

High-throughput Genomics: bringing order to typing, classification, ecology and evolution

Evolution of fitness, successful lineages and even emergence of novel bacteria was always an unsolved riddle for biologists until the advent of era of modern genomics. Bacterial genomes reflect overall events that introduced variability with time in response to the host or environment resulting in successful lineages of bacteria following a particular lifestyle. The genus *Xanthomonas* represents versatile bacterial isolates having vast array of pathogenic potential displaying host-specificity even at infra-specific level (pathovar level). It comprises diversity of isolates and their successful devastating lineages first reported from India hence, having ecological connection as well. Furthermore, even non-pathogenic *Xanthomonas* (NPX) isolates that seem to be counterparts or ancestors to pathogenic counterparts need to be studied in context. Hence, *Xanthomonas* makes a great model for eco-evo studies and in the present work, genomic insights are focused on inter-species, intra-species and infra-specific (pathovar) level for various pathogens isolated from vast array of hosts or temporal strains of same pathovar or strains following different lifestyles (pathogenic or non-pathogenic) on the same host.

Evolutionary and ecological insights into pathovar diversification of *Xanthomonas citri*

Host specialization is one of the characteristic features of highly evolved pathogens such as the *Xanthomonas* group of phytopathogenic bacteria. Since the hosts involve staple crops and economically important fruits such as citrus, detailed understanding of the diversity and evolution of such strains infecting diverse plants is important for quarantine purposes. In this section, genomic investigation of members of a phylogenetically and ecologically defined group of *Xanthomonas* strains pathogenic to diverse plants, like citrus, mango, pomegranate, grape, cotton etc., were carried out. This group includes the oldest *Xanthomonas* pathovars and also recently emerged pathovars in a particular country where they are endemic. Our high-throughput genomic study has provided novel insights into the evolution of a unique lineage consisting of serious pathogens and their ecological relatives, suggesting the nature, scope, and pattern of rapid and recent diversification. Further, from the level of species to that of clonal variants, the study revealed interesting genomic patterns in diversification of a *Xanthomonas* lineage and perhaps will inspire careful study of the host range of the included pathovars. Present genomic knowledge and resources will be invaluable in surveillance of *X. citri* and its relatives.

Hence, according to analysis till now, *X. citri* is found to be a versatile species comprising of diverse array of closely related pathovars from diverse hosts. Genomic insights at species level were deduced, however, what is happening at infra-specific level? Does a single clone of pathovar from a single host infects the same host or there are different isolates of the pathogen which are phylogenomically distinct and causing disease to the same host down the timeline? To answer these questions, further, genomic insights into the pathogenic isolates from pomegranate but at different time scales were analysed.

Evolutionary forces at infra-specific level leading to devastating XPIN clone

Here, underlying cause of evolution of a successful devastating pathogen *X. citri* pv. *punicae* is deduced. Eventhough pomegranate is cultivated worldwide, the pathogen was first reported in India and is still an epidemic problem in India. Strikingly, according to typing analysis, all the temporal strains were found to be clonal, eventhough, strains isolated after 1980's were reported to cause huge economic losses in contrast to the pre-1980's isolates. It is possible that clonal primitive and benign population which is least destructive suddenly led to have 60% - 80% crop loss in later decades. Complete genomic investigation revealed that a strain isolated during the epidemic (NCPPB3563) was found to have tremendous genomic rearrangements, most diverged as compared to other temporal isolates. It comprised highest number of mutations (leading to most non-synonymous SNPs) and recombination events. Moreover, this strain was reported to have non-synonymous mutations even in DNA repair related genes as well. Hence, this may be a hypermutator strain, which has introduced more dynamics at genome level. Strikingly, strains after this hypermutator strain are carrying *rpoB* gene mutation, conferring rifampicin resistance to these strains but may primarily arise to modify the global transcriptome of the host required for rapid evolution. There are plenty of reports studying effect of different *rpoB* mutations in evolvability of the descendants. Since, some of the *rpoB* mutations may be lethal to the bacterial strains, which has to be compensated by downstream mutations, resulting in more fit, devastating and successful clone. Further, a 10 kb genomic marker encoding terhalose utilization cassette was found to be unique to all the temporal strains at the pathovar level, and was not present in *X. citri* at species level. Hence, this is something unique to the pomegranate pathovar and can be utilized for identification and drug target expetiments. Interestingly, this section has highlighted how a mutator strain is introduced into the population to cause genomic turmoil even at infra-specific level resulting in rise of more devastating clone.

