Jan 2008.
[160] S.-C. Su, C.-C. J. Kuo, and T. Chen, "Single nucleotide polymorphism data analysis-state-of-the-art review on this emerging field from a signal processing viewpoint," *IEEE Signal Processing Magazine,* vol. 24, pp. 75-82, 2007.
[161] K. Smith, "Genetic Polymorphism and SNPs: Genotyping, Haplotype Assembly Problem, Haplotype Map," *Functional Genomics and Proteomics,* 2002.
[162] H. M. Wenz, S. Baumhueter, S. Ramachandra, and M. Worwood, "A rapid automated SSCP multiplex capillary electrophoresis protocol that detects the two common mutations implicated in hereditary hemochromatosis (HH)," *Hum Genet,* vol. 104, pp. 29-35, Jan 1999.
[163] I. C. Gray, D. A. Campbell, and N. K. Spurr, "Single nucleotide polymorphisms as tools in human genetics," *Hum Mol Genet,* vol. 9, pp. 2403-8, Oct 2000.
[164] Z. B. Alwi, "The Use of SNPs in Pharmacogenomics Studies," *Malays J Med Sci,* vol. 12, pp. 4-12, Jul 2005.
[165] M. C. O'Donovan, P. J. Oefner, S. C. Roberts, J. Austin, B. Hoogendoorn, C. Guy, G. Speight, M. Upadhyaya, S. S. Sommer, and P. McGuffin, "Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection," *Genomics,* vol. 52, pp. 44-9, Aug 15 1998.
[166] P. Y. Kwok and X. Chen, "Detection of single nucleotide polymorphisms," *Curr Issues Mol Biol,* vol. 5, pp. 43-60, Apr 2003.
[167] H. B. Rasmussen, "Restriction fragment length polymorphism analysis of PCR-amplified fragments (PCR-RFLP) and gel electrophoresis-valuable tool for genotyping and genetic fingerprinting," in *Gel Electrophoresis-Principles and Basics*, ed: InTech, 2012.
[168] Y. M. Lo, "The amplification refractory mutation system," *Methods Mol Med,* vol. 16, pp. 61-9, 1998.
[169] H. K. Koopaee and A. E. Koshkoiyeh, "SNPs Genotyping technologies and their applications in farm animals breeding programs," *Brazilian Archives of Biology and Technology,* vol. 57, pp. 87-95, 2014.
[170] S. Jenkins and N. Gibson, "High-throughput SNP genotyping," *Comp Funct Genomics,* vol. 3, pp. 57-66, 2002.
[171] R. M. Twyman, "Single nucleotide polymorphism (SNP) genotyping techniques—an overview," *Encyclopedia of Diagnostic Genomics and Proteomics. New York, USA: Marcel Dekker, Inc,* pp. 1202-1207, 2005.
[172] W. Tan, J. Xia, Y. Dan, M. Li, S. Lin, X. Pan, H. Wang, Y. Tang, N. Liu, S. Tan, M. Liu, W. He, W. Zhang, Q. Mao, Y. Wang, and G. Deng, "Genome-wide association study identi fi es HLA-DR variants conferring risk of HBV-related acute-on-chronic liver failure," *Gut,* Jan 27 2017.
[173] M. L. Wayne and L. M. McIntyre, "Combining mapping and arraying: An approach to candidate gene identification," *Proc Natl Acad Sci U S A,* vol. 99, pp. 14903-6, Nov 12 2002.
[174] J. M. Kwon and A. M. Goate, "The candidate gene approach," *Alcohol Res Health,* vol. 24, pp. 164-8, 2000.

[175] J. M. Drazen, C. N. Yandava, L. Dube, N. Szczerback, R. Hippensteel, A. Pillari, E. Israel, N. Schork, E. S. Silverman, D. A. Katz, and J. Drajesk, "Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment," *Nat Genet,* vol. 22, pp. 168-70, Jun 1999.

[176] J. Poirier, M. C. Delisle, R. Quirion, I. Aubert, M. Farlow, D. Lahiri, S. Hui, P. Bertrand, J. Nalbantoglu, B. M. Gilfix, and S. Gauthier, "Apolipoprotein E4 allele as a predictor of cholinergic deficits and treatment outcome in Alzheimer disease," *Proc Natl Acad Sci U S A,* vol. 92, pp. 12260-4, Dec 19 1995.

[177] M. Slatkin, "Linkage disequilibrium--understanding the evolutionary past and mapping the medical future," *Nat Rev Genet,* vol. 9, pp. 477-85, Jun 2008.

[178] A. Tenesa, S. A. Knott, A. D. Carothers, and P. M. Visscher, "Power of linkage disequilibrium mapping to detect a quantitative trait locus (QTL) in selected samples of unrelated individuals," *Ann Hum Genet,* vol. 67, pp. 557-66, Nov 2003.

[179] B. Sobrino and A. Carracedo, "SNP typing in forensic genetics: a review," *Methods Mol Biol,* vol. 297, pp. 107-26, 2005.

[180] M. Parkes, J. C. Barrett, N. J. Prescott, M. Tremelling, C. A. Anderson, S. A. Fisher, R. G. Roberts, E. R. Nimmo, F. R. Cummings, D. Soars, H. Drummond, C. W. Lees, S. A. Khawaja, R. Bagnall, D. A. Burke, C. E. Todhunter, T. Ahmad, C. M. Onnie, W. McArdle, D. Strachan, G. Bethel, C. Bryan, C. M. Lewis, P. Deloukas, A. Forbes, J. Sanderson, D. P. Jewell, J. Satsangi, J. C. Mansfield, L. Cardon, and C. G. Mathew, "Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility," *Nat Genet,* vol. 39, pp. 830-2, Jul 2007.

[181] L. Henckaerts, I. Cleynen, M. Brinar, J. M. John, K. Van Steen, P. Rutgeerts, and S. Vermeire, "Genetic variation in the autophagy gene ULK1 and risk of Crohn's disease," *Inflamm Bowel Dis,* vol. 17, pp. 1392-7, Jun 2011.

[182] J. Dang, J. Li, Q. Xin, S. Shan, X. Bian, Q. Yuan, N. Liu, X. Ma, Y. Li, and Q. Liu, "Gene-gene interaction of ATG5, ATG7, BLK and BANK1 in systemic lupus erythematosus," *Int J Rheum Dis,* Sep 30 2015.

[183] M. Salem, O. H. Nielsen, K. Nys, S. Yazdanyar, and J. B. Seidelin, "Impact of T300A Variant of ATG16L1 on Antibacterial Response, Risk of Culture Positive Infections, and Clinical Course of Crohn's Disease," *Clin Transl Gastroenterol,* vol. 6, p. e122, 2015.

[184] C. D. Intemann, T. Thye, S. Niemann, E. N. Browne, M. Amanua Chinbuah, A. Enimil, J. Gyapong, I. Osei, E. Owusu-Dabo, S. Helm, S. Rusch-Gerdes, R. D. Horstmann, and C. G. Meyer, "Autophagy gene variant IRGM -261T contributes to protection from tuberculosis caused by Mycobacterium tuberculosis but not by M. africanum strains," *PLoS Pathog,* vol. 5, p. e1000577, Sep 2009.

[185] H. Saitsu, T. Nishimura, K. Muramatsu, H. Kodera, S. Kumada, K. Sugai, E. Kasai-Yoshida, N. Sawaura, H. Nishida, A. Hoshino, F. Ryujin, S. Yoshioka, K. Nishiyama, Y. Kondo, Y. Tsurusaki, M. Nakashima, N. Miyake, H. Arakawa, M. Kato, N. Mizushima, and N. Matsumoto, "De novo mutations in the autophagy gene WDR45 cause static encephalopathy of childhood with neurodegeneration in adulthood," *Nat Genet,* vol. 45, pp. 445-9, 449e1, Apr 2013.

[186] L. J. Martin, J. Gupta, S. S. Jyothula, M. Butsch Kovacic, J. M. Biagini Myers, T. L. Patterson, M. B. Ericksen, H. He, A. M. Gibson, T. M. Baye, S. Amirisetty, A. M. Tsoras, Y. Sha, N. T. Eissa, and G. K. Hershey, "Functional variant in the autophagy-related 5 gene promotor is associated with childhood asthma," *PLoS one,* vol. 7, p. e33454, 2012.

[187] D. L. Pham, S. H. Kim, P. Losol, E. M. Yang, Y. S. Shin, Y. M. Ye, and H. S. Park, "Association of autophagy related gene polymorphisms with neutrophilic airway inflammation in adult asthma," *Korean J Intern Med,* vol. 31, pp. 375-85, Mar 2016.

[188] M. Zheng, H. Yu, L. Zhang, H. Li, Y. Liu, A. Kijlstra, and P. Yang, "Association of ATG5 Gene Polymorphisms With Behcet's Disease and ATG10 Gene Polymorphisms With VKH Syndrome in a Chinese Han Population," *Invest Ophthalmol Vis Sci,* vol. 56, pp. 8280-7, Dec 2015.

[189] Y. C. Lin, P. F. Chang, H. F. Lin, K. Liu, M. H. Chang, and Y. H. Ni, "Variants in the Autophagy Related Gene IRGM Confer Susceptibility to Nonalcoholic Fatty Liver Disease by Modulating Lipophagy," *J Hepatol,* Jul 11 2016.

[190] S. Metzger, M. Saukko, H. Van Che, L. Tong, Y. Puder, O. Riess, and H. P. Nguyen, "Age at onset in Huntington's disease is modified by the autophagy pathway: implication of the V471A polymorphism in Atg7," *Hum Genet,* vol. 128, pp. 453-9, Oct 2010.

[191] A. Huijbers, T. S. Plantinga, L. A. Joosten, K. K. Aben, J. Gudmundsson, M. den Heijer, L. A. Kiemeney, M. G. Netea, A. R. Hermus, and R. T. Netea-Maier, "The effect of the ATG16L1 Thr300Ala polymorphism on susceptibility and outcome of patients with epithelial cell-derived thyroid carcinoma," *Endocr Relat Cancer,* vol. 19, pp. L15-8, Jun 2012.

