Whole genome sequence of Wilsonomyces carpophilus, the causal agent of shot hole disease of stone fruits: insights into secreted proteins of a necrotrophic fungal repository

Asha Nabi
Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir
Sehla K
Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir
Bilal A. Padder
Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir
T.A. Sofi
Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir
Khalid Z. Masoodi
Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir
Mehraj D. Shah (mehraj547@rediffmail.com)
Sher-E-Kashmir University of Agricultural Sciences and Technology Kashmir https://orcid.org/0000-0003-3325-4229

Research Article

Keywords: Necrotroph, Secretome, Illumina HiSeq, PacBio

Posted Date: November 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2194224/v1

Version of Record: A version of this preprint was published at Molecular Biology Reports on March 6th, 2023. See the published version at https://doi.org/10.1007/s11033-023-08243-5.

Abstract

Background

Shot hole is one of the common fungal diseases in stone fruits *viz*., peach, plum, apricot and cherry, and almond in nuts, and is caused by *Wilsonomyces carpophilus*. Our previous research found that both disease incidence and severity is significantly decreased after fungicide application. The pathogenicity studies proved the wide host range of the pathogen infecting all stone fruits and almond among nut crops, however, the mechanisms underlying the host-pathogen interaction was still limited. Additionally, the polymerase chain reaction (PCR) based molecular detection of the pathogen infecting different stone fruits using simple sequence repeat (SSR) markers was also unknown due to the unavailability of pathogen genome.

Methods and Results

We examined the *Wilsonomyces carpophilus* morphology, pathology, and genomics. Whole genome sequencing of the *W. carpophilus* was carried out by Illumina HiSeq and PacBio next generation sequencing (NGS) plate-forms by hybrid assembly. Molecular mechanisms employed by the pathogens to cause disease are altered as a result of constant selection pressure. The studies revealed that the necrotrophs are more lethal, with a complex pathogenicity mechanism and little-understood effector repositories. As a result, we underpin the information about necrotrophic plant pathogenic fungus *W. carpophilus*causing shot hole disease in stone fruits such as peach, plum, apricot and cherry, and almonds among the nut crops. The isolates collected from five different hosts showed a significant difference in their morphology, however, the probability value (p=0.29) suggests in-significant difference in pathogenicity. Here, we report a *W. carpophilus* draft genome assembly of 29.9 megabase (Accession number: PRJNA791904). A total of 10,901 protein-coding genes have been predicted, including heterokaryon incompatibility genes, cytochrome-p450 genes, kinases, sugar transporters, among others. In the genome assembly, we found 2851 simple sequence repeats (SSRs). The most prominent proteins showing the necrotrophic lifestyle of the *W. carpophilus* pathogen were hydrolases, polysaccharide-degrading enzymes, esterolytic, lipolytic, and proteolytic enzymes accounted for 225 released proteins. In the pathogen genome, we discovered tRNAs, rRNAs, and pseudogenes. Among 223 selected fungal species, the top-hit species distribution revealed the majority of hits against the *Pyrenochaeta* species followed by *Ascochyta rabiei* and*Alternaria alternate.*

Conclusions

These findings suggested that the draft genome of the *Wilsonomyces carpophilus* is 29.9 MB based on Illumina HiSeq and PacBio hybrid sequencing assembly. The studies revealed that the necrotrophs are more lethal, with a complex pathogenicity mechanism and little-understood effector repositories. The isolates collected from five different hosts showed a significant difference in their morphology. Total 10901 protein-coding genes have been predicted in the genome including heterokaryon incompatibility and cytochrome-p450 genes, kinases, sugar transporters etc. We also found 2851 simple sequence

repeats (SSRs). The prominent proteins showing necrotrophic lifestyle were hydrolases, polysaccharide-degrading enzymes, esterolytic, lipolytic and proteolytic enzymes accounted for 225 released proteins. In addition, we discovered tRNAs, rRNAs, and pseudogenes and the top-hit species distribution revealed the majority hits against the *Pyrenochaeta* species followed by *Ascochyta rabiei.*

