Cytochrome P450 (CYP) is a multigene family of constitutive and inducible enzymes which catalyze oxidative metabolism of a wide variety of xenobiotics, of both endogenous and exogenous origin. Though several distinct CYP gene families are expressed in mammals, four of these CYP gene families (CYP 1-4) are recognized to be the major forms of CYPs that metabolize foreign chemicals including drugs, environmental chemicals and other xenobiotics. Alteration in the expression of these CYPs has been often found to be associated with exposure of drugs and environmental chemicals. Specific and selective expression profiles of these CYPs are observed following treatment of specific drugs and environmental chemicals. Genetic polymorphism is reported for several of the xenobiotic metabolizing CYPs and the susceptible genotypes are reported to be at increased risk to the toxicity of drugs and chemicals. Genetic variations in the CYPs are often an emerging resource for studying various complex diseases and have been identified as biomarker of susceptibility. Much of the variability in the CYPs is in the form of stable substitution of a single base termed as the single nucleotide polymorphism (SNP). Although SNPs in CYPs have been found to be present in all the population, the frequencies of these polymorphisms differ between populations. In general, SNPs occurring in the coding region have the potential to alter protein structure and function and therefore could be the causative factor influencing disease susceptibility, drug efficacy and xenobiotic response. SNPs in CYPs can cause abolished, quantitatively altered or enhanced xenobiotic metabolism leading to enhanced or diminished toxicity of a chemical or adverse drug effects. In addition, recent studies
have shown that CYPs are expressed in blood lymphocytes and could be used as a phenotypic tool and biomarker to predict exposure of drugs and environmental changes.

Cytochrome P450 2E1 (CYP2E1) is a major component of the microsomal ethanol oxidizing system in the liver. It exhibits a high degree of inter-individual variability in its catalytic activity. To date, several mutations have been described in CYP2E1 gene which are well characterized in the Caucasian as well as in Oriental population. Further, these SNPs in CYP2E1 play a major role in determining individual's susceptibility to alcoholic liver diseases such as alcoholic liver cirrhosis, a chronic liver disease results from damage to liver cells from heavy alcohol consumption. As compared to the data available in other population worldwide, not much is known about the SNPs in CYP2E1 in Indian population. Only few case-control studies are available while no population based study have been carried out to investigate the genotype status of CYP2E1 gene in Indian populations. Studies have shown that Indian population has more human diversity than any other comparable global population in the world. It is likely that the SNPs found in the genes involved in the ethanol metabolism pathway may be conserved and these could be exploited to identify the risk factors for the environment induced diseases. Further the CYP2E1 protein activity and mRNA expression is also known to increase several fold in the liver of experimental animals following alcohol administration suggesting that CYP2E1 is regulated at various levels such as pretranscriptional, pretranslational, translational, and posttranslational. This increase in CYP2E1 activity results in accumulation of toxic acetaldehyde and reactive oxygen species in the cell which may lead to mitochondrial damage, DNA modification and lipid peroxidation. Studies have also shown that freshly isolated peripheral blood lymphocytes of human and laboratory animal express measurable levels of CYP2E1 mRNA and protein. Expression of CYP2E1 mRNA and protein in the peripheral blood lymphocytes is known to be influenced by the same factors that affect the concentration of hepatic enzymes including exposure to xenobiotics and certain physiological states particularly those associated with higher plasma cholesterol. Since expression of CYP2E1 is altered by many factors affecting chemical metabolism and toxicity, monitoring of CYP2E1 levels in individuals, who are regular alcohol users could help in identifying individuals who are at risk to develop alcohol induced toxicity.
The candidate in the present dissertation has therefore attempted to identify the SNPs, including novel ones in CYP2E1 gene in the different sub-population of India. Attempts were also made by the candidate in a case-control study to investigate association of functionally important SNPs in CYP2E1 gene identified in Indian population along with the other candidate genes such as glutathione S-transferases (GSTs) and manganese superoxide dismutase (MnSOD), responsible for detoxifying electrophilic intermediates including free radicals generated during the ethanol metabolism. Studies were also initiated by the candidate to develop peripheral blood lymphocyte CYP2E1 expression as a possible biomarker of alcoholic liver cirrhosis by investigating mRNA and protein expression of CYP2E1 and associated catalytic activity in freshly prepared blood lymphocytes isolated from alcoholic cirrhotic patients.