[192] F. Morani, R. Titone, L. Pagano, A. Galetto, O. Alabiso, G. Aimaretti, and C. Isidoro, "Autophagy and thyroid carcinogenesis: genetic and epigenetic links," *Endocr Relat Cancer,* vol. 21, pp. R13-29, Feb 2014.

[193] Q. Tan, M. Wang, M. Yu, J. Zhang, R. G. Bristow, R. P. Hill, and I. F. Tannock, "Role of Autophagy as a Survival Mechanism for Hypoxic Cells in Tumors," *Neoplasia,* vol. 18, pp. 347-55, Jun 2016.

[194] Z. Qin, J. Xue, Y. He, H. Ma, G. Jin, J. Chen, Z. Hu, X. Liu, and H. Shen, "Potentially functional polymorphisms in ATG10 are associated with risk of breast cancer in a Chinese population," *Gene,* vol. 527, pp. 491-5, Sep 25 2013.

[195] P. P. Cai, H. X. Wang, J. C. Zhuang, Q. B. Liu, G. X. Zhao, Z. X. Li, and Z. Y. Wu, "Variants of autophagy-related gene 5 are associated with neuromyelitis optica in the Southern Han Chinese population," *Autoimmunity,* vol. 47, pp. 563-6, Dec 2014.

[196] X. J. Zhou, X. L. Lu, J. C. Lv, H. Z. Yang, L. X. Qin, M. H. Zhao, Y. Su, Z. G. Li, and H. Zhang, "Genetic association of PRDM1-ATG5 intergenic region and autophagy with systemic lupus erythematosus in a Chinese population," *Ann Rheum Dis,* vol. 70, pp. 1330-7, Jul 2011.

[197] D. Chen, C. Zhu, X. Wang, X. Feng, S. Pang, W. Huang, R. G. Hawley, and B. Yan, "A novel and functional variant within the ATG5 gene promoter in sporadic Parkinson's disease," *Neurosci Lett,* vol. 538, pp. 49-53, Mar 22 2013.

[198] S. C. Garman, "Structure-function relationships in alpha-galactosidase A," *Acta Paediatr,* vol. 96, pp. 6-16, Apr 2007.

[199] C. G. Doss, B. Rajith, N. Garwasis, P. R. Mathew, A. S. Raju, K. Apoorva, D. William, N. R. Sadhana, T. Himani, and I. P. Dike, "Screening of mutations affecting protein stability and dynamics of FGFR1-A simulation analysis," *Appl Transl Genom,* vol. 1, pp. 37-43, Dec 01 2012.

[200] C. G. Doss and N. Nagasundaram, "Investigating the structural impacts of I64T and P311S mutations in APE1-DNA complex: a molecular dynamics approach," *PLoS one,* vol. 7, p. e31677, 2012.

[201] P. Kumar, S. Henikoff, and P. C. Ng, "Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm," *Nat Protoc,* vol. 4, pp. 1073-81, 2009.

[202] I. A. Adzhubei, S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov, and S. R. Sunyaev, "A method and server for predicting damaging missense mutations," *Nat Methods,* vol. 7, pp. 248-9, Apr 2010.

[203] I. Adzhubei, D. M. Jordan, and S. R. Sunyaev, "Predicting functional effect of human missense mutations using PolyPhen-2," *Curr Protoc Hum Genet,* vol. Chapter 7, p. Unit7 20, Jan 2013.

[204] J. Bendl, J. Stourac, O. Salanda, A. Pavelka, E. D. Wieben, J. Zendulka, J. Brezovsky, and J. Damborsky, "PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations," *PLoS Comput Biol,* vol. 10, p. e1003440, Jan 2014.

[205] G. A. Thorisson, A. V. Smith, L. Krishnan, and L. D. Stein, "The International HapMap Project Web site," *Genome Res,* vol. 15, pp. 1592-3, Nov 2005.
[206] J. Ren, L. Wen, X. Gao, C. Jin, Y. Xue, and X. Yao, "CSS-Palm 2.0: an updated software for palmitoylation sites prediction," *Protein Eng Des Sel,* vol. 21, pp. 639-44, Nov 2008.
[207] Z. Liu, J. Cao, X. Gao, Q. Ma, J. Ren, and Y. Xue, "GPS-CCD: a novel computational program for the prediction of calpain cleavage sites," *PLoS one,* vol. 6, p. e19001, Apr 20 2011.
[208] W. Liu, Y. Xie, J. Ma, X. Luo, P. Nie, Z. Zuo, U. Lahrmann, Q. Zhao, Y. Zheng, Y. Zhao, Y. Xue, and J. Ren, "IBS: an illustrator for the presentation and visualization of biological sequences," *Bioinformatics,* vol. 31, pp. 3359-61, Oct 15 2015.
[209] A. Vij, R. M. Yennamalli, and H. Changotra, "Non-synonymous single nucleotide polymorphisms of ATG5 destabilize ATG12–ATG5/ATG16L1 complex: An enzyme with E3 like activity of ubiquitin conjugation system," *Meta Gene,* vol. 13, pp. 38-47, 2017.
[210] V. Ramensky, P. Bork, and S. Sunyaev, "Human non-synonymous SNPs: server and survey," *Nucleic Acids Res,* vol. 30, pp. 3894-900, Sep 01 2002.
[211] B. Li, V. G. Krishnan, M. E. Mort, F. Xin, K. K. Kamati, D. N. Cooper, S. D. Mooney, and P. Radivojac, "Automated inference of molecular mechanisms of disease from amino acid substitutions," *Bioinformatics,* vol. 25, pp. 2744-50, Nov 01 2009.
[212] H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, and N. Ben-Tal, "ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules," *Nucleic Acids Res,* vol. 44, pp. W344-50, Jul 08 2016.
[213] E. Capriotti, P. Fariselli, and R. Casadio, "I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure," *Nucleic Acids Res,* vol. 33, pp. W306-10, Jul 01 2005.
[214] L. T. Huang, M. M. Gromiha, S. F. Hwang, and S. Y. Ho, "Knowledge acquisition and development of accurate rules for predicting protein stability changes," *Comput Biol Chem,* vol. 30, pp. 408-15, Dec 2006.
[215] L. T. Huang and M. M. Gromiha, "Reliable prediction of protein thermostability change upon double mutation from amino acid sequence," *Bioinformatics,* vol. 25, pp. 2181-7, Sep 01 2009.
[216] D. E. Pires, D. B. Ascher, and T. L. Blundell, "DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach," *Nucleic Acids Res,* vol. 42, pp. W314-9, Jul 2014.
[217] V. Parthiban, M. M. Gromiha, and D. Schomburg, "CUPSAT: prediction of protein stability upon point mutations," *Nucleic Acids Res,* vol. 34, pp. W239-42, Jul 01 2006.
[218] C. Magyar, M. M. Gromiha, G. Pujadas, G. E. Tusnady, and I. Simon, "SRide: a server for identifying stabilizing residues in proteins," *Nucleic Acids Res,* vol. 33, pp. W303-5, Jul 01 2005.
[219] M. M. Gromiha, G. Pujadas, C. Magyar, S. Selvaraj, and I. Simon, "Locating the stabilizing residues in (alpha/beta)8 barrel proteins based on hydrophobicity, long-range interactions, and sequence conservation," *Proteins,* vol. 55, pp. 316-29, May 01 2004.
[220] M. Ryan, M. Diekhans, S. Lien, Y. Liu, and R. Karchin, "LS-SNP/PDB: annotated non-synonymous SNPs mapped to Protein Data Bank structures," *Bioinformatics,* vol. 25, pp. 1431-2, Jun 01 2009.
[221] N. Guex and M. C. Peitsch, "SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling," *Electrophoresis,* vol. 18, pp. 2714-23, Dec 1997.
[222] T. A. de Beer, K. Berka, J. M. Thornton, and R. A. Laskowski, "PDBsum additions," *Nucleic Acids Res,* vol. 42, pp. D292-6, Jan 2014.
[223] Y. Zhang and J. Skolnick, "TM-align: a protein structure alignment algorithm based on the TM-score," *Nucleic Acids Res,* vol. 33, pp. 2302-9, 2005.

