Chapter 8. Summary and future implications

8.1 Summary

Allele specific gene silencing (ASGS) by ASP-siRNAs has demonstrated remarkable potential in recent times (Lombardi et al., 2009). This approach is potentially valuable to cure GOF genetic maladies by suppressing disease-causing SNPs and rescue of normal counterpart. Its potential has been successfully reported in various in-vivo experiments including clinical trials (Leachman et al., 2010). Thus, it is anticipated as a better remedial method for presently fatal and incurable disorders, SNPs linked to cancer, viral drug resistance and much more. Despite ample experimental studies on ASP-siRNAs for genetic disorders, a manually curated repository and tool for the prediction of inhibitory efficiency was lacking, which motivated us to make a database (ASPsiDb) and prediction algorithm (ASPsiPred) in this arena. ***ASPsiDb*** is the primary all-inclusive database of 4543 experimentally demonstrated ASP-siRNAs; out if which 422 were ASP-siRNAs with chemical modifications. It delivers complete experimental knowledge of ASP-siRNA sequences accompanied by efficacy details to suppress fully complementary mutant allele (Eff^{mut}) and wild-type counterpart (Eff^{wild}).

ASPsiPred is a two-tier system to predict inhibitory efficiency of two alleles associated with one siRNA *i.e.* Eff^{mut} and Eff^{wild} by $\boldsymbol{ASPsiPred^{SVM}}$ and $\boldsymbol{ASPsiPred^{matrix}}$ respectively. In first tier, we have developed predictive models using different siRNA features computed on experimentally verified 922 ASP-siRNAs with numerical efficiencies to predict Eff^{mut}. Further, in second tier of system, we have built a two-dimensional matrix namely $\mathbf{ASPsiPred^{matrix}}$ from high throughput and experimental rule-derived discoveries stating the influence of a single residue clash between siRNA and mRNA at 19-different positions of siRNA on Eff^{wild}. Besides, to decide optimal selective siRNA having least off-targets, we amalgamated sequence-based (seed- and full) off-targets tool *ASP-siOffTar*. Thus, present algorithm would be enormously valuable to suppress not only disorder causing mutations, but also to know the natural function of allele/s that are not compulsory associated with a disease.

Further, another chief regulatory molecule involved in thesis work is microRNA. They control the expression of host as well as its own genes for its benefit and existence inside

host (Skalsky and Cullen, 2010). In recent past, several miRNAs encoded by numerous viruses have been documented in literature along with experimental information regarding respective target genes (Gottwein, 2013). State-of-the-art evidences also propose that apart from its own miRNAs, virus exploits host-encoded miRNAs to create an appropriate atmosphere inside host (Bruscella et al., 2017). On the other hand, cellular miRNAs are reported to play antiviral defense role during viral infections entitled as antiviral miRNAs (Bruscella et al., 2017). Keeping the importance of miRNA, diverse repositories have been published in recent past for miRNAs and their cognate targets for various organisms.

But specialized archive covering all aspects of role of viral miRNA in pathogenesis was missing. Hence, we have offered an all-inclusive resource "*VIRmiRNA*" comprehending sub-repositories viz "*VirmiRNA*", "*VIRmiRtar*" *and* "*AVIRmir*" for viral miRNAs, targets (host/virus both) and collection of antiviral miRNAs. In the present work, we have delivered updated, annotated and thorough knowledge on experimentally authenticated viral miRNAs along with their isoforms, the biggest assemblage of their cognate targets and the first database on antiviral miRNAs. This all-inclusive reserve along with the exploration tools would be beneficial in decrypting host-virus connections and the development of miRNA derived drugs against the pathogenic viruses.

Similarly, algorithms for classification of precursor and mature viral miRNAs were lacking. Although, numerous tools have published on predicting the precursor and mature miRNAs in various organisms, most of these machine-learning techniques (MLTs) are designed to identify animal and plant miRNAs (Cui et al., 2015; Leclercq et al., 2013). Further, a few of them can predict viral pre-miRNAs, but they have trained their models exclusively on general dataset having limited number of viral pre-miRNAs. Therefore, they didn't perform well on exclusively on viral dataset. Moreover, these algorithms have not been updated for viral dataset in recent past. Hence, we ahve developed ***VIRmiRPred***, combining two prediction systems for viral precursor- (***VIRmiRprePred***) and mature miRNAs (***VIRmiRmatPred***). It is an integrated pipeline developed by using improved sequence, composition, structural and thermodynamic based features.

8.2 Future implications of work done

Thoughtful information on individual aspects of allele specific gene silencing may be exploited in the cure of presently fatal gain-of-function (GOF) genetic ailments. ASPsiDb encompasses manually curated, updated and highly annotated data on ASP-siRNAs and along with knowledge about associated gene mutation, genetic and clinical data. It is also integrated with beneficial tools *viz* ASP-siBLAST and Map. We anticipate that current archive would be valuable to accelerate RNAi based therapeutics for currently fatal genetic disorders.

Like this, in-silico design and prediction of the optimum allele-selective siRNAs would be of immense importance for scientific community. ASPsiPred would be advantageous for *in-silico* designing and predicting efficiency of allele selective siRNAs. This algorithm would be beneficial for investigators who are working on the experimental determination of function of allelic variants. Presently, ASPsiPred is limited to one-mismatch dataset. Nonetheless, the future usage of siRNA selectivity would not only be treasured to suppress disease linked SNPs, but also can also valuable as investigation tool to suppress one splice variant from other. In conclusion, present work being the first prediction algorithm in the field of ASP-RNAi will be helpful in designing of highly efficacious and selective ASP-siRNAs to increase the rate of RNAi centered drugs for human genetic diseases.

Therapeutic progress in the arena of viral infections can be accelerated, if we provide most of the regulatory molecules and victim targets on one comprehensive platform. This resource would be beneficial to uncover, analyze and interpret common tactics adopted by most of the viruses like incorporating cellular miRNAs in viral genome (orthologous miRNAs), exhibiting multiple isomirs, regulating complex host-pathogen by multiplicity and cooperativity. Therefore, Integrated resource for all mentioned three modules along with exploratory analysis would be beneficial for scientific community to accelerate microRNA based antiviral therapeutics. In conclusion, we hope this first comprehensive depository in the arena of viral miRNA biology might be beneficial in expanding understanding of scientific community towards role of miRNAs in deadly viral infections. It would also be useful as initial point to generate a whole new class of drugs. Further, collection of novel miRNAs identified in viruses will benefit in the

expansion of antiviral therapeutics (Jamal et al., 2012).

Computational prediction of viral miRNAs with new feature or methods represents an opportunity to predict novel miRNA candidates including non-classical ones. VIRmiRPred is an improved algorithm incorporating sequence, structure, base-pair related and thermodynamic features for the classification of viral pre- and mature miRNA for improved viral miRNA gene prediction. Being the first pipeline for viral miRNAs, it would not only be helpful to expand novel viral miRNA repertoire, but also for understanding intricacies of host-pathogen interactions.