Introduction

With the expansion of our knowledge about threats posed by the plant pathogens, it is necessary to elucidate the molecular basis of plant disease mechanism. The use of fungicide is a common practice and management of plant diseases alone costs around $220 billion dollars annually [1]. Besides cost, fungicide resistance is a major concern to plant pathologists [2]. Therefore, to devise ecofriendly management capsule, it is mandatory to understand the underlying molecular mechanism and signaling events during the disease development. There are many studies related to secretome and virulence strategies of biotrophs[3], however, there is dearth of information about necrotroph genome repository. Although the necrotrophs are far more damaging than biotrophs and cause tremendous production loss in crops [4]. Among others, the *Wilsonomyces carpophilus* is the necrotrophic plant pathogenic fungus that cause notable damage to majority of *Prunus* species such as *P. persica* (Peach), *P. domestica* (Plum), *P. armeniaca* (Apricot), *P. avium* (Sweet cherry) and *P. dulcis* (Almond). The trees infected with the fungus show 30 to 90% production loss [5]. On the fruits, the disease symptoms appears as pinkish to reddish necrotic lesions. On leaves reddish to purplish lesions are formed, which fall down and generate shot holes, giving the tree a ragged appearance due to the abscission layer production [6]. The disease has been documented from Africa, Asia, North, South, and Central America, Australia, and Oceania [7], and recently reported from western Tianshan Mountains of China [8]. Taxanomically *W. carpophilus* belongs to the Dothideomycetes, Pleosporomycetidae, Pleosporales, Dothidotthiaceae group of fungi [9]. The taxonomy of the genus has been controversial over the long time and pathogen is considered to have number of synonyms such as *Stigminacarpophila, Coryneumbeijerinkii, Clasterosporium carpophilum, Thyrostroma carpophilum, Sciniatosporium carpophilum, and Sporocadus carpophilus* [10, 11], however, recent multigene phylogeny considered *Wilsonomyces* as separate genus [9].

Despite of the huge damage and taxonomic crises, there are only few studies in the literature. There is no information on fungal biology, secretome, SSRs (simple sequence repeats) and other metabolites related to pathogenicity. Furthermore, the pathogen necrotrophic life style and lack of host specificity necessitate the identification of pathogenicity arsenals, toxins, enzymes, and other secreted compounds for disease development. It has been discovered that the metabolites or toxins produced by necrotrophs excite rather than prevent programmed cell death (PCD), which can provide the pathogen with metabolizable substrate [12]. From this perspective, defining the role of pathogen-encoded secreted proteins in stimulating or dampening the plant defense system becomes essential.

To gain the insights about the poorly studied *W. carpophilus* fungus, we used the high throughput sequencing toolbox to better understand the life style and developmental dynamics of fungus. In present study, we studied plant pathogenic fungus associated with various *Prunus* species at morphological,

pathological and molecular level to discern the metabolites, toxins and enzymes released by the pathogen and to ascertain its necrotrophic behavior. Although we provide an arrayed list of genes that might play a crucial role during the successful host colonization but many questions related to pathosystem oblige additional research. However, recent optimization of random insertional mutagenesis protocol of the pathogen [13] can be used for functional validation of these pathogenicity arsenals that we report in the present study. The present draft genome assembly is the first genome draft of *W. carpophilus* and will be helpful in understanding the interface in future.