[224] E. Mo, S. E. Peters, C. Willers, D. J. Maskell, and I. G. Charles, "Single, double and triple mutants of Salmonella enterica serovar Typhimurium degP (htrA), degQ (hhoA) and degS (hhoB) have diverse phenotypes on exposure to elevated temperature and their growth in vivo is attenuated to different extents," *Microb Pathog,* vol. 41, pp. 174-82, Oct-Nov 2006.
[225] H. Gatanaga, H. Ode, A. Hachiya, T. Hayashida, H. Sato, and S. Oka, "Combination of V106I and V179D polymorphic mutations in human immunodeficiency virus type 1 reverse transcriptase confers resistance to efavirenz and nevirapine but not etravirine," *Antimicrob Agents Chemother,* vol. 54, pp. 1596-602, Apr 2010.
[226] Y. Dehouck, J. M. Kwasigroch, M. Rooman, and D. Gilis, "BeAtMuSiC: Prediction of changes in protein-protein binding affinity on mutations," *Nucleic Acids Res,* vol. 41, pp. W333-9, Jul 2013.
[227] S. A. Miller, D. D. Dykes, and H. F. Polesky, "A simple salting out procedure for extracting DNA from human nucleated cells," *Nucleic Acids Res,* vol. 16, p. 1215, Feb 11 1988.
[228] Z. Li, Z. Zhang, Z. He, W. Tang, T. Li, Z. Zeng, L. He, and Y. Shi, "A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis (http://analysis.bio-x.cn)," *Cell Res,* vol. 19, pp. 519-23, Apr 2009.
[229] Y. Y. Shi and L. He, "SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci," *Cell Res,* vol. 15, pp. 97-8, Feb 2005.
[230] J. Shen, Z. Li, J. Chen, Z. Song, Z. Zhou, and Y. Shi, "SHEsisPlus, a toolset for genetic studies on polyploid species," *Sci Rep,* vol. 6, p. 24095, Apr 06 2016.
[231] T. Mavroconstanti, S. Johansson, I. Winge, P. M. Knappskog, and J. Haavik, "Functional properties of rare missense variants of human CDH13 found in adult attention deficit/hyperactivity disorder (ADHD) patients," *PLoS one,* vol. 8, p. e71445, 2013.
[232] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *J Mol Biol,* vol. 215, pp. 403-410, 1990.
[233] R. Owczarzy, A. V. Tataurov, Y. Wu, J. A. Manthey, K. A. McQuisten, H. G. Almabrazi, K. F. Pedersen, Y. Lin, J. Garretson, N. O. McEntaggart, C. A. Sailor, R. B. Dawson, and A. S. Peek, "IDT SciTools: a suite for analysis and design of nucleic acid oligomers," *Nucleic Acids Res,* vol. 36, pp. W163-9, Jul 01 2008.
[234] C. M. Yates and M. J. Sternberg, "The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein-protein interactions," *J Mol Biol,* vol. 425, pp. 3949-63, Nov 01 2013.
[235] A. Rodriguez-Casado, "In silico investigation of functional nsSNPs an approach to rational drug design," *Research and Reports in Medicinal Chemistry,* vol. 2, pp. 31-42, 2012.
[236] D. E. Reich, M. Cargill, S. Bolk, J. Ireland, P. C. Sabeti, D. J. Richter, T. Lavery, R. Kouyoumjian, S. F. Farhadian, R. Ward, and E. S. Lander, "Linkage disequilibrium in the human genome," *Nature,* vol. 411, pp. 199-204, May 10 2001.
[237] G. Duan and D. Walther, "The roles of post-translational modifications in the context of protein interaction networks," *PLoS Comput Biol,* vol. 11, p. e1004049, Feb 2015.
[238] A. Mastrangelo, T. Colasanti, C. Barbati, A. Pecani, D. Sabatinelli, M. Pendolino, S. Truglia, L. Massaro, R. Mancini, F. Miranda, F. R. Spinelli, F. Conti, and C. Alessandri, "The Role of Posttranslational Protein Modifications in Rheumatological Diseases: Focus on Rheumatoid Arthritis," *J Immunol Res,* vol. 2015, p. 712490, 2015.
[239] B. Sun, M. Zhang, P. Cui, H. Li, J. Jia, Y. Li, and L. Xie, "Nonsynonymous Single-Nucleotide Variations on Some Posttranslational Modifications of Human Proteins and the Association with Diseases," *Comput Math Methods Med,* vol. 2015, p. 124630, 2015.
[240] S. Benzeno, F. Lu, M. Guo, O. Barbash, F. Zhang, J. G. Herman, P. S. Klein, A. Rustgi, and J. A. Diehl, "Identification of mutations that disrupt phosphorylation-dependent nuclear export of cyclin D1," *Oncogene,* vol. 25, pp. 6291-303, Oct 12 2006.

[241] J. Ren, C. Jiang, X. Gao, Z. Liu, Z. Yuan, C. Jin, L. Wen, Z. Zhang, Y. Xue, and X. Yao, "PhosSNP for systematic analysis of genetic polymorphisms that influence protein phosphorylation," *Mol Cell Proteomics,* vol. 9, pp. 623-34, Apr 2010.

[242] P. Radivojac, P. H. Baenziger, M. G. Kann, M. E. Mort, M. W. Hahn, and S. D. Mooney, "Gain and loss of phosphorylation sites in human cancer," *Bioinformatics,* vol. 24, pp. i241-7, Aug 15 2008.

[243] G. M. Ryu, P. Song, K. W. Kim, K. S. Oh, K. J. Park, and J. H. Kim, "Genome-wide analysis to predict protein sequence variations that change phosphorylation sites or their corresponding kinases," *Nucleic Acids Res,* vol. 37, pp. 1297-307, Mar 2009.

[244] C. Salaun, J. Greaves, and L. H. Chamberlain, "The intracellular dynamic of protein palmitoylation," *J Cell Biol,* vol. 191, pp. 1229-38, Dec 27 2010.

[245] Y. Fukata and M. Fukata, "Protein palmitoylation in neuronal development and synaptic plasticity," *Nat Rev Neurosci,* vol. 11, pp. 161-75, Mar 2010.

[246] A. Zohaib, X. Duan, B. Zhu, J. Ye, S. Wan, H. Chen, X. Liu, and S. Cao, "The Role of Ubiquitination in Regulation of Innate Immune Signaling," *Curr Issues Mol Biol,* vol. 18, pp. 1-10, 2016.

[247] D. Popovic, D. Vucic, and I. Dikic, "Ubiquitination in disease pathogenesis and treatment," *Nat Med,* vol. 20, pp. 1242-53, Nov 2014.

[248] M. P. O'Connell, J. L. Fiori, K. M. Baugher, F. E. Indig, A. D. French, T. C. Camilli, B. P. Frank, R. Earley, K. S. Hoek, J. H. Hasskamp, E. G. Elias, D. D. Taub, M. Bernier, and A. T. Weeraratna, "Wnt5A activates the calpain-mediated cleavage of filamin A," *J Invest Dermatol,* vol. 129, pp. 1782-9, Jul 2009.

[249] H. Patzke and L. H. Tsai, "Calpain-mediated cleavage of the cyclin-dependent kinase-5 activator p39 to p29," *J Biol Chem,* vol. 277, pp. 8054-60, Mar 08 2002.

[250] M. Songane, J. Kleinnijenhuis, B. Alisjahbana, E. Sahiratmadja, I. Parwati, M. Oosting, T. S. Plantinga, L. A. Joosten, M. G. Netea, T. H. Ottenhoff, E. van de Vosse, and R. van Crevel, "Polymorphisms in autophagy genes and susceptibility to tuberculosis," *PLoS one,* vol. 7, p. e41618, 2012.

[251] M. G. Cucu, I. Streața, A. L. Riza, A. L. Cimpoeru, S. Șerban-Șoșoi, A. Ciocoiu, R. M. Pleșea, E. L. Popescu, Ș. Dorobanțu, and A. Anghel, "Polymorphisms in autophagy genes and active pulmonary tuberculosis susceptibility in Romania," *Revista Romana de Medicina de Laborator,* vol. 25, pp. 47-53, 2017.

[252] V. Gateva, J. K. Sandling, G. Hom, K. E. Taylor, S. A. Chung, X. Sun, W. Ortmann, R. Kosoy, R. C. Ferreira, G. Nordmark, I. Gunnarsson, E. Svenungsson, L. Padyukov, G. Sturfelt, A. Jonsen, A. A. Bengtsson, S. Rantapaa-Dahlqvist, E. C. Baechler, E. E. Brown, G. S. Alarcon, J. C. Edberg, R. Ramsey-Goldman, G. McGwin, Jr., J. D. Reveille, L. M. Vila, R. P. Kimberly, S. Manzi, M. A. Petri, A. Lee, P. K. Gregersen, M. F. Seldin, L. Ronnblom, L. A. Criswell, A. C. Syvanen, T. W. Behrens, and R. R. Graham, "A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus," *Nat Genet,* vol. 41, pp. 1228-33, Nov 2009.

[253] J. Fernandez-Mateos, R. Seijas-Tamayo, J. C. A. Klain, M. P. Borgonon, E. Perez-Ruiz, R. Mesia, E. Del Barco, C. S. Coloma, A. R. Dominguez, J. C. Daroqui, E. Fernandez Ruiz, J. J. Cruz-Hernandez, and R. Gonzalez-Sarmiento, "Analysis of autophagy gene polymorphisms in Spanish patients with head and neck squamous cell carcinoma," *Sci Rep,* vol. 7, p. 6887, Jul 31 2017.

[254] K. Buffen, M. Oosting, J. Quintin, A. Ng, J. Kleinnijenhuis, V. Kumar, E. van de Vosse, C. Wijmenga, R. van Crevel, E. Oosterwijk, A. J. Grotenhuis, S. H. Vermeulen, L. A. Kiemeney, F. L. van de Veerdonk, G. Chamilos, R. J. Xavier, J. W. van der Meer, M. G. Netea, and L. A. Joosten, "Autophagy controls BCG-induced trained immunity and the response to intravesical BCG therapy for bladder cancer," *PLoS Pathog,* vol. 10, p. e1004485, Oct 2014.

[255] O. Iemets, "The Single-Nucleotide Polymorphism rs510432 in ATG5 Gene and the Development of Atopic March in Children," *CHILDS HEALTH,* pp. 14-20, 2016.

[256] K. A. White, L. Luo, T. A. Thompson, S. Torres, C. A. Hu, N. E. Thomas, J. Lilyquist, H. Anton-Culver, S. B. Gruber, L. From, K. J. Busam, I. Orlow, P. A. Kanetsky, L. D. Marrett, R. P. Gallagher, L. Sacchetto, S. Rosso, T. Dwyer, A. E. Cust, C. B. Begg, and M. Berwick, "Variants in autophagy-related genes and clinical characteristics in melanoma: a population-based study," *Cancer Med,* vol. 5, pp. 3336-3345, Nov 2016.

[257] J. Yuan, R. Han, A. Esther, Q. Wu, J. Yang, W. Yan, X. Ji, Y. Liu, Y. Li, W. Yao, and C. Ni, "Polymorphisms in autophagy related genes and the coal workers' pneumoconiosis in a Chinese population," *Gene,* Aug 24 2017.

[258] M. Dezelak, K. Repnik, S. Koder, I. Ferkolj, and U. Potocnik, "A Prospective Pharmacogenomic Study of Crohn's Disease Patients during Routine Therapy with Anti-TNF-alpha Drug Adalimumab: Contribution of ATG5, NFKB1, and CRP Genes to Pharmacodynamic Variability," *OMICS,* vol. 20, pp. 296-309, May 2016.

[259] Y. Shao, F. Chen, Y. Chen, W. Zhang, Y. Lin, Y. Cai, Z. Yin, S. Tao, Q. Liao, J. Zhao, H. Mai, Y. He, J. He, and L. Cui, "Association between genetic polymorphisms in the autophagy-related 5 gene promoter and the risk of sepsis," *Sci Rep,* vol. 7, p. 9399, Aug 24 2017.

