SUMMARY

Diabetic Retinopathy (DR) is the major cause of blindness worldwide (Fong et al., 2004). Uncontrolled hyperglycemia along with dislipidemia, hypertension and other lifestyle risk factors are responsible for the DR development (Frank, 2004). Several biochemical pathways have been reported which play a cumulative role in the DR progression such as aldose reductase pathway, protein kinase C, increased formation of advanced glycation end products due to oxidative stress and upregulation of endoplasmic stress (Pusparajah et al., 2016) and inflammatory pathway (Tang et al., 2013). With rapid industrialisation, the tendency towards physical inactivity, leading to overweight and obesity are on the rise among the populations of Punjab in Northern India as well as in other ethnicities worldwide. In North Indian population, the increased prevalence of diabetes and its related complications have been observed due to the transition from agriculture to a sedentary lifestyle associated with the consumption of high calorie and fat-rich diet (Gutch et al., 2014). DR being a complex disease, the mutual effects of gene-gene and/or gene-environment interactions play a key role in the disease pathogenesis. Thus, it is important to characterise the genetic and environmental factors that influence the risk of DR across different populations. Unfortunately, there is a paucity of genetic data in the literature from the North Indian populations that are largely restricted to the analysis of few single nucleotide polymorphisms (SNPs) in relatively smaller cohorts. Therefore, the present study was designed to investigate the various clinical and genetic determinants associated with DR in a large cohort from Punjab and to undertake genotype-phenotype correlation in order to understand their involvement in DR susceptibility.

The study design and the corresponding protocol was approved by the Ethics Committee of Guru Nanak Dev University, Amritsar. A detailed questionnaire was developed for sampling, including the demographic and clinical parameters and the potential risk factors for DM and DR, which were based on significant inputs from the literature. The present study comprised of 414 DR cases, who were enrolled based on a stringent inclusion criteria. The diagnosis was based on a comprehensive ophthalmological examination, including fundus imaging based on three 45° field tests.
per eye done by an ophthalmologist at Amritsar, Punjab. This was independently confirmed by another ophthalmologist at the LV Prasad Eye Institute, Hyderabad (who was masked to the initial diagnosis), based on the available clinical details and the fundus images of the patients. Gender and ethnicity matched healthy individuals (>50 years of age) from Punjab, without any signs or symptoms of diabetes or DR or any other systemic conditions, were enrolled as controls. A written informed consent was obtained from each subject prior to their sample collection, as per Indian Council of Medical Research (ICMR) guidelines.

DNA was extracted using standard protocols (inorganic method) and quantified by Nanodrop and further quality checked by agarose gel electrophoresis. Plasma was separated from the blood for determining the lipid profile of the subjects.

The overall genetic screening was based on 126 SNPs across 57 candidate genes and intergenic regions that were selected based on their role in the aetiology of DR, T2D and other metabolic disorders. Further information pertaining to the population frequencies of these SNPs, their functional implications and potential involvement in other etiologies were obtained from the literature and the public databases. Genotyping was accomplished with the Sequenom iPLEX Gold for MassARRAY that works on the principle of allele specific products with distinct masses. In the iPLEX extension PCR assay, the locus-specific extension primer and amplified target DNA were incubated with mass-modified dideoxynucleotide terminators. The primer extension was made according to the sequence of the variant site and the mass of the extended primer was determined with the use of Matrix Associated Laser Desorption Ionization (MALDI-TOF) spectrometry. The Sequenom SpectroTYPER software translated the mass of the observed primers into a genotype for each reaction. The quality of each genotype was assessed using a clustering algorithm.