Materials And Methods

Disease status and investigation

The survey was conducted to observe shot hole disease status of *P. persica, P. domestica, P. armeniaca, P. avium* and *P. dulcis* in the University orchard of SKUAST-K, Shalimar, Srinagar, Kashmir (34.0837° N, 74.7973° E). We randomly selected infected trees and each tree was divided into north, south, east and west direction. About thirty leaves were collected from all the directions in each host species and disease incidence was also calculated.

$$Percentdiseaseincidence = \frac{Numberofleavesinfected}{Totalnumberofleavesexamined} X100$$

Fungal culture preparation

Fifty isolates were brought to laboratory to raise fungal cultures. The infected leaves were thoroughly washed under running tap water to remove any dirt. After drying, the margins of leaf lesions were cut into small bits (3x3 mm) along with some healthy portion. The bits were disinfected by immersing in 0.01%mercuric chloride for 30sec. The bits were rinsed twice in sterile distilled water and dried in sterilized petri dishes. Four bits of diseased tissue were placed in a 90 mm petri dish containing Asthana and Hawker's medium [5]. Each plate was incubated at 24 ± 1ºC under a 12 hour photoperiod in a light incubator for seven days. A 5 mm disc from each fungal culture was transferred to Asthana and Hawker's medium as well as on Potato dextrose agar (PDA) medium [14]. To purify the cultures, we used single spore isolation method as described previously [15].

Assessment of Morphological characteristics

Fifty isolates from five hosts were used for morphological assessments such as growth pattern, colony colour and texture. The cultures incubated at 24 ± 1ºC under the 12 hour photoperiod were evaluated every seven days after inoculation. Conidial characteristics such as shape, size, colour and septation of different isolates were assessed using microscope previously calibrated with ocular and stage micrometer.

Pathogenicity test

Pathogenic variability of five isolates from five different hosts was studied. Total of twenty five pathogenicity tests were conducted by inoculating each isolate on their respective hosts and on other host species following detached leaf technique [16]. The freshly collected healthy leaves (3 hours old) were thoroughly washed with sterilized water and placed in a 200 x 200 mm moist chamber. Inoculations were carried out with spore (conidia) suspensions containing 10^5 spores/ml by drop placement method [17] with the help of 20µl micropipette. The moist chambers were then incubated at 24 ± 1ºC with 12 hour photoperiod. The control groups were mock inoculated with distilled water. The one way ANOVA (Analysis of variance) was used to determine pathological variability between the isolates of five hosts.

DNA isolation for genome sequencing

We selected *P. persica* isolates for whole genome sequencing. The DNA of the pathogen isolate was extracted using XcelGen DNA isolation Kit (Xceleris, Ahmedabad, India) according to the manufacturer's instructions. The quality and quantity of extracted DNA was checked using a Qubit 2.0 Fluorimeter (Life Technologies Ltd., Paisley, UK). The integrity of DNA (DIN) was checked using Bioanalyser 2100 (Agilent Technologies, Santa Clara, CA).

Library preparation and genome sequencing

The DNA Library was prepared using NEBNext Ultra DNA Library Prep Kit (Biolabs, England). The library preparation process was initiated with 200ng DNA. The adapters were ligated to both ends of the DNA fragments. These adapters contain sequences essential for binding dual-barcoded libraries to a flow cell for sequencing and PCR amplification. To ensure maximum yield from a limited amounts of starting material, a high-fidelity amplification step was performed using PCR Master Mix.

The whole genome of plant pathogenic fungus *W. carpophilus* was decoded using Illumina HiSeq and PacBio sequencing technologies. *De Novo* assembly of high quality paired end reads was accomplished using Velvet v1.2.10 [18] and the assembly was optimized at Kmer-79 (Supplementary Fig. 1). Further, scaffolding was performed on pre-assembled contigs taking long reads of PacBio using SSPACE-LongRead v1.1 [19]. We aligned Illumina short reads on PacBio long reads (a hybrid approach) using PBJelly software [20] and GapCloser v1.12 to increase the precision of base calling.