[260] A. Poon, D. Eidelman, C. Laprise, and Q. Hamid, "ATG5, autophagy and lung function in asthma," *Autophagy,* vol. 8, pp. 694-5, Apr 2012.

Chapter 8

APPENDICES

APPENDIX-A

Table 8.1: The linkage disequilibrium table representing important SNPs.

Sr. No	Name	ObsHET	PhredHET	Hwpval	%Geno	MAF	Alleles
1	rs9386514	0.339	0.347	1	90	0.223	T:C
2	rs6920944	0.169	0.155	1	97.8	0.085	C:T
3	rs9486301	0.328	0.274	0.3426	96.7	0.164	T:C
4	rs4945747	0.383	0.31	0.1458	100	0.192	C:T
5	rs11754416	0.15	0.139	1	100	0.075	C:T
6	rs9486302	0.276	0.238	0.6031	97.8	0.138	C:T
7	rs12201458	0.167	0.18	0.9127	100	0.1	C:A
8	rs1885450	0.383	0.31	0.1458	100	0.192	T:C
9	rs698029	0.467	0.486	0.9107	100	0.417	G:A
10	rs11759048	0.164	0.15	1	86.7	0.082	C:T
11	rs9486305	0.293	0.25	0.5066	93.3	0.147	C:T
12	rs9373839	0.339	0.347	1	90	0.223	T:C
13	rs11751513	0.083	0.08	1	100	0.042	C:A
14	rs6930834	0.145	0.135	1	88.9	0.073	G:C
15	rs2743556	0.517	0.489	0.9112	100	0.425	G:A
16	rs9486306	0.321	0.27	0.3925	90	0.161	T:G
17	rs2245214	0.525	0.491	0.8338	98.9	0.432	C:G
18	rs2299860	0.214	0.191	1	90	0.107	T:C
19	rs11756379	0.15	0.139	1	100	0.075	T:C
20	rs617994	0.483	0.489	1	100	0.425	G:T
21	rs9372120	0.39	0.371	1	98.9	0.246	T:G
22	rs2299863	0.179	0.191	0.9644	90	0.107	T:G
23	rs1626224	0.15	0.139	1	100	0.075	T:C
24	rs17587319	0.217	0.193	0.9665	100	0.108	C:G
25	rs4945748	0.217	0.193	0.9665	100	0.108	G:C
26	rs9386516	0.35	0.349	1	100	0.225	A:G
27	rs665791	0.482	0.496	0.9936	90	0.455	A:G

28	rs1769972	0.143	0.133	1	90	0.071	T:C
29	rs1769971	0.098	0.093	1	81.1	0.049	A:G
30	rs1769970	0.17	0.155	1	88.9	0.085	G:T
31	rs594819	0.508	0.488	1	98.9	0.424	T:G
32	rs2787521	0.138	0.128	1	95.6	0.069	T:C
33	rs2787525	0.143	0.133	1	90	0.071	T:C
34	rs12529488	0.228	0.202	0.9257	94.4	0.114	C:T
35	rs3827644	0.321	0.337	0.9537	90	0.214	G:C
36	rs3804329	0.35	0.349	1	100	0.225	A:G
37	rs1475270	0.482	0.496	0.9936	88.9	0.455	G:A
38	rs2787528	0.167	0.153	1	100	0.083	G:C
39	rs2743561	0.017	0.017	1	100	0.008	C:T
40	rs13203112	0.017	0.017	1	97.8	0.009	C:T
41	rs2787529	0.15	0.139	1	100	0.075	G:A
42	rs2743562	0.15	0.139	1	100	0.075	G:C
43	rs796707	0.161	0.148	1	90	0.08	G:A
44	rs13191943	0.017	0.017	1	98.9	0.008	T:C
45	rs1769974	0.15	0.139	1	100	0.075	C:T
46	rs12526431	0.214	0.191	1	90	0.107	C:T
47	rs2743564	0.153	0.141	1	97.8	0.076	C:T
48	rs1769981	0.143	0.133	1	90	0.071	G:T
49	rs3804332	0.2	0.18	1	88.9	0.1	T:C
50	rs9398075	0.316	0.332	0.9246	95.6	0.211	G:A
51	rs10484576	0.217	0.193	0.9665	100	0.108	G:A
52	rs3851210	0.161	0.148	1	90	0.08	C:T
53	rs9486312	0.317	0.267	0.3675	100	0.158	A:C
54	rs10484577	0.35	0.349	1	100	0.225	A:C
55	rs2757133	0.15	0.139	1	100	0.075	C:A
56	rs3804333	0.321	0.337	0.9537	90	0.214	C:T
57	rs1624701	0.143	0.133	1	90	0.071	C:T
58	rs6906688	0.083	0.11	0.338	100	0.058	T:A
59	rs524428	0.5	0.494	1	90	0.446	C:T

60	rs17513225	0.136	0.126	1	97.8	0.068	A:C
61	rs2757132	0.15	0.139	1	100	0.075	A:G
62	rs9373842	0.351	0.352	1	95.6	0.228	T:G
63	rs633724	0.518	0.481	0.818	90	0.402	C:T
64	rs9372121	0.35	0.349	1	100	0.225	A:G
65	rs9486314	0.411	0.384	0.9352	90	0.259	C:T
66	rs484621	0.482	0.492	1	90	0.438	C:T
67	rs2787538	0.161	0.148	1	90	0.08	A:G
68	rs12529626	0.232	0.205	0.9114	90	0.116	A:G
69	rs2787540	0.143	0.133	1	90	0.071	A:G
70	rs547738	0.473	0.492	0.9324	88.9	0.436	A:G
71	rs1624009	0.161	0.148	1	90	0.08	A:T
72	rs490010	0.482	0.492	1	90	0.438	G:A
73	rs1766200	0.161	0.148	1	90	0.08	C:T
74	rs1769962	0.153	0.141	1	97.8	0.076	T:C
75	rs10499049	0.2	0.18	1	100	0.1	A:G
76	rs9486315	0.339	0.282	0.3173	90	0.17	T:C
77	rs2763219	0.518	0.473	0.7223	90	0.384	A:G
78	rs543465	0.15	0.139	1	100	0.075	C:T
79	rs12212740	0.179	0.191	0.9644	90	0.107	G:A
80	rs17066720	0.018	0.018	1	90	0.009	G:C
81	rs2787545	0.15	0.139	1	100	0.075	G:A
82	rs9320155	0.317	0.267	0.3675	100	0.158	C:G
83	rs1766208	0.37	0.302	0.2296	87.8	0.185	C:T
84	rs2791116	0.133	0.124	1	100	0.067	T:G
85	rs9373843	0.368	0.301	0.2078	96.7	0.184	A:G
86	rs626664	0.15	0.139	1	100	0.075	T:C
87	rs3804334	0.333	0.278	0.2984	100	0.167	C:T
88	rs3804335	0.133	0.124	1	100	0.067	C:T
89	rs9399980	0.133	0.124	1	100	0.067	T:C
90	rs2787548	0.136	0.126	1	98.9	0.068	T:C
91	rs2787549	0.15	0.139	1	100	0.075	G:C

92	rs9486319	0.298	0.254	0.4924	95.6	0.149	G:A
93	rs2763214	0.133	0.153	0.6697	100	0.083	T:A
94	rs671116	0.518	0.473	0.7223	90	0.384	A:G
95	rs17066765	0.017	0.017	1	100	0.008	T:C
96	rs2763212	0.161	0.148	1	90	0.08	A:G
97	rs2787552	0.121	0.113	1	95.6	0.06	C:T
98	rs573775	0.357	0.392	0.6745	90	0.268	G:A
99	rs803360	0.483	0.493	1	100	0.442	C:G
100	rs17066773	0.018	0.018	1	90	0.009	T:C
101	rs580187	0.354	0.291	0.357	82.2	0.177	T:A
102	rs12526668	0.196	0.177	1	90	0.098	C:T
103	rs597443	0.464	0.494	0.8005	90	0.446	C:T
104	rs2763221	0.167	0.153	1	83.3	0.083	T:C
105	rs3804338	0.089	0.085	1	90	0.045	C:T

Table 8.2. Amino Acid Conservation Scores obtained for ATG5 from the ConSurf web server.