Evolutionary insights into non-pathogenic *Xanthomonas* strains associated with rice and *Citrus* sp.

Xanthomonas isolates that are non-pathogen are reported from *Citrus* sp. and rice plants. It is pertinent to note that *Citrus* sp. and rice plants are native to India. Further, India is the perfect choice in terms of climatic conditions, agricultural practices (monoculture) and diverse habitats for a bacterium to evolve into successful lineages and for the emergence of novel pathovars. How bacterial strains adopt to different lifestyles is always a topic of interest to evolutionary biologists. It is important to carry out genomic relationship of these ecological counterparts to *Xanthomonas* pathovars from India (XPINs).

Towards this I attempted to investigate underlying genomic signatures of pathogenic and non-pathogenic (NPX) strains known to be associated with rice and *Citrus* sp. Overall, phylogenomically, NPX associated with both the hosts represented six different species, out of which four corresponded to potential novel species, reflecting relatively unexplored side of non-pathogenic *Xanthomonas* as compared to vastly studied pathogenic counterparts. Herein, among rice NPX there were two and in *Citrus* NPX there were four different species. Evidences for different evolutionary trajectories followed by bacteria of different lifestyles in general and to my surprise, even among NPX isolated from same host were prominent from present phylogenomic analysis.

Does the different lifestyles followed by different bacterial isolates is reflected in their gene content? Present study allowed systematic scanning of other candidate gene and gene clusters that might have played important role in evolution of pathogenic *Xanthomonas* clusters and genes required for adaptation of bacteria to the host. Primarily, genes required for pathogenicity were T3SS and its effectors, T6SS and HrpG, HrpX regulon. However, amongst them, T6SS, HrpG, HrpX were also present in rice NPX *X. maliensis* and HrpG, HrpX regulon was present in *Citrus* NPX LMG8993. Hence, *X. maliensis* can be considered as intermediary strain, related to pathogenic strains and not to non-pathogenic ones as reaffirmed by phylogenomic analysis as well. On the other hand, T4SS and hence, rate of recombination is more in rice associated non-pathogenic strains, which is absent from RPX (rice pathogenic *Xanthomonas*). Whereas, this was found to be present in both pathogenic and non-pathogenic strains in case of *Citrus* NPX. Further, both hosts associated NPX showed reductive evolution for some genes (*raxX* and *raxST*) of T1SS. Differential repertoire for cell wall degrading enzymes and adhesins were also reported indicating evolution of different

lifestyles. Further, T2SS, *rpf* cluster, iron uptake regulation and siderophores cluster, various two component systems, *gigX* cluster, exopolysaccharide production cluster were present in all the strains under study. As a consequence, all these are related to the adaptation and survival of the bacteria irrespective of the lifestyle followed by it. Hence present study gives us insights of evolution of distinct lifestyles of bacteria which will be crucial for future detection and management of pathogenic strains and exploitation of non-pathogenic strains in plant improvement and protection.

Further, in addition to these, I looked at the misclassifications in six additional XPINs. Interestingly, five were reported in the last century but showed phylogenomic relationship to *X. fuscans* a closely related species to *X. citri*. Sixth one was reported recently as a pathovar of *Commiphora wightii* (guggal plant) and was classified as *X. axonopodis* pv. *commiphoreae*. Unlike other XPINs, my analysis revealed that the XPIN infecting guggal belongs to *X. euvesicatoria*, a species known to have isolates mainly from America, infecting tomato and pepper. Hence, this will be first XPIN in *X. euvesicatoria*, unlike rice or citrus which have centre of diversity in India and are known in India from time immemorial.

Overall, this study is an attempt to genomically investigate *X. citri* and its diverse ecological relatives from different time scales, following different lifestyles. This study has focused either on diverse array of pathovars showing distinct host specificities even at species level (*X. citri*) or genomic turmoil at infra-specific level with time resulting in devastating clone (*X. citri* pv. *punicae*) or different species of *Xanthomonas* associated with same host, but following different lifestyles (*X. citri*, *X. oryzae*, *X. maliensis* and *Xanthomonas* sp. nov.). In addition to these, I have also focused on close relatives of *X. citri* i.e. *X. fuscans*, *X. axonopodis* along with other relatives i.e. *X. oryzae* and *X. euvesicatoria*. Multiple misclassifications underneath genomic turmoil hindered systematic studies of *Xanthomonas* pathogens. Hence, I carried out an eco-evo study, wherein I genomically investigated bacterial versatility, adaptability, ecology, evolution of successful lineages and novel strains taking *Xanthomonas* as the model organism.