Gene prediction and annotation

The assembled genome was subjected to gene prediction using Augustus v2.5.5 for the identification of coding sequences. The predicted protein coding genes were subjected to similarity search against NCBI's non-redundant (nr) database using *Uniprot, KOG* and *Pfam*database of BLASTP algorithm with an e-value threshold of 1e-5. Simultaneously, all the proteins were searched for similarity against BLASTP with an e-value threshold of 1e-5. Comparative analysis of gene annotation in different database was carried out using http://www.interactivenn.net/. Gene Ontology (GO) annotation was obtained using nr database through Blast2GO command line v-1.4.1. GO sequence distributions helps in specifying all the annotated nodes comprising of GO functional groups. Genes associated with the similar functions were assigned to

same GO functional group. The GO sequence distribution was analyzed for all the three GO domains i.e. biological processes, molecular function and cellular components.

Secretome mining

The secretool was used to predict *W. carpophilus* secretome that enables secretome predictions out of amino acid sequence files (http://genomics.cicbiogune.es/SECRETOOL/Secretool.php). The Signal-P (v4.1) and WoLFPSORT (v0.2) were used to identify signal peptides and extracellular localizations in total of 10901 protein coding genes. The TMHMM (v2.0) and PredGPI (http://gpcr.biocomp.unibo.it/predgpi/pred.htm) were used to eliminate sequences with transmembrane domains, ER-retention signal and GPI (glycosylphosphatidyl inositol) anchors, respectively (Supplementary Fig. 2).

Simple sequence repeats

A high-throughput SSR search using MIcroSAtellite (MISA) was performed to identify mono to hexa nucleotide SSR motifs (http://pgrc.ipk-gatersleben.de/misa/download/misa.pl). The default parameters were used so that di-nucleotide pattern should appear at least six times, whereas tri, tetra, penta and hexa nucleotide motifs should appear five times.

Pathway analysis and identification of tRNAs and rRNAs

Pathway analysis, ortholog assignment and mapping of genes to the biological pathways were performed using KEGG automatic annotation server (KAAS). All the gene sequences were compared against the KEGG database using BLASTP with threshold bit-score value of 60 (default).

To identify probable tRNA genes, we used tRNAscan-SE that allows detection of unusual tRNA species with accurate prediction of secondary structures. It includes both prokaryotic and eukaryotic selenocysteinetRNA genes, tRNA-derived repetitive elements and pseudogenes. The RNAmmer 1.2 was used for rRNA gene identification.

Results

Field evaluation

The shot hole disease survey showed highest disease incidence in *P. domestica* (62.5%) followed by *P. persica* (54.1%). We found shot hole disease incidence of 30.2, 35.2, and 12.2% in *P. armeniaca, P. avium and P. dulcis*, respectively. Being deciduous the leaves fall in winter and on the onset of spring the new flush of leaves appear on the tree. The disease incidence in April was low which gradually increased upto June. In the mid of June the lesions started falling down and left holes over the leaves and give ragged appearance to trees.

Diversity in Morphological and pathological characteristics

The fifty isolates cultured on two different media (Asthana hawker's medium and PDA) showed difference in their colony characteristics. Our goal was to observe difference between the isolates of five different hosts however; the isolates of different hosts overlapped in morphological characteristics. We categorized 50 isolates into five groups (each for Asthana hawker's medium and PDA) with Group I having velvety growth with uniform margins on both the media. However, the colony colour on Asthana hawker's medium was brown covered with blackish mass of spores and on PDA it was brownish white. The Group II showed flat to cottony growth with dark golden brown coloration on Asthana hawker's medium. On PDA they showed fluffy colonies with uniform margins having grayish white centre surrounded by olivaceous regions. The Group III displayed dull white cottony growth on Asthana hawker's medium and white fluffy cottony colonies with irregular margins on PDA. The Group IV showed brown colonies on Asthana hawker's medium and cottony dull white to light brown centres surrounded by black mass of spores on PDA. The Group V showed white fluffy growth on Asthana hawker's medium and grayish fluffy colonies with uniform margins on PDA (Supplementary Table 1.). The pathogen produces light brown to dark brown oval septate (3–5 septa) conidia and conidial size ranges from 15.98–40.20 x 7.20-15.28 µm (Fig. 1.).