POS	SEQ	3LATOM	SCORE (normalized)	COLOR	CONFIDENCE INTERVAL	CONFIDENCE INTERVAL COLORS	MSA DATA	RESIDUE VARIETY
1	D	ASP3:B	-0.474	7	-0.681,-0.364	7,6	107/150	N,E,S,D
2	D	ASP4:B	-1.196	9	-1.275,-1.162	9,9	136/150	D,N
3	K	LYS5:B	-0.577	7	-0.734,-0.436	8,7	136/150	K,R,D,M,A,E,N,S,T,Y
4	D	ASP6:B	-0.950	8	-1.087,-0.879	9,8	137/150	N,Q,D,A,E
5	V	VAL7:B	-0.522	7	-0.681,-0.436	7,7	139/150	L,M,Q,V,A,I
6	L	LEU8:B	-0.442	7	-0.625,-0.286	7,6	139/150	S,T,I,E,L,R,Q,M,P,C
7	R	ARG9:B	-0.922	8	-1.048,-0.832	9,8	140/150	Q,R,K,E,H
8	D	ASP10:B	-0.180	6	-0.436,-0.006	7,5	140/150	I,A,H,E,R,K,Q,M,D,G,C,V,Y,S,N
9	V	VAL11:B	-0.244	6	-0.436,-0.108	7,5	140/150	M,T,L,V,I
10	W	TRP12:B	-1.100	9	-1.236,-1.048	9,9	140/150	W,R
11	F	PHE13:B	-0.184	6	-0.436,-0.006	7,5	140/150	N,D,S,K,E,Y,F
12	G	GLY14:B	-0.915	8	-1.048,-0.832	9,8	140/150	A,S,G
13	R	ARG15:B	-0.288	6	-0.503,-0.201	7,6	141/150	S,T,Y,R,K,Q,A,H
14	I	ILE16:B	-0.265	6	-0.503,-0.108	7,5	141/150	V,I,L
15	P	PRO17:B	-1.031	9	-1.162,-0.966	9,8	141/150	P,G,A
16	T	THR18:B	-0.541	7	-0.681,-0.436	7,7	141/150	V,I,A,S,L,T,C
17	C	CYS19:B	0.286	4	-0.006, 0.394	5,4	141/150	S,Y,V,Q,R,L,K,E,A,H,C
18	F	PHE20:B	-0.879	8	-1.007,-0.784	9,8	141/150	L,V,I,F
19	T	THR21:B	0.725	2	0.394, 0.808	4,2	141/150	C,I,H,E,L,R,K,M,Q,N,V,F,S,T
20	L	LEU22:B	-0.864	8	-1.007,-0.784	9,8	141/150	L,T,V,F,I,A
21	Y	TYR23:B	-0.001	5	-0.286, 0.110	6,5	141/150	N,Y,T,S,C,E,H,A,D,Q,L
22	Q	GLN24:B	1.055	1	0.577, 1.115	3,1	141/150	L,K,R,D,Q,I,A,E,G,P,S,T,V,N
23	D	ASP25:B	0.257	4	-0.006, 0.394	5,4	141/150	S,T,V,Y,N,R,L,Q,D,A,H,E,C
24	E	GLU26:B	-0.860	8	-1.007,-0.784	9,8	139/150	N,G,D,Q,E
25	I	ILE27:B	0.466	3	0.110, 0.577	5,3	138/150	G,P,C,E,A,H,I,M,D,L,Y,F,V,T,S
26	T	THR28:B	0.757	2	0.394, 1.115	4,1	138/150	N,V,S,T,G,A,I,L,Q,M,F,Y,P,C,H,E,D
27	E	GLU29:B	0.201	4	-0.108, 0.394	5,4	115/150	C,G,D,L,E,A,N,T,S,F,V
28	R	ARG30:B	0.304	4	-0.006, 0.577	5,3	136/150	V,S,T,I,H,E,K,R,Q,D,G,C
29	E	GLU31:B	0.680	3	0.241, 0.808	4,2	139/150	V,S,T,N,I,A,L,K,R,Q,G,F,H,E,D,P
30	A	ALA32:B	-0.848	8	-1.007,-0.784	9,8	139/150	T,S,V,W,A,C,P
31	E	GLU33:B	-0.092	5	-0.364, 0.110	6,5	140/150	N,S,P,A,E,K,L,R,D
32	P	PRO34:B	-0.316	6	-0.503,-0.201	7,6	140/150	N,S,T,V,P,R,L,K,Q,D,A,E
33	Y	TYR35:B	1.066	1	0.577, 1.586	3,1	140/150	L,S,V,I,F,Y,C
34	Y	TYR36:B	-0.386	6	-0.625,-0.286	7,6	140/150	C,Y,F,V,T
35	L	LEU37:B	-0.519	7	-0.734,-0.364	8,6	141/150	L,M,T,A,I,V,Y

36	L	LEU38:B	-0.792	8	-0.923,-0.734	8,8	143/150	M,L,E,V,A,F,I
37	L	LEU39:B	-0.601	7	-0.784,-0.503	8,7	143/150	C,V,A,I,L
38	P	PRO40:B	-0.762	8	-0.923,-0.681	8,7	143/150	P,R,S,L,M
39	R	ARG41:B	-1.130	9	-1.236,-1.087	9,9	143/150	P,G,E,R,S
40	V	VAL42:B	0.674	3	0.241, 0.808	4,2	143/150	I,L,M,Q,G,C,F,V,Y,S,N
41	S	SER43:B	-0.999	9	-1.087,-0.966	9,8	143/150	N,G,I,A,T,S
42	Y	TYR44:B	-1.236	9	-1.296,-1.199	9,9	143/150	Y
43	L	LEU45:B	-0.471	7	-0.681,-0.364	7,6	143/150	I,F,L
44	T	THR46:B	-0.582	7	-0.784,-0.503	8,7	143/150	P,T,M,S,A
45	L	LEU47:B	-0.836	8	-1.007,-0.734	9,8	143/150	V,A,I,L,M
46	V	VAL48:B	-0.485	7	-0.681,-0.364	7,6	143/150	T,F,V,Y,C,L,I,A,H
47	T	THR49:B	-0.497	7	-0.681,-0.364	7,6	144/150	L,M,A,W,I,C,S,T,V,Y
48	D	ASP50:B	-0.282	6	-0.503,-0.108	7,5	144/150	S,T,N,D,Q,A,E,G,P
49	K	LYS51:B	-0.831	8	-0.966,-0.734	8,8	144/150	A,E,K,R,Q,P
50	V	VAL52:B	-0.586	7	-0.734,-0.503	8,7	144/150	C,I,A,V,M,L,S
51	K	LYS53:B	0.565	3	0.241, 0.808	4,2	144/150	I,H,E,K,L,R,Q,N,F,V,Y,S
52	K	LYS54:B	-0.172	6	-0.436,-0.006	7,5	143/150	V,X,T,S,N,E,D,Q,R,K,G,P
53	H	HIS55:B	-0.389	6	-0.566,-0.286	7,6	144/150	Y,F,T,S,E,I,H,Q,R,L
54	F	PHE56:B	-0.821	8	-0.966,-0.734	8,8	143/150	A,I,M,L,Y,V,F,S
55	Q	GLN57:B	1.343	1	0.808, 1.586	2,1	142/150	N,V,T,X,S,G,A,H,I,Q,M,R,L,K
56	K	LYS58:B	0.331	4	-0.006, 0.577	5,3	142/150	V,S,T,N,I,H,A,R,K,Q,D,G
57	V	VAL59:B	1.053	1	0.577, 1.115	3,1	142/150	C,L,R,D,M,Q,A,H,I,S,T,F,V,Y
58	M	MET60:B	0.288	4	-0.006, 0.394	5,4	143/150	C,G,M,D,L,K,E,A,I,T,S,V
59	R	ARG61:B	0.994	1	0.577, 1.115	3,1	143/150	F,V,S,T,N,A,I,H,E,L,K,R,D,Q,P
60	Q	GLN62:B	2.393	1	1.115, 2.828	1,1	143/150	H,E,D,P,C,Y,A,I,K,R,L,M,Q,G,V,S,T,N
61	E	GLU63:B	1.095	1	0.577, 1.586	3,1	143/150	G,P,D,L,K,R,E,A,I,T,S,F,V
62	D	ASP64:B	0.833	2	0.394, 1.115	4,1	137/150	N,T,S,V,Q,R,L,K,I,A,F,C,P,D,E,H
63	I	ILE65:B	1.063	1	0.577, 1.115	3,1	143/150	V,S,T,N,A,I,R,K,M,Q,G,F,Y,H,E,D,P
64	S	SER66:B	1.643	1	1.115, 1.586	1,1	143/150	Y,S,T,N,H,I,A,E,K,R,M,D,Q,G,P
65	E	GLU67:B	-0.396	6	-0.566,-0.286	7,6	144/150	K,Q,D,H,A,E,P,S,T,V,N
66	I	ILE68:B	-0.172	6	-0.364,-0.006	6,5	144/150	P,M,L,V,F,I,A
67	W	TRP69:B	-0.991	9	-1.162,-0.879	9,8	144/150	W,P,G
68	F	PHE70:B	-0.113	5	-0.364,-0.006	6,5	144/150	V,F,Y,L,R,M,I,A,C
69	E	GLU71:B	-0.602	7	-0.784,-0.503	8,7	144/150	K,R,S,Q,D,V,E
70	Y	TYR72:B	1.021	1	0.577, 1.115	3,1	144/150	N,F,V,Y,S,T,C,H,L,Q
71	E	GLU73:B	-0.376	6	-0.566,-0.286	7,6	143/150	E,A,D,Q,L,R,K,G,X,S,N
72	G	GLY74:B	-0.197	6	-0.436,-0.006	7,5	144/150	S,T,N,K,R,L,H,G,P
73	T	THR75:B	0.727	2	0.394, 0.808	4,2	144/150	S,T,F,V,N,L,K,Q,M,I,A,H,E
74	P	PRO76:B	-0.900	8	-1.048,-0.832	9,8	144/150	T,L,A,P