To identify SNPs in Indian population, different population were identified based on diverse geographical zones (North, South, East, West, Central and North-East) and linguistic backgrounds (Indo Europeans-IE, Dravidians-DR, Tibeto Burmans-TB and Austro Asiatics-AA) to maximize novel SNP discovery. Molecular diversity studies have revealed that Austro-Asiatic (AA) language speakers are exclusively tribals and are likely to have been the most ancient inhabitants of India. Populations of Indo-European (IE) and Dravidian (DR) lineage groups account for almost 85% of the total Indian population, while the Tibeto-Burman (TB) lineage groups are supposedly immigrants to India from Tibet and Burma. These diverse set of sub-population does increase the heterogeneity in terms of SNP discovery as compared to a set of samples from a single sub-population. These population were identified after consultation with the expert anthropologists. Blood samples were collected from different sub-population of India by the Indian Genome Variation (IGV) Consortium members including IITR, Lucknow. Genomic DNA samples were provided by IGV Consortium for identifying SNPs in CYP2E1. Five pairs of primers for CYP2E1 gene were designed with the aim to cover the exonic region as well as the regions carrying functionally important SNPs using the DNA Star software.

The first step involved discovering SNPs in CYP2E1 gene in one representative DNA sample drawn from 43 different sub-populations and designated as SNP Discovery Panel. Though 43 individuals per se do not represent the entire Indian population, such a diverse set does increase the heterogeneity in terms of SNP discovery as compared to a set of samples from a single sub-population. The SNPs identified in these Discovery
panel samples were than validated in the second step in a larger sample size of these populations. SNPs in CYP2E1 gene in SNP Discovery Panel were identified by bidirectional sequencing of PCR products using primer pairs that covered functionally important polymorphism of CYP2E1 gene. Five polymorphic sites including a novel one were identified in CYP2E1 gene and were designated as SNP1 (-930A/G; rs3813870) in 5'UTR, SNP2 (6396G/A; Novel) in exon 6, SNP3 (7126G/A; rs8192775), SNP4 (7632T/A; rs6413432) and SNP5 (8057C/A; rs2864984) in intron 6. Out of these five SNPs, SNP2 (6396G/A) was considered as novel because no information or any literature was available in any existing databases about this SNP. SNP1 was found to be polymorphic in all the four major sub-population (AA DR, IE and TB) with minor allele frequency (MAF) of 0.26. SNP2 was found to be novel and polymorphic in only one sub-population (DR) with very low minor allele frequency (0.01). SNP3 which is present in intron-6 was found to be polymorphic in the three sub-population (AA, DR and IE) with the minor allele frequency of 0.05. Similarly, SNP4 commonly known as Dral polymorphism or CYP2E1*6, was found to be polymorphic only in IE population with minor allele frequency of 0.01. Likewise, SNP5 also located at intron-6 was found to be polymorphic in three sub population (AA, DR and IE) with minor allele frequency of 0.05. SNPs identified as SNP1-5 at -930A/G (5'UTR), 6396G/A (exon 6), 7126G/A (intron 6), 7632T/A (intron 6) and 8057C/A (intron 6) in CYP2E1 gene were considered for further validation in a larger sample size because of their higher minor allele frequency. Interestingly, one of the major functionally important SNP, designated as Rsal or CYP2E1*5B was not present in any of the 43 representative DNA sample. Since the frequency of this polymorphism is very rare in Indian population (1%), it may be possible that the 43 representative DNA sample from different sub-population of India may not be carrying this SNP. It may be possible that if larger samples were screened, this SNP could have been identified.

