CHAPTER 6

SUMMARY AND CONCLUSION

- Type II diabetes is a complex genetic disease comprising many metabolic disorders with a common phenotype of glucose intolerance. Patients with type II diabetes inherit a variety of different genetic factors. These genetic factors coupled with the environmental factors become the primary cause of the disease. In order to identify the genes responsible for the disease, different methodological approaches have been incorporated in this study.

- In the first phase of study, the Fisher linear discriminant analysis was employed. In this analysis, genes $G0S2$, $DECR2$, $THRAP6$ and $ZAK$ were identified as predominantly causing type II diabetes as per the high Fisher score. Despite a very high expression value, these identified genes were having a marked discrimination in their expression value between the normal and the disease samples. Genes $G0S2$ and $DECR2$ were reported to play a prominent role in causing diabetes, whereas gene $THRAP6$ and $ZAK$ were involved in metabolic regulation activities.

- The second phase of this work eliminated the redundancy in the microarray data samples. 90% of the genes were found to be similar based on the similarity index. These genes were eliminated by the greedy algorithm approach. The algorithm eliminated all the genes that were having a similarity index score higher than 0.8.

- In the third phase of the study, the SVMRFE approach was used in order to shortlist the dissimilar genes which could discriminate between the normal and the diabetic individual.
• 47 genes in all the four datasets under the study were found to be the most discriminatory in classifying the diabetic and normal individual.

• Among these 47 discriminatory genes, 20 genes were identified to be non-essential genes.

• The protein-protein interaction and pathway analysis of these 20 genes revealed the most prominent 9 genes, possibly responsible for the type II diabetes.

• The FCGR2B, DLG3, SCNN1G, FUT3, HPRT1, APCDD1, USP6NL, ProSAPiP1 and RNASEK are the genes that have been identified to be the potential and less explored/novel drug targets for curing/preventing type II diabetes.

• Genes FCGR2B, DLG3 and SCNN1G from the dataset “Effect of insulin infusion on human skeletal muscle” have shown their interaction with the proteins that cause diabetes.

• Genes FUT3 and HPRT1 from the dataset “human pancreatic islets from normal and type II diabetic subjects (A)” may also be responsible in causing type II diabetes as their interacting protein partners are involved in insulin signalling and hyperglycaemic affects.

• APCDD1 and RNASEK from the dataset “human pancreatic islets from normal and type II diabetic subjects (B)” may also be considered as a drug target in case of the type II diabetes. The APCDD1 has shown its interaction with PPARγ that is already a well-known target for type II diabetes. The RNASEK gene has not been reported previously in any research as a cause for type II diabetes.

• The USP6NL and the ProSAPiP1 from dataset “human skeletal muscle-type II diabetes” are the genes that have been found to be prominent in causing type II diabetes. The chance of USP6NL to be involved in the type II diabetes is quite high. This is because its interacting protein partners are reported to be highly involved in causing the type II diabetes. The
*ProSAPiP1* involvement in causing the type II diabetes has not been reported till date in any study.

- Genes *RNASEK* and *ProSAPiP1* may act as a novel drug target in case of the type II diabetes. This is because their involvement in causing this disease has not been studied in any research previously.

- The *SCNN1G, APCDD1, USP6NL* and *ProSAPiP1* do not possess any crystallographic protein structure.

- As a future extension of the work, the 3-D protein structure of these genes may be identified. Then, a molecular dynamic study can be performed to understand their folding and functional behaviour.

- In this thesis, the nature of the genes involved in causing diabetes has been studied extensively. Experiments such as the modulation of expression of above mentioned genes or the generation of the knockout mouse model have been employed for the research.