Based on the pathogenicity test, incubation time varied from four to seven days for the development of disease symptoms. The lesions developed were brown with circular to slightly irregular shape. The *P. persica* isolates showed incubation period of 4, 4, 5, 5 and 4 days on Peach, Plum, Apricot, Sweet cherry and Almond hosts respectively. The isolates obtained from *P. domestica*showed incubation period of 5, 4, 4, 7 and 6 days on Peach, Plum, Apricot, Sweet cherry and Almond hosts respectively. Isolates collected from *P. armeniaca*isolates showed incubation period of 5, 5, 4, 7 and 7 days on Peach, Plum, Apricot, Sweet cherry and Almond hosts respectively. Incubation period of 5, 5, 6, 6 and 7 days on Peach, Plum, Apricot, Sweet cherry and Almond hosts, respectively was shown by the isolates collected from *P. avium* whereas *P.dulcis*isolates showed incubation period of 5, 4, 6, 6 and 5 days on Peach, Plum, Apricot, Sweet cherry and Almond hosts respectively (Fig. 1). We could not find any significant pathological variability between the isolates of five different hosts that is evident from probability value ($p = 0.29$).

Genome sequencing

The whole genome sequence of *W. carpophilus*generated total of 9.0 Gb data comprising of 7.9Gbfrom IlluminaHiSeq and 1.1 Gb from PacBio sequencing platforms. The total genome size of *W. carpophilus*was found to be 29.9Mb (Accession number: PRJNA791904). Under Biosample number (SAMN24657336) we classified fungus as Ascomycota, Saccharomyceta, Pezizomycotina, Leotiomyceta, Dothideomyceta, Dothideomycetes, *Wilsonomyces.* We found 130 scaffolds with maximum scaffold length of 2.1 Mb. The N50 (scaffold) of 662.3Kb was calculated in the genome. The total GC content of the genome was calculated to be 49.77%. The minimum read length was observed to be more than 200bp and less than 600bp and the maximum read length was greater than 5000bp (Supplementary Fig. 3.).

Gene prediction and annotation

We predicted a total of 10901 genes with a total gene length of 13.7 Mb. The average gene length was found to be 1263bp with the maximum gene length of 31.8 Kb. The maximum number of genes having gene length between 1000bp to 5000bp and minimum number of genes with gene length between 200 to 300bpwere observed (Supplementary Fig. 4). Gene Ontology (GO) annotation performed for all three GO domains revealed 4584 genes involved in biological processes, 4858 genes were associated with molecular functions and 3551 genes associated with the cellular components (Fig. 2). Under biological process, different categories such as, cellular metabolic process, response to oxidative stress, signaling and others were highly represented. However, other processes that include growth, carbon utilization, and biological adhesion were the least significant. On molecular grounds, the most significant activities were catalytic, binding, transporter, transcriptional regulator, and anti-oxidant activity, whereas, nutrient reservoir and toxin activities were the least represented. In the cellular component class, cell membrane, organelles, proteins containing complex etc. were highly abundant, and extracellular region, supramolecular complex and nucleoid were the least represented.

Prediction and secretome analysis

For successful infection, the plant pathogenic fungi release arsenal of secreted proteins, particularly effectors. A comprehensive pipeline was designed to carry out the prediction analysis of *W. carpophilus*secretome. Out of 10901 protein encoding genes, we found 225 potentially secreted proteins in *W. carpophilus* genome. We found abundant cell wall degrading enzymes (CWDE) such as cellulases, pectate lyases, cutinases, glycosidases, glucanases, carboxylesterases, laccases, endo-1, 4-beta-galactosidase, rhamnogalacturonanlyase, pectin-esterase, lipolytic protein etc. Besides CWDEs, we found chaperone proteins, mycelial catalase, ricin beta lectin, cathepsin, histidine kinases etc. in the genome (Supplementary Table 2)., These findings indicate that *W. carpophilus*is a necrotrophic plant pathogenic fungus and possess a set of enzymes that is suitable for degradation of cell wall components in the host plants.