Validation data of SNPs in CYP2E1 gene in different Indian sub-population further provided evidence that CYP2E1 gene is polymorphic in all the Indian sub-population. The five SNPs discovered in Discovery Panel were found to exist in Indian population with varied frequency. SNP1 was found to be present in all the linguistically and geographically diverse population of India. As the SNP1 was present in most of the population of India, it can be concluded that this SNP was conserved in almost all the
population of India and must be the oldest SNP present in *CYP2E1* gene in Indian population. This SNP has also been shown to be present in other world population though there is no frequency data available in NCBI or HapMap databases. Similarly, SNP2 which is novel and results in amino acid substitution from Glycine to Serine was also found to be polymorphic in AA, DR and IE lineages. Minor allele frequency of this SNP was found to be similar in AA, DR and IE population. As this SNP is novel, no allele or genotype frequency data is available in the HapMap or NCBI database. However, as this SNP is located at coding region and resulted in amino acid substitution from Glycine to Serine, it might be responsible for alteration in protein activity and need to be further studied. SNP3 was also found to be polymorphic in all sub-population of India belonging to different linguistic lineages. This polymorphism was found to be reported in Caucasians with comparable minor allele frequency with IE lineage. Interestingly, the MAF reported in Oriental population (Chinese and Japanese) was almost similar to Dravidian but found to be higher than other sub population of India. However, the minor allele frequency of this polymorphism in Sub-Saharan African is lower than minor allele frequency observed in all the sub-population of India. Likewise, our validation data have shown that SNP4 (Dral polymorphism) of *CYP2E1* was found to be polymorphic in most of the sub-population of India except for few IE and one TB lineages. The MAF of Dral polymorphism observed in various sub-population of India was quite comparable to Caucasian population. However, this SNP was found to be present in African, Hispanic and Pacific Rim with higher minor allele frequency than found in different sub-population of India. Similarly, SNP5 was found to be polymorphic in all sub-population of India with varied frequency of minor allele. This polymorphism has also been reported in Caucasian, African American and Chinese population. Minor allele frequency reported in Caucasian and African American was found to be lower than found in different sub population of India whereas its frequency in Chinese population is similar to the found in AA and Out-group population but higher than IE and TB population. Haplotype analysis revealed the formation of nine major haplotypes. Some of these haplotypes were found to be present in all the four major linguistic groups and in one Out-group population though the frequency was found to vary between these groups. Haplotype analysis of SNPs of *CYP2E1* also revealed that most of the SNPs did not exhibit Linkage Disequilibrium (LD) in majority of Indian population. However few SNPs (SNP3 & SNP5; SNP1 &
SNP3; SNP3 & SNP4) exhibit LD in few sub-population of India. The differences in the frequencies of these haplotypes could be of significance in determining the individual’s susceptibility towards the environment induced diseases.

To further identify the involvement of these SNPs with alcoholic liver disease, a case-control study was carried out to investigate the association of functionally important SNPs in CYP2E1 gene and other candidate genes with alcoholic liver cirrhosis in North Indian population. Even though Rsal (CYP2E1*5B) and Mspl polymorphism of CYP2E1 gene were not identified in SNP Discovery Panel of 43 samples, they were also included in the study as they were reported to be functionally important in Caucasians and Oriental population. Due to their very rare frequency in Indian population, these SNPs were not detected in 43 representative samples of Discovery Panel. However, because of their clinical significance, they were included in the case-control study to investigate their association with the disease. Along with CYP2E1, polymorphism in glutathione S-transferases (GSTs) and manganese superoxide dismutase (MnSOD), involved in scavenging free radicals, were also studied for their association with alcoholic liver cirrhosis. The study group consisted of 175 patients suffering from alcoholic liver cirrhosis and 140 number of patients suffering from non-alcoholic liver cirrhosis visiting the OPD facility of Gastroenterology Department of Sanjay Gandhi Post Graduate Institute of Medical Sciences (SGPGIMS), Lucknow, India was included in the study. Patients with alcoholic and non-alcoholic liver cirrhosis were diagnosed on the basis of their liver biopsy. Control group consisted of non-alcoholic (n=255) and alcoholic (n=140) healthy men having no evidence of liver disease as judged by physical examination and normal liver function test. All the patients and controls included in the study belonged to same geographical location (Northern India) and same ethnicity.

