Conclusion

Deciphering the genetic basis of the infection mechanism and host-specific lifestyle of fungal pathogens is crucial for a better understanding of host-pathogen interactions and to develop effective and novel strategies to combat the pathogen. The genome sequencing of chilli anthracnose fungus, *C. truncatum* was taken up to generate genomic resource for this important pathogen with subcuticular intramural necrotrophic lifestyle, which is the only major member of truncatum clade of *Colletotrichum* genus. The high values of sequencing statistics and existence of conserved fungal genes and core genes from the genus *Colletotrichum* in *C. truncatum* genome reflect the high quality of assembly and gene annotations that were achieved in the present study. RNA-Seq analysis not only served as evidence for the gene structure prediction, but also provided the preliminary cues to design high throughput experiments to study gene regulation and a list of species-specific candidate genes for experimental validation and functional characterization. The comparative genomic analyses of the gene categories relevant for fungal pathogenicity revealed the conserved genes among *Colletotrichum* species with a broad host range as well as some unique genes specific to *C. truncatum*. An expansion in all the major putative pathogenicity gene categories encoding effectors and secreted proteins, CAZymes, proteases, SM associated genes and PHI homologues, was observed in its genome. Many of the genes could be associated with its host-specific subcuticular intramural necrotrophic lifestyle, which need to be functionally characterized.

A refined genome sequence of *C. truncatum* was generated using PacBio SMRT sequencing that substantially reduced the number of gaps in the draft assembly. The genome architecture of six other *Colletotrichum* species along with *C. truncatum* was explored by examining the repetitive element landscape, mainly TEs and SSRs. Retrotransposons represented by Gypsy and/or Copia elements formed the largest fraction of TEs in all species. A recent burst of LTR amplification was observed in the genomes of *C. truncatum*, *C. higginsianum* and *C. scovillei*. The absence of GC-bias or repeat-rich regions in *C. truncatum* contrasted the two-speed genome hypothesis proposed for many of the filamentous fungi and oomycetes, including *C. graminicola* and *C. orbiculare*. TEs in *C. truncatum* were significantly associated with secretory genes, effectors and genes within SM clusters as compared to the random genes. The most prevalent TE families showed signatures of RIP, but absence of homologues of genes required for RIP and lack of sexual stages suggested ancestral activity of RIP machinery. Though there was no direct evidence for the DNA methylation in the TEs, the
presence of genes like cytosine methyltransferase suggested that this could be the active TE silencing mechanism in *C. truncatum*. SSRs formed a small fraction of total genome and were mainly concentrated in intergenic regions. The exons of all the species showed predominance of trinucleotide repeats, except for *C. truncatum*.

In conclusion, this study provides a high quality reference genome sequence and annotation of genes with putative roles in pathogenicity and host interaction of an important *Colletotrichum* species, which would serve as a genomic resource to facilitate further functional and evolutionary studies of this agronomically important fungal pathogen and to develop novel disease control measures in future.