Simple sequence repeats

Microsatellites (SSRs) have an active role in upholding genetic variations in genome evolution. These are the tandem repeats of nucleotide motifs of 2–6 base pair (bp) size. They are highly polymorphic and ubiquitously present in all the known genomes. SSRs were identified in different scaffold sequences with the help of MISA perl script. We identified 2851 SSRs with 937 dinucleotides, 1362 trinucleoids, 265 tetranucleotides and few penta and hexanucleotides. Trinucleotide motifs were the most abundant in number comprising of 47.7% of total SSRs in the genome. SSRs with 150 bp flanking region (upstream as well as downstream) were fetched with in-house python script which can be further used for primer designing. We found 2592 SSRs out of 2851 having 150bp flanking region.

Comparison with other fungal genomes

A total of 10901predicted proteins of different genes were identified, out of which, 9427 proteins showed hits in nrdatabase. Among 223 selected fungal species, the top-hit species distribution revealed that the

majority of hits were found to be against the *Pyrenochaeta* species followed by *Ascochyta rabiei* (Fig. 3.). The *Pyrenochaeta* species and *A. rabiei* showed 1812 and 1391 hits, respectively against *W. carpophilus* genome. In *Uniprot* database we found 6331 protein hits, in *KOG* database 4378 protein hits were obtained and only 4272 proteins hits were observed in case of *Pfam* database (Fig. 4). The most abundant proteins identified in *Pfam* database were cytochrome p450, AMP (Anti-microbial peptides) binding genes, adh-shorts, heterokaryon incompatibility genes and kinases etc. Comparative analysis showed that genes annotated in *UniProt*, *Pfam* and *KOG* databases are in concordance with nr database. Among 10901 genes coding proteins, only 3142 genes were common to all databases (Fig. 5).

Scrutiny of metabolism

The proteins mapped to biological pathways represented several metabolic pathways of major biomolecules such as carbohydrates, lipids, nucleotides, amino acids, glycans, cofactors, vitamins, terpenoids, polyketides etc. We found 328 genes involved in carbohydrate metabolism followed by amino acid and lipid metabolism in which 287 and173 genes were involved respectively. However, very less number of genes were involved in terpenoid and polyketide metabolism. In case of genetic information of the pathogen genome, the high number genes were involved in translation, degradation, replication and repair. We also found a significant number of genes that play an important role in signal transduction, transport, catabolism and many other cellular processes of the pathogen (Supplementary Table 3).

Uncovering non-coding RNA species within the genome

We predicted a total of 146 tRNAs out of which only fifteen tRNAs were found on the scaffold one and rest of the tRNAs on the scaffold two (Supplementary Table. 4). We predicted thirty-three 8S rRNA, fifteen18S rRNA, three 28S rRNA and one 5S rRNA. We found these rRNA genes on 1, 10, 12, and 124 scaffolds respectively. In addition, we also predicted pseudogenes in the genome. We represented the all scaffolds, genes, tRNAs and rRNAs of *W. carpophilus*genome in a circos plot (Fig. 6).

Discussion

We investigated *W. carpophilus* on a morphological, pathological, and genetic level to discover its necrotrophic fungal features. The pathogen displayed morphological variation, which is usually attributed to differences in geographic locations and environmental effects; however, we collected the isolates from only one location and assume that it could be due to differences in plant age, plant resistance, and the presence of different pathotypes[21]. The another possible reason could be repeat induced point (RIP) mutation in the genes that leads to variation in isolates and is evident from the presence of pseudogenes in the *W. carpophilus* genome [22, 23]. The non-significant pathological variability in the pathogen indicates that the pathogen can easily breach the surface of all the five hosts [8]. Besides, we explored the pathogen on genomic grounds using high throughput sequencing technology (Illumina HiSeq and PacBio) to gain insights on the complex molecular mechanism of the pathogen. We found 10901 protein coding genes using gene prediction algorithm in the genome. When these genes were blast against the *nr* database, most of the tophits were against the *Pyrenochaeta* spp. followed by *Ascochyta rabiei*. The