It was found that the frequency of minor allele (1%) of CYP2E1*5B (Rsal) was almost similar to that reported earlier in North Indian population but slightly higher than South Indian controls. Similar frequency of the minor allele of CYP2E1*5B was also reported in the Caucasians population while the Oriental population carry much higher frequency of the minor allele. Similarly, the minor allele frequency (16%) of CYP2E1*6 (Dral) was quite similar to that reported earlier in the North Indian or South Indian population. This polymorphism is also reported in the Caucasians, though its frequency was relatively less when compared to our study or Oriental population. Our data further
revealed that MspI polymorphism is present in North Indian population, though its minor allele frequency (6%) is quite higher than reported in Caucasians population. Similar to the CYP2E1, the frequency of null genotype of GSTM1 (31.5%) was found to be similar to that reported earlier in the North Indian population. However, this was much lower than found in other Asian populations, Caucasians and Pacific-Rims. Likewise, the GSTT1 null genotype frequency was found to be 17% in the controls which was similar to that seen in other North Indian, Caucasians and African-Americans but lower than reported in Orientals and Pacific Rim. Similarly the frequency of minor alleles of GSTP1 was found to be 22% which was equal to that observed with other North Indian and Caucasians but more than that reported with Japanese and African-Americans and less than Hispanics and Pacific-Rim population. Likewise, the frequency of Alanine encoding allele of MnSOD in our study was found to be 43% which is slightly lower than that reported in Caucasian but much higher than that found in Orientals.

It was also found that Rsal polymorphism (CYP2E1*5B) of CYP2E1 gene were observed with higher frequency in alcoholic liver cirrhotic patients when compared with non-alcoholic controls or alcoholic controls or non-alcoholic cirrhotic patients. The higher frequency of Rsal polymorphism of CYP2E1 was found to be significantly associated with alcoholic liver cirrhosis risk. Interestingly, Dral and Mspl polymorphism of CYP2E1 did not show statistically significant risk for alcoholic liver cirrhosis. Haplotype analysis has revealed that Rsal and Dral polymorphism of CYP2E1 do not exhibit LD in either controls or cases. Haplotype approach further revealed that haplotype T-A-T was found to be associated with more than 5-fold increase in risk for alcoholic liver cirrhosis. As observed with CYP2E1*5B, the present study also demonstrated that polymorphism in GSTs modifies the susceptibility to alcoholic liver cirrhosis. An elevated risk to alcoholic liver cirrhosis was observed in the patients with null genotype of GSTM1 when compared with the non-alcoholic controls. Similar significant risk was also observed in alcoholic cirrhotic patients carrying variant genotype (Ile/Val + Val/Val) of GSTP1 when compared with the non-alcoholic controls. Interestingly, no significant increase in the frequency of null genotype in GSTT1 was observed in alcoholic cirrhotic patients when compared with the non-alcoholic controls or non-alcoholic cirrhotic patients or alcoholic controls. As observed with GSTs, a statistically significant risk was observed in alcoholic cirrhotic patients with variant genotype of MnSOD, an enzyme
specifically involved in the detoxification of mitochondrial reactive oxygen species such as superoxide anion, when compared with non-alcoholic controls.

It was also found that combination of null genotypes of \textit{GSTM1} and \textit{GSTT1} or variant genotype of \textit{GSTP1} with null genotype of \textit{GSTM1} or \textit{GSTT1} confer an even higher risk to alcoholic liver cirrhosis. Likewise, a much greater risk for alcoholic cirrhosis was also detected in the patients carrying the combination of null genotype of \textit{GSTM1}, \textit{GSTT1} and variant genotype of \textit{GSTP1}. Our data also demonstrated that combination of SNPs particularly those which are involved in generating free radicals such as CYP2E1, a phase I enzyme, and GSTs or MnSOD, which detoxify there free radicals increase the risk in alcoholic cirrhotic patients. Several fold increase in risk was observed in the alcoholic liver cirrhosis patients with the combination of variant genotypes of \textit{CYP2E1*5B} and \textit{GSTP1} or null genotype of \textit{GSTM1} or \textit{GSTT1}. Similarly, several fold increase in the risk was observed in patients with alcoholic liver cirrhosis carrying combination of variant genotypes of \textit{CYP2E1*5B} and \textit{MnSOD} when compared with non-alcoholic controls or alcoholic controls or non-alcoholic cirrhotic patients. Further evidence for the importance of genotypic combinations in the development of alcoholic liver cirrhosis was provided by the present study indicating significantly increased risk in the cases simultaneously carrying the null or variant genotypes of \textit{GSTs} and \textit{MnSOD} when compared to the patients with the risk genotypes of \textit{GSTs} or \textit{MnSOD} alone.