75	L	LEU77:B	-0.692	7	-0.879,-0.566	8,7	144/150	V,I,T,Q,M,L
76	K	LYS78:B	-0.876	8	-1.007,-0.784	9,8	144/150	Q,R,K,P,N
77	W	TRP79:B	-0.477	7	-0.734,-0.286	8,6	144/150	G,M,T,L,I,V,W
78	H	HIS80:B	-0.939	8	-1.048,-0.879	9,8	149/150	N,Q,I,H,Y
79	Y	TYR81:B	0.135	5	-0.201, 0.394	6,4	149/150	C,M,L,W,H,I,N,Y,F
80	P	PRO82:B	-1.209	9	-1.275,-1.162	9,9	149/150	R,P
81	I	ILE83:B	-0.738	8	-0.879,-0.681	8,7	150/150	C,V,I,L
82	G	GLY84:B	-1.241	9	-1.296,-1.236	9,9	150/150	G
83	L	LEU85:B	-0.394	6	-0.566,-0.286	7,6	149/150	I,V,A,F,Y,L
84	L	LEU86:B	-1.007	9	-1.125,-0.923	9,8	149/150	C,V,I,L,S
85	F	PHE87:B	0.071	5	-0.201, 0.241	6,4	149/150	Y,W,F,H,T,M,L
86	D	ASP88:B	-1.235	9	-1.296,-1.199	9,9	149/150	D,T
87	L	LEU89:B	-0.153	6	-0.364,-0.006	6,5	149/150	S,Y,F,C,M,Q,L,I
88	L	LEU90:B	1.398	1	0.808, 1.586	2,1	149/150	N,S,T,F,V,Y,C,L,Q,M,A,H,I
89	A	ALA91:B	0.876	2	0.394, 1.115	4,1	149/150	R,K,Q,M,W,I,A,G,S,T,V,N,D,H,E,C,P,F
90	S	SER92:B	0.915	2	0.577, 1.115	3,1	148/150	T,S,N,I,A,Q,M,L,K,R,G,Y,F,E,H,D,P
91	S	SER93:B	0.546	3	0.241, 0.808	4,2	149/150	N,V,S,T,P,G,H,E,K,L,R,Q,D
92	S	SER94:B	1.221	1	0.808, 1.586	2,1	149/150	N,T,S,V,G,M,R,L,K,I,A,F,P,D,E,H
93	A	ALA95:B	1.307	1	0.808, 1.586	2,1	149/150	G,I,A,L,K,Q,M,N,V,S,T,P,C,H,E,D,F
94	L	LEU96:B	-0.818	8	-0.966,-0.734	8,8	149/150	P,Q,D,L,I,H,T,S,V
95	P	PRO97:B	-1.112	9	-1.199,-1.048	9,9	149/150	T,L,I,P
96	W	TRP98:B	-0.750	8	-0.966,-0.625	8,7	149/150	I,V,W,L,M
97	N	ASN99:B	0.793	2	0.394, 1.115	4,1	149/150	V,Y,S,T,N,I,H,A,E,K,R,Q,P,G,C
98	I	ILE100:B	-0.558	7	-0.734,-0.436	8,7	149/150	L,V,I
99	T	THR101:B	-0.647	7	-0.784,-0.566	8,7	149/150	V,F,S,T,N,I,L,D
100	V	VAL102:B	-0.959	8	-1.087,-0.879	9,8	149/150	I,A,V,M,L
101	H	HIS103:B	-0.882	8	-1.007,-0.832	9,8	149/150	N,H,V,R,K,M,Q
102	F	PHE104:B	-0.897	8	-1.048,-0.832	9,8	149/150	S,T,I,V,F,Y
103	K	LYS105:B	0.259	4	-0.006, 0.394	5,4	149/150	S,T,N,H,E,R,K,Q,D,G
104	S	SER106:B	-0.158	6	-0.364,-0.006	6,5	150/150	D,Q,K,R,E,A,I,H,C,G,P,T,S,V,N
105	F	PHE107:B	-0.531	7	-0.734,-0.436	8,7	150/150	S,F,V,Y,G,P,R,L,A,W
106	P	PRO108:B	-1.209	9	-1.275,-1.162	9,9	150/150	S,P
107	E	GLU109:B	0.476	3	0.110, 0.577	5,3	150/150	V,T,S,N,E,A,I,Q,M,D,R,K,P
108	K	LYS110:B	1.107	1	0.577, 1.586	3,1	148/150	N,S,T,G,E,K,R,Q,D
109	D	ASP111:B	0.406	4	0.110, 0.577	5,3	134/150	G,E,I,A,M,D,Q,K,L,N,V,T,S
110	L	LEU112:B	-0.259	6	-0.503,-0.108	7,5	142/150	L,S,I,V
111	L	LEU113:B	0.045	5	-0.201, 0.241	6,4	148/150	M,L,Y,F,A,V,I
112	H	HIS114:B	0.060	5	-0.201, 0.241	6,4	149/150	S,A,I,H,Q,L,K,R,G,P
113	C	CYS115:B	0.160	4	-0.108, 0.394	5,4	148/150	Y,F,T,S,N,W,A,I,M,L,R,P,C
114	P	PRO116:B	2.172	1	1.115, 2.828	1,1	140/150	P,D,H,E,Y,G,K,R,L,Q,I,A,N,S,T,V

115	S	SER117:B	-0.269	6	-0.436,-0.201	7,6	150/150	N,S,T,C,G,K,R,D,A,E
116	K	LYS118:B	0.131	5	-0.108, 0.241	5,4	150/150	I,E,L,K,R,D,G,V,S,T,N
117	D	ASP119:B	0.346	4	-0.006, 0.577	5,3	150/150	E,A,Q,D,L,K,N,V,T,S
118	A	ALA120:B	0.310	4	-0.006, 0.394	5,4	150/150	Y,V,T,S,N,E,W,I,A,M,D,L
119	I	ILE121:B	-0.367	6	-0.566,-0.286	7,6	150/150	T,M,L,V,A,I
120	E	GLU122:B	-0.989	9	-1.087,-0.923	9,8	150/150	E,K,Q,D
121	A	ALA123:B	-0.353	6	-0.566,-0.286	7,6	150/150	G,M,D,Q,I,W,A,E,S,T,F
122	H	HIS124:B	-0.314	6	-0.503,-0.201	7,6	150/150	S,Y,F,N,Q,M,L,E,I,H,C
123	F	PHE125:B	-0.409	6	-0.625,-0.286	7,6	150/150	F,W,Y,L,M,N
124	M	MET126:B	-0.566	7	-0.734,-0.503	8,7	150/150	M,L,F,V,I,N
125	S	SER127:B	-0.605	7	-0.734,-0.503	8,7	150/150	N,C,V,H,A,Q,S
126	C	CYS128:B	0.334	4	-0.006, 0.577	5,3	150/150	N,V,S,T,G,C,A,K,M
127	M	MET129:B	0.241	4	-0.006, 0.394	5,4	150/150	V,I,M,L
128	K	LYS130:B	-1.254	9	-1.296,-1.236	9,9	150/150	K
129	E	GLU131:B	-1.159	9	-1.236,-1.125	9,9	150/150	Q,D,E
130	A	ALA132:B	-1.219	9	-1.275,-1.199	9,9	150/150	A,S,K
131	D	ASP133:B	-1.075	9	-1.162,-1.007	9,9	150/150	S,K,D,G,N
132	A	ALA134:B	0.152	4	-0.108, 0.241	5,4	150/150	N,Y,F,V,T,S,G,C,I,W,A,Q,M,L
133	L	LEU135:B	-0.825	8	-0.966,-0.734	8,8	150/150	L,T,M,F,V,I
134	K	LYS136:B	-1.046	9	-1.162,-1.007	9,9	150/150	K,R
135	H	HIS137:B	-0.970	8	-1.087,-0.923	9,8	150/150	N,Y,H,T,R
136	K	LYS138:B	-0.596	7	-0.784,-0.503	8,7	150/150	Q,D,R,K,V,G
137	S	SER139:B	0.186	4	-0.108, 0.394	5,4	150/150	K,D,M,A,I,E,C,G,S,T,N
138	Q	GLN140:B	-0.134	5	-0.364,-0.006	6,5	150/150	Y,V,T,N,E,A,I,H,M,D,Q,K,R,L,P
139	V	VAL141:B	-0.140	6	-0.364,-0.006	6,5	149/150	I,A,M,K,L,R,V,T
140	I	ILE142:B	-0.312	6	-0.503,-0.201	7,6	149/150	F,A,V,I,L,M,T
141	N	ASN143:B	-0.614	7	-0.784,-0.503	8,7	149/150	N,T,S,M,Q,K,L,E,H,A
142	E	GLU144:B	0.169	4	-0.108, 0.394	5,4	149/150	E,A,Q,D,L,K,G,C,V,T,S,N
143	M	MET145:B	-1.084	9	-1.162,-1.048	9,9	149/150	I,F,L,M
144	Q	GLN146:B	-0.778	8	-0.923,-0.681	8,7	149/150	V,T,S,G,M,Q,L,R,K
145	K	LYS147:B	-0.275	6	-0.503,-0.108	7,5	149/150	S,V,P,R,K,Q,M,I,A,E
146	K	LYS148:B	-0.511	7	-0.681,-0.436	7,7	149/150	G,Q,D,K,R,E,H,N,T,S
147	D	ASP149:B	-0.881	8	-1.007,-0.784	9,8	149/150	E,D,Q,G
148	H	HIS150:B	-0.772	8	-0.923,-0.681	8,7	149/150	Y,V,F,S,H,I,Q,R,L,C
149	K	LYS151:B	-0.227	6	-0.436,-0.108	7,5	149/150	N,S,T,G,K,R,D,H,A,I
150	Q	GLN152:B	-0.975	8	-1.087,-0.923	9,8	149/150	T,S,Q,M,L,K,E,A
151	L	LEU153:B	-1.153	9	-1.236,-1.125	9,9	149/150	L,M
152	W	TRP154:B	-0.799	8	-1.007,-0.681	9,7	148/150	F,W,L,T
153	M	MET155:B	0.343	4	-0.006, 0.577	5,3	148/150	N,V,T,S,E,A,I,H,D,M,Q,L,R