Statistical analyses were performed using the SPSS software (version 20.0). Posthoc power of the study was calculated using the Quanto power calculator. The continuous variables were represented as means±standard deviations (SD) and were compared between the groups using Student’s t-test. Univariate and multivariate logistic regression analyses were performed to elucidate the risk factors associated with DR. A p-value of less than 0.05 was considered statistically significant.
analyses, Hardy-Weinberg Equilibrium (HWE) and Minor allele frequency (MAF), Linkage disequilibrium and haplotype analysis were done with the help of Haploview software (version 4.2) that uses the EM algorithm. The p-values for allele and haplotype frequencies were corrected with the 10,000 permutation test. The associations of different genetic models were examined using the $\chi^2$ test along with Odds ratios (OR) and 95% confidence intervals (CI). The test of significance was further evaluated by Bonferroni correction [i.e. p-value/no. of SNPs (0.05/126=0.0004)]. The present data were compared with the HapMap and 1000 genome consortium databases and other studies in the literature that included the distributions of allele frequencies among various populations at these 126 loci using Fisher exact test.

Clinical Analysis: The mean duration of T2D and DR were 155.85±89.32 months and 17.27±28.31 months, respectively. Univariate analysis of the baseline characteristics indicated significant association for age (p<0.001); BMI (p=0.002); WC (p=<0.001); WHR (p<0.010); WHtR (p<0.001); triglycerides (p=0.005); SBP (p<0.001) and DBP (p<0.001). Multivariate analysis revealed some independent risk factors like the history of hypertension [p=<0.001; OR=13.64 (9.15-20.34)], male gender [p=0.003; OR=1.86 (1.22-2.81)], increasing age [p<0.001; OR=1.08 (1.05-1.11) and BMI [p<0.001; OR=1.31 (1.17-1.46)] to be associated with the risk of DR, while the WHtR was protective [p<0.001; OR=0.189 (0.09-0.41)]. There were no associations to HC, cholesterol and HDL with DR.

Genetic Analysis: Based on quality control assessments, 3 SNPs with call rates <95% and 10 SNPs that deviated from HWE (p<0.05) were excluded from further analysis. The PAI-1 rs1799889 and TGF-β1 rs1800820 were observed to be monomorphic in this cohort.

The strongest associations were observed with the TCF7L2 rs12255372 [p<0.0001; OR=1.81 (1.44-2.27)] and rs11196205 [p=0.0008; OR=1.62 (1.32-1.99)] SNPs that withstood Bonferroni correction. The rs12255372 genotypes provided 1.85 fold and 3 fold risks toward DR under the dominant and recessive models, respectively. Likewise, the risks for the rs11196205 genotypes ranged between 1.5-2.2 fold risk for the dominant and recessive genotypes, respectively. The haplotype analysis of the TCF7L2
SNPs indicated that the C-T-T [p=0.001; OR=1.59 (1.21-2.07)] and C-T-G [p=0.004; OR=1.68 (1.2-2.3)] haplotypes conferred significant risk of DR, while the G-G-T [p=0.047; OR=0.76 (0.61-0.95) and G-G-G [p=0.005; OR=0.69 (0.54-0.88) haplotypes were protective.

The MAF for AKR1B1 rs9640883; PTPN-1 rs3787345 and rs754118; HFE rs1800562 and intergenic variant rs4516615 were significantly different amongst cases and controls after the 10,000 permutation test. However, the allele frequency distribution for ADIPOQ (rs266729 and rs17366568); TLR4 (rs1927911) and 2 intergenic variants (rs869494 and rs4465961) failed to withstand 10,000 permutation test. The genotype/model analysis for all the aforementioned SNPs didn’t withstand Bonferroni correction. Haplotypes generated from the 2 SNPs of PTPN-1 (rs3787345 and rs754118) indicated a 1.3 fold risk of DR for the T-C haplotype, while those generated with the variants of ADIPOQ, AKR1B1 and TLR-4 did not show any significant association (p>0.05).