In addition to identify the association of susceptibility genotypes of \textit{CYP2E1} with alcoholic liver cirrhosis, attempts were also made to develop and validate \textit{CYP2E1} as a phenotypic marker which can eventually be used in monitoring studies to screen individuals for alcohol exposure. Since exposure to alcohol has been shown to result in increased hepatic and blood lymphocyte \textit{CYP2E1} levels in laboratory animals, these studies have been extended to humans who are regular alcohol users and suffering from alcoholic liver cirrhosis with the aim to validate blood lymphocytes \textit{CYP2E1} levels as a possible biomarker of exposure. Study was therefore initiated to investigate mRNA and protein expression of \textit{CYP2E1} and associated catalytic activity in freshly prepared blood lymphocytes isolated from non-alcoholic healthy controls, non-alcoholic cirrhotic and alcoholic cirrhotic patients. Eight patients suffering from alcoholic liver cirrhosis (non-cholestatic) and equal number of patients suffering from non-alcoholic liver cirrhosis,
Summary

diagnosed on the basis of their liver biopsy and hepatitis B surface antigen and antibodies against hepatitis C virus, visiting the OPD facility of Gastroenterology Department of Sanjay Gandhi Post Graduate Institute of Medical Sciences (SGPGIMS), Lucknow, India was included in the study. Eight healthy non-alcoholic individuals were also enrolled in the study and served as controls. Semi-quantitative, quantitative RT-PCR, ELISA and NDMA-d activity were used for studying mRNA and protein expression of CYP2E1 and associated catalytic activity in freshly prepared blood lymphocytes of non-alcoholic healthy controls, non-alcoholic cirrhotic and alcoholic cirrhotic patients. Significant increase in the CYP2E1 mRNA and protein expression was observed in the freshly prepared blood lymphocytes isolated from non-cholestatic alcoholic liver cirrhotic patients when compared to non-alcoholic controls and non-alcoholic cirrhotic patients. This increase in blood lymphocyte CYP2E1 expression was associated with an increase in NDMA-d activity in blood lymphocytes isolated from alcoholic liver cirrhotic patients when compared to non-alcoholic controls and non-alcoholic cirrhotic patients. Our data indicating increase in CYP2E1 expression and catalytic activity in freshly prepared blood lymphocytes of alcoholic liver cirrhotic patients with early stage of alcoholic cirrhosis have suggested suitability of using blood CYP2E1 mRNA expression profile as a biomarker to detect the hepatic level of CYP2E1.

In conclusion, the present dissertation had demonstrated that \textit{CYP2E1} is highly polymorphic in Indian population. Indian population when stratified either on language basis or on the basis of geographical zones, were found to exhibit polymorphism in \textit{CYP2E1}. The presence of SNPs in \textit{CYP2E1} even in the tribal population, have further provided evidence that these SNPs in \textit{CYP2E1} were conserved during evolution and may have a physiological role in addition to metabolizing drugs and environmental chemicals. Case-control studies have further shown that polymorphism in \textit{CYP2E1} particularly \textit{CYP2E1*5B} and the phase II genes such as \textit{GSTs}, \textit{MnSOD}, are associated with an increased susceptibility to alcoholic cirrhosis. Risk to alcoholic cirrhosis was further found to increase several fold in the alcoholic cirrhotic cases carrying combinations of risk genotype of \textit{GSTs} or \textit{CYP2E1*5B} or \textit{MnSOD} demonstrating the role of gene-gene interactions in modulating the risk to alcoholic liver cirrhosis. The candidate in the present dissertation has further showed a significant increase in the CYP2E1 mRNA and protein expression and associated NDMA-d activity in the freshly prepared blood
lymphocytes isolated from non-cholestatic alcoholic liver cirrhotic patients indicating suitability of using blood CYP2E1 mRNA expression profile as a biomarker to predict the hepatic level of CYP2E1.