154	G	GLY156:B	-0.777	8	-0.923,-0.681	8,7	148/150	G,A,S
155	L	LEU157:B	-0.357	6	-0.566,-0.201	7,6	148/150	F,V,I,L,M
156	Q	GLN158:B	0.736	2	0.394, 0.808	4,2	148/150	C,E,A,H,W,I,Q,M,L,K,R,N,V,T,S
157	N	ASN159:B	-0.590	7	-0.734,-0.503	8,7	148/150	S,T,N,K,M,Q,D,H,E
158	D	ASP160:B	-0.631	7	-0.784,-0.503	8,7	148/150	S,V,F,Y,N,L,D,H,E,G
159	R	ARG161:B	-0.709	8	-0.879,-0.625	8,7	148/150	H,K,S,R,N
160	F	PHE162:B	-0.824	8	-0.966,-0.734	8,8	148/150	T,F,A,Y
161	D	ASP163:B	-0.669	7	-0.832,-0.566	8,7	148/150	E,D,L,K,N
162	Q	GLN164:B	-0.842	8	-0.966,-0.784	8,8	148/150	K,R,D,Q,Y,E,N
163	F	PHE165:B	-1.000	9	-1.125,-0.923	9,8	148/150	Y,F
164	W	TRP166:B	-1.102	9	-1.236,-1.048	9,9	148/150	F,W
165	A	ALA167:B	-0.442	7	-0.625,-0.364	7,6	148/150	E,H,I,A,Q,D,K,G,P,V,F,T,S
166	I	ILE168:B	-0.291	6	-0.503,-0.201	7,6	148/150	L,M,T,V,A,I,E
167	N	ASN169:B	-1.218	9	-1.275,-1.199	9,9	148/150	V,K,N
168	R	ARG170:B	-0.147	6	-0.364,-0.006	6,5	148/150	E,H,Q,D,R,L,K,G,Y,V,S,N
169	K	LYS171:B	-0.371	6	-0.566,-0.286	7,6	149/150	H,M,Q,L,K,R
170	L	LEU172:B	-1.010	9	-1.125,-0.923	9,8	149/150	L,R,I,F
171	M	MET173:B	-1.042	9	-1.125,-1.007	9,9	149/150	C,A,V,M,T,L
172	E	GLU174:B	-0.642	7	-0.832,-0.566	8,7	148/150	L,Q,D,A,E,P,G,S,T,F,N
173	Y	TYR175:B	1.228	1	0.808, 1.586	2,1	148/150	D,H,C,P,F,Y,L,R,Q,I,A,G,S,T,V,N
174	P	PRO176:B	0.696	3	0.241, 0.808	4,2	144/150	N,T,S,V,G,Q,R,L,A,I,Y,F,P,D,E,H
175	A	ALA177:B	1.561	1	0.808, 1.586	2,1	139/150	C,G,P,D,Q,A,E,N,S,T,V
176	E	GLU178:B	0.753	2	0.241, 1.115	4,1	76/150	G,K,D,H,I,E,N,S,T
177	E	GLU179:B	1.486	1	0.808, 1.586	2,1	149/150	S,T,Y,N,K,L,M,Q,D,A,I,E,C,P,G
178	N	ASN180:B	1.005	1	0.577, 1.115	3,1	149/150	K,R,L,D,Q,H,E,G,P,S,V,N
179	G	GLY181:B	1.377	1	0.808, 1.586	2,1	149/150	K,L,Q,A,G,S,T,V,N,D,H,E,C,P,F,Y
180	F	PHE182:B	-0.433	7	-0.625,-0.286	7,6	149/150	V,F,N,W,A,I,M,L,P
181	R	ARG183:B	-0.198	6	-0.436,-0.006	7,5	149/150	F,Y,C,H,A,I,L,R,K,M
182	Y	TYR184:B	0.310	4	-0.006, 0.577	5,3	150/150	C,D,A,H,N,S,F,Y
183	I	ILE185:B	-0.611	7	-0.784,-0.503	8,7	150/150	L,T,I,V
184	P	PRO186:B	-1.166	9	-1.275,-1.125	9,9	150/150	P,S
185	F	PHE187:B	0.083	5	-0.201, 0.241	6,4	149/150	C,Y,V,I,F,M,L
186	R	ARG188:B	-0.944	8	-1.048,-0.879	9,8	149/150	H,E,S,K,R
187	I	ILE189:B	-0.262	6	-0.503,-0.108	7,5	149/150	V,W,F,I,L,C
188	Y	TYR190:B	-0.666	7	-0.832,-0.566	8,7	149/150	G,V,F,H,Y
189	Q	GLN191:B	1.507	1	0.808, 1.586	2,1	150/150	C,H,E,D,F,Y,I,A,K,R,L,M,Q,N,V,S,T
190	T	THR192:B	2.735	1	1.586, 2.828	1,1	46/150	S,T,V,Q,M,D,A,P
191	T	THR193:B	2.530	1	1.586, 2.828	1,1	45/150	I,M,T,R,S,L
192	T	THR194:B	1.358	1	0.808, 1.586	2,1	146/150	V,T,S,N,A,I,Q,M,L,R,K,G,E,H,D,P

193	E	GLU195:B	0.074	5	-0.201, 0.241	6,4	149/150	H,E,K,D,Q,G,P,F,Y,S,N
194	R	ARG196:B	1.490	1	0.808, 1.586	2,1	149/150	H,E,D,C,Y,A,L,R,K,Q,M,G,V,S,T,N
195	P	PRO197:B	0.810	2	0.394, 1.115	4,1	146/150	A,H,I,D,R,L,K,G,P,C,Y,V,T,S,N
196	F	PHE198:B	0.935	2	0.394, 1.115	4,1	148/150	L,Q,M,I,H,C,P,T,V,F,Y
197	I	ILE199:B	0.709	2	0.394, 0.808	4,2	148/150	H,I,W,L,K,R,M,Q,C,V,Y,S,T,N
198	Q	GLN200:B	-1.107	9	-1.199,-1.048	9,9	148/150	E,L,M,T,Q
199	K	LYS201:B	0.705	2	0.241, 0.808	4,2	148/150	A,I,W,Q,M,R,K,G,V,T,S,N,E,D,P,C,Y,F
200	L	LEU202:B	-0.649	7	-0.832,-0.566	8,7	148/150	F,T,I,A,Q,L,P,C
201	F	PHE203:B	0.306	4	-0.006, 0.577	5,3	149/150	L,Q,M,W,H,I,C,T,V,F,Y
202	R	ARG204:B	0.072	5	-0.201, 0.241	6,4	149/150	T,S,N,A,Q,K,R,G,P,C
203	P	PRO205:B	-0.659	7	-0.832,-0.566	8,7	148/150	N,F,V,Y,S,T,P,A,I
204	V	VAL206:B	2.389	1	1.115, 2.828	1,1	146/150	C,H,F,Y,G,L,R,K,M,Q,I,W,N,S,T,V
205	A	ALA207:B	1.198	1	0.808, 1.586	2,1	146/150	L,K,R,D,H,A,E,G,S,T,V,N
206	A	ALA208:B	1.391	1	0.808, 1.586	2,1	142/150	G,P,Q,D,K,E,A,N,T,S,V
207	D	ASP209:B	1.320	1	0.808, 1.586	2,1	147/150	N,T,S,P,G,E,H,A,D,Q,K,R
208	G	GLY210:B	-0.426	7	-0.625,-0.286	7,6	146/150	V,S,N,E,D,K,G,C
209	Q	GLN211:B	2.406	1	1.115, 2.828	1,1	146/150	Q,D,L,R,K,E,H,A,C,G,T,S,Y,V,N
210	L	LEU212:B	2.161	1	1.115, 2.828	1,1	146/150	N,S,T,F,V,C,G,P,L,R,K,Q,A,H,E
211	H	HIS213:B	1.098	1	0.577, 1.586	3,1	147/150	H,E,F,Y,G,I,A,R,K,L,Q,M,N,V,S,T
212	T	THR214:B	-0.608	7	-0.734,-0.503	8,7	146/150	P,I,R,L,K,N,V,S,T
213	L	LEU215:B	-0.194	6	-0.436,-0.006	7,5	146/150	L,F,W,A,V,I,Y
214	G	GLY216:B	1.984	1	1.115, 2.828	1,1	145/150	N,S,F,Y,G,L,K,R,D,Q,A,I,H,E
215	D	ASP217:B	-0.161	6	-0.364,-0.006	6,5	145/150	C,E,H,A,D,Q,K,N,T,S
216	L	LEU218:B	-0.416	6	-0.625,-0.286	7,6	145/150	M,L,S,I,A,F,V
217	L	LEU219:B	0.156	4	-0.108, 0.394	5,4	145/150	C,I,L,M,N,V,F,T
218	K	LYS220:B	2.562	1	1.586, 2.828	1,1	145/150	N,T,S,V,G,M,D,Q,L,K,R,E,I,H,A
219	E	GLU221:B	0.660	3	0.241, 0.808	4,2	144/150	E,A,I,M,D,Q,L,R,N,V,T,S
220	V	VAL222:B	0.873	2	0.394, 1.115	4,1	144/150	Y,F,V,T,S,N,W,A,I,M,L,C
221	C	CYS223:B	1.519	1	0.808, 1.586	2,1	141/150	G,C,E,I,A,H,D,Q,L,N,Y,V,F,T,S
222	P	PRO224:B	0.216	4	-0.108, 0.394	5,4	142/150	G,Q,M,K,L,A,I,N,T,S,V,P,D,H,Y,F
223	S	SER225:B	2.725	1	1.586, 2.828	1,1	142/150	T,S,V,F,N,M,Q,D,K,R,L,E,A,H,G
224	A	ALA226:B	1.906	1	1.115, 2.828	1,1	133/150	T,S,F,V,N,M,D,Q,K,R,L,E,I,A,C
225	I	ILE227:B	2.794	1	1.586, 2.828	1,1	98/150	R,L,K,M,I,A,G,S,T,V,N,D,E,P,F,Y
226	D	ASP228:B	2.822	1	1.586, 2.828	1,1	137/150	N,S,T,V,G,R,K,L,Q,I,A,F,Y,C,P,D,H,E
227	K	LYS235:B	2.824	1	1.586, 2.828	1,1	137/150	T,S,V,N,Q,R,L,K,I,A,G,F,D,E,H,P
228	N	ASN236:B	2.823	1	1.586, 2.828	1,1	136/150	G,M,L,K,A,I,W,N,T,S,V,C,P,D,E,H,Y,F
229	Q	GLN237:B	1.992	1	1.115, 2.828	1,1	136/150	N,T,S,V,G,Q,M,R,L,K,A,I,Y,P,E,H
230	V	VAL238:B	0.068	5	-0.201, 0.241	6,4	136/150	L,I,A,C,P,Y,V,F
231	M	MET239:B	0.873	2	0.394, 1.115	4,1	136/150	I,H,K,L,R,Q,M,V,F,Y,T
232	I	ILE240:B	-0.662	7	-0.832,-0.566	8,7	136/150	C,T,S,L,I,V,A