The allele frequencies of ADIPOQ, AKRIBI, TLR-4, PTPN-1, HFE and TCF7L2 variants were compared with other populations worldwide. The distribution of allele frequencies were similar to the Gujarati Indians living in Houston (GIH; HapMap database) and other Indian populations (1000 Genome project). The frequencies of TCF7L2 rs12255372 (Cicacci et al., 2013) and rs11196205 (Zhang et al., 2015) in the normal controls were similar to the GIH, other Indian and Caucasian populations, but significantly different from the Chinese population. This could be due to different ethnic background of the subjects included in the study. Interestingly, the Chinese population exhibited a higher MAF for rs266729 amongst the DR cases similar to our study, but no significant associations were observed (Li et al., 2015). This difference in the DR susceptibility in the Indian and Chinese cohorts can be due to the differences in the sample groups included in the two cohorts. Similar to our results, the AKRIB1 rs9640883 conferred risk of DR in an Australian cohort (Abhary et al., 2010). In the Polish population, no significant associations were observed with PTPN-1 rs3787345 and rs754118 for DR development/progression however, the MAFs of both the SNPs were similar to the present cohort. This difference in the DR susceptibility could be due to the smaller sample size or the differences in the clinical phenotype of the subjects.
Summary

(Malecki et al., 2008) compared to the present cohort. Although the TLR-4 rs1927911 variant exhibited a relatively higher frequency of G-allele amongst the cases compared to the controls in a North Indian (Singh et al., 2014) and a Chinese population (Xu et al., 2015), no association was observed to DR. Interestingly, we observed a higher A-allele frequency among DR cases in North Indian population, which may indicate a ‘flip-flop’ phenomenon (Lin et al., 2007). Another reason for these differences could be due to the smaller sample sizes in these studies compared to the present cohort. A higher frequency of the HFE rs1800562 A-allele was observed in the DR cases of the present cohort and in a Caucasian population from Slovenian origin (Peterlin et al., 2003) In contrast, the rs1800562 was monomorphic in a South Indian cohort (Balasubbu et al., 2010), which could be due to ethnic differences to their North-Indian counterparts (Riech et al., 2009, Basu et al., 2016). Overall, the MAF distributions were similar to GIH and Caucasian populations for majority of the SNPs. The differences with some population groups could be due to the diverse ethnic background, vagaries of sample sizes, variation in the disease prevalence and the clinical phenotype of the enrolled subjects.

Genotype-phenotype correlation was undertaken to understand the effect of the associated genotypes on the clinical and demographic parameters implicated in DR. Multivariate logistic regression analysis revealed the significant associations of the genotypes of ADIPOQ rs266729 with triglycerides, ADIPOQ rs17366768 with HC, AKR1B1 rs9640883 with triglycerides and HDL, HFE rs1800562 and intergenic variant rs869494 with HC and family history of T2D, respectively. Significant associations were also noted for the TCF7L2 rs11196205 with increasing age and the TLR-4 rs1927911 with the duration of DR and triglycerides. The rs869494 TT genotype provided protection towards family history of T2D [p=0.04; OR=0.34 (0.12-0.95). Interestingly, the other genotypes did not confer any susceptibility to the phenotype parameters in this cohort.

In conclusion, Male Gender, History of Hypertension, BMI were the major lifestyle risk factors in DR. The strongest association was observed with TCF7L2 rs12255372 and rs11196205 (p=0.0008 and p<0.0001, respectively). Further, TCF7L2 haplotypes C-T-T and C-T-G conferred 1.6-1.7 fold risk of DR. Moderate to weak associations were
observed in ADIPOQ (rs266729 and rs17366568); AKRIBI (rs9640883); TLR4 (rs1927911); PTPN-1 (rs3787345 and rs754118); HFE rs1800562 and the intergenic variants (rs869494, rs4465961, rs4516615) with DR. The posthoc power of study was >80% for TCF7L2 rs12255372 and rs11196205, AKRIBI (rs9640883), TLR4 (rs1927911), ADIPOQ (rs17366568), HFE rs1800562 and the intergenic variants (rs869494, rs4465961, rs4516615) with DR. However, the PTPN-1 (rs3787345 and rs754118) exhibited <80% power. The genotype-phenotype analysis did not reveal any significant correlation for DR susceptibility in this cohort.