233	H	HIS241:B	-0.896	8	-1.007,-0.832	9,8	137/150	Q,H,F
234	G	GLY242:B	-1.096	9	-1.199,-1.048	9,9	138/150	N,G,S,V
235	I	ILE243:B	-0.646	7	-0.784,-0.566	8,7	137/150	T,M,L,I,A,V
236	E	GLU244:B	1.673	1	1.115, 1.586	1,1	137/150	F,D,E,H,C,P,T,S,V,N,M,Q,R,L,K,I,A,G
237	P	PRO245:B	0.420	3	0.110, 0.577	5,3	137/150	L,T,A,I,V,P
238	M	MET246:B	-0.456	7	-0.681,-0.364	7,6	136/150	V,X,S,E,H,A,Q,M,L,P
239	L	LEU247:B	1.408	1	0.808, 1.586	2,1	136/150	F,D,E,H,P,X,T,S,V,N,M,Q,L,R,K,A,W
240	E	GLU248:B	0.612	3	0.241, 0.808	4,2	137/150	N,T,S,V,G,P,D,Q,K,E,H
241	T	THR249:B	-0.937	8	-1.048,-0.879	9,8	137/150	V,S,T,I,A,E,M,C
242	P	PRO250:B	-0.756	8	-0.923,-0.681	8,7	137/150	N,P,Q,D,S,H,V
243	L	LEU251:B	0.034	5	-0.201, 0.241	6,4	137/150	C,T,M,L,V,I
244	Q	GLN252:B	-0.248	6	-0.436,-0.108	7,5	137/150	G,C,H,I,A,R,L,Q,N,V,S,T
245	W	TRP253:B	-1.013	9	-1.162,-0.923	9,8	136/150	W,F,Y
246	L	LEU254:B	-0.585	7	-0.734,-0.503	8,7	136/150	L,M,A,V,I
247	S	SER255:B	-0.849	8	-0.966,-0.784	8,8	136/150	C,G,P,L,A,I,S,V,Y
248	E	GLU256:B	-0.563	7	-0.734,-0.436	8,7	136/150	T,V,Y,K,R,L,D,Q,I,E
249	H	HIS257:B	-1.005	9	-1.125,-0.966	9,8	136/150	N,R,S,Y,H,W
250	L	LEU258:B	0.167	4	-0.108, 0.394	5,4	136/150	C,F,Y,L,M
251	S	SER259:B	-1.079	9	-1.162,-1.048	9,9	135/150	A,S,T,C
252	Y	TYR260:B	-0.659	7	-0.832,-0.566	8,7	134/150	C,Y,H,F
253	P	PRO261:B	-0.826	8	-0.966,-0.734	8,8	133/150	X,L,A,P
254	D	ASP262:B	-1.257	9	-1.296,-1.236	9,9	133/150	D,X
255	N	ASN263:B	-1.217	9	-1.275,-1.199	9,9	134/150	N,S,H
256	F	PHE264:B	-1.238	9	-1.296,-1.236	9,9	134/150	F
257	L	LEU265:B	-1.021	9	-1.125,-0.966	9,8	134/150	I,V,L
258	H	HIS266:B	-1.113	9	-1.199,-1.087	9,9	134/150	Y,H,T
259	I	ILE267:B	-0.445	7	-0.625,-0.364	7,6	133/150	M,L,V,F,I
260	S	SER268:B	0.777	2	0.394, 1.115	4,1	132/150	S,T,V,F,A,I,C
261	I	ILE269:B	0.407	4	0.110, 0.577	5,3	130/150	V,A,I,L,M
262	I	ILE270:B	2.526	1	1.586, 2.828	1,1	102/150	N,T,V,F,C,P,R,L,M,Q,H,A,I
263	P	PRO271:B	0.611	3	0.110, 0.808	5,2	70/150	P,Q,M,K,H,A,T,S,Y
264	Q	GLN272:B	1.859	1	0.808, 2.828	2,1	64/150	P,E,A,Q,R,K,L,V,T,S
265	P	PRO273:B	0.197	4	-0.201, 0.394	6,4	53/150	D,K,A,G,P,T,S,V
266	T	THR274:B	1.980	1	0.808, 2.828	2,1	40/150	N,T,S,P,E,A,D,R,K
267	D	ASP275:B	0.083	5	-0.436, 0.394	7,4	31/150	G,N,D,F,V

*Below the confidence cut-off - The calculations for this site were performed on less than 6 non-gaped homologue sequences, or the confidence interval for the estimated score is equal to- or larger than- 4 color grades.

APPENDIX-B

8.3. Buffers and reagents

a) Buffers and reagents used in DNA isolation

- **Tris (hydroxymethyl) aminomethane-cholride (Tris-Cl)**

Ingredients	**Amount (g)**	**Final concentration**
Tris base (pH 8.0)	12.11	1 M
pH was set to 8.0 by using 1N HCl Added Mili Q (MQ) water to make up the final volume to 100 ml. Filtered the solution using Whatmann filter paper Stored at room temperature (RT)		

- **Ammonium Chloride (NH_4Cl)**

Ingredients	**Amount (g)**	**Final concentration**
Ammonium Chloride (NH_4Cl)	5.35	1 M
Dissolved in 80 ml of MQ water and final volume was made up to 100 ml. Stored at RT		

- **Di-sodium ethylene diamine tetra acetate (Na_2EDTA)**

Ingredients	**Amount (g)**	**Final concentration**
EDTA	18.61	0.5 M
Dissolved in 50 ml of MQ water and placed over magnetic stirrer. 10 M NaOH was added to the above mixture Allowed the salt to dissolve completely. Final volume was made up to 100 ml with MQ water drop-wise to set the pH at 8.0. Stored at RT		

- **Red blood cell (RBC) lysis buffer**

Ingredients	**Concentration**	**Volume (ml)**
Tris (pH 8.0)	1 M	10
Ethylene diamine tetra acetic acid (EDTA)	0.5 M	2
Ammonium chloride (NH_4Cl) (pH 8.0)	1 M	125
Added Mili Q water to make up the final volume to 1000 ml Stored at RT.		

- **Tris-EDTA (TE) buffer (pH 8.0)**

Ingredients	**Concentration**	**Volume (ml)**
Tris Cl	1 M	10
Ethylene diamine tetra acetic acid (EDTA) (pH 8.0)	0.5 M	2
Mixed completely in 700 ml of MQ water initially. Then final volume was made to 1000 ml by adding MQ water. Stored at RT.		

- **Tris-EDTA (TE) buffer (pH 7.3)**

Ingredients	**Concentration**	**Volume (ml)**
Tris Cl (pH 7.3)	1 M	10
Ethylene diamine tetra acetic acid (EDTA)	0.5 M	2
Mixed completely in 700 ml of MQ water initially. Then final volume was made to 1000 ml by adding MQ water. Stored at RT.		

- **Sodium dodecyl Sulphate (SDS) (10%)**

Ingredients	**Amount (g)**	**Final concentration**
Sodium dodecyl Sulphate (SDS)	10	10 %
Dissolved in 70 ml of MQ water Allowed the salt to dissolve completely Final volume was made up to 100 ml with MQ water Stored at RT		

- **Ammonium acetate (7.5 M)**

Ingredients	**Amount (g)**	**Final concentration**
Ammonium acetate	28.9	7.5 M
Dissolved in 20 ml of MQ water Allowed the salt to dissolve completely Final volume was made up to 50 ml with MQ water Stored at RT		

- **Dehydrated Ethyl alcohol (chilled)**

  Undiluted dehydrated ethyl alcohol is kept at -20°C deep freezer till use.

- **Ethyl alcohol (70%)**

Ingredients	**Volume (ml)**	**Final concentration**
Dehydrated ethyl alcohol	70	70%
Mixed completely in 30 ml of MQ water to obtain the final volume of 100 ml. Stored at RT.		

b) <u>Buffers used for Agarose gel electrophoresis</u>

- **Tris-acetic acid-EDTA (TAE) buffer (50x)**

Ingredients	**Amount (g)**	**Final concentration**
Tris base	242	50x
Dissolved completely in 500 ml of MQ water Added 57.1 ml of glacial acetic acid Added 100 ml of 0.5 M EDTA (pH 8.0) Final volume was made up to 100 0 ml with MQ water Filtered and stored at RT		

- **Agarose gel (0.8%, 2%, 2.5%)**

Ingredients	**Amount (g)**	**Final concentration**
Agarose	0.8	0.8%
Added 100 ml of 1x TAE buffer Dissolved completely by heating in microwave oven		

Ingredients	**Amount (g)**	**Final concentration**
Agarose	2	2%
Added 100 ml of 1x TAE buffer Dissolved completely by heating in microwave oven		

Ingredients	**Amount (g)**	**Final concentration**
Agarose	2.5	2.5%
Added 100 ml of 1x TAE buffer Dissolved completely by heating in microwave oven		

- **Ethidium bromide (10 mg/ml)**

Ingredients	**Amount (mg)**	**Final concentration**
Ethidium bromide	50	10 mg/ml
Added 5 ml of MQ water Dissolved completely by heating in microwave oven		

c) <u>Buffers used for Agarose gel loading</u>

- **Gel loading dye (6x)**

  0.25 % Bromophenol Blue (BPB) with 40% Sucrose dissolved in water for genomic DNA analysis.

- **Gel loading dye (10x)**

0.5 % (w/v) xylene cyanol, dissolved in distilled water mixed with equal volume of glycerol. Used for PCR product analysis.

List of publications

Avni Vij, Rohit Randhawa, Jyoti Parkash, Harish Changotra (2016). Investigating Regulatory Signatures of Human Autophagy-Related Gene 5 (ATG5) through Functional in Silico Analysis. Meta Gene, 9, 237-248**.**

Avni Vij, Raghu M. Yennamalli, Harish Changotra (2017). Non-Synonymous Single Nucleotide Polymorphisms of ATG5 Destabilize ATG12 - ATG5 / ATG16L1 Complex: an enzyme with E3 like activity of ubiquitin conjugation system. Meta Gene, 13 (2017), 38-47.

Harish Changotra and **Avni Vij**. Rotavirus virus-like particles (RV-VLPs) vaccines: An update. Rev Med Virol. 2017;e1954 [I.F = 5.4]

Conference poster presentations

Avni Vij and Harish Changotra, "In silico prediction and analysis of non synonymous single nucleotide polymorphisms (nsSNPs) in Autophagy-related gene 5 (ATG5)"at National Conference on Recent Trends in Biomedical Engineering, Cancer Biology, Bioinformatics and Applied Biotechnology (BECBAB-2015) in November, 2015 at Jawaharlal Nehru University, New Delhi.

Avni Vij and Harish Changotra, "Role of Autophagy-related gene 5 (ATG5) single nucleotide polymorphism rs2245214 (C/G) with HBV susceptibility in North Indian population", at International Conference on Recent Research in Biomedical Engineering, Cancer Biology, Stem Cells, Bioinformatics and Applied Biotechnology (BECBAB-2017) in February, 2017 at Jawaharlal Nehru University, New Delhi.