protein-coding genes in *W. carpophilus* are almost similar in number (10,901) compared to the genome of the necrotrophic fungal pathogen *Ascochyta rabiei*(10,596) [3], however, *Pyrenochaeta* spp have relatively higher protein coding genes and it could be due to wide range of species in the genus infecting both plants and humans [24, 25].

In nature, the sexual stage of *W. carpophilus* rarely exists [26, 27]. It has been observed that the pathogen with broader host range loose sexuality over a time and exhibit a greater degree of asexuality than those with narrow host ranges [28]. Thus the variability in the fungus could be potentially due to the vegetative hyphal fusion or anastomosis, and is evident from the presence of abundant HET (Heterokaryon incompatibility) genes in the *W. carpophilus* genome. The viable heterokaryon is formed when individuals have same set of HET genotype, whereas individuals with different HET genotypes form incompatible vegetative heterokaryon[29]. The other key genes deployed in the pathogen genome were antimicrobial peptide (AMP) binding genes that are used by the pathogen to overcome host defense. These AMPs are the part of plant's innate immunity system against the pathogen attack. The pathogen has an ability to bind these genes in order to surpass plant defense system. The abundant AMP binding genes in the pathogen suggest that it encounters many host AMPs that ultimately facilitates pathogen to infect wide range of species [30]. The other abundant genes in the pathogen genome were Cytochrome P450, pKinases and sugar transporters etc. Among them Cytochrome P450 genes are involved in the degradation of plant derived toxins and therefore plays an important role in fungal colonization eventually in pathogenesis [31]. Higher number of CYPs also indicates the wide host range of the pathogen requiring more toxins to overcome the phytoalexins[3]. The presence of abundant sugar transporters in the pathogen genome suggests accumulation of sugars [32] which inturn increases abscisic acid levels in an infected cell [33] that can be the possible reason for abscised shot holes symptoms on the leaves of host though it needs an additional research for validation. The protein kinases in the pathogen genome plays key role in various processes of the fungal life cycle such as growth direction, nutrient uptake, stress responses and reproduction [34] thus can be an important factor for survival of the pathogen.

Nowadays necrotrophs are gaining attention due to new theories put forth on the mechanism of the host pathogen dialogue of necrotrophs. The *W. carpophilus* can be an excellent model to study pathogenicity mechanism of wide host range necrotrophs. The presences of cell wall degrading enzymes such as glycosidases, glucanases, pectatelyases and cellulases etc. in the *W. carpophilus*genome are more often predicted in necrotrophs and shows its stronger resemblance to the necrotrophic plant pathogens. Other secretory proteins such as chaperone proteins called as heat shock proteins play an important role in maintaining the integrity of the pathogen in the adverse conditions, thus making it more viable for infection[35]. Laccasse precursor that plays an important role in lignin depolymerization of infected host[36]and number of other enzymes such as endo-1, 4-beta-galactosidase, cutinase, rhamnogalacturonamlyase, pectin-esterase, pectate lyase, lipolytic protein etc. plays decisive role in the pathogenicity mechanism. However, these pathogenicity determinants need further characterization and validation. The present study has revealed the number of pathogenicity genes that can be easily targeted to render pathogen ineffective. Our research opened new opportunities for the comprehensive genomic

study of a variety of biological, metabolic and pathological aspects that make the *W. carpophilus*a successful necrotrophic pathogen. Therefore, it is an opportune time to go beyond the conventional neutral genetics by identifying, analyzing, site specific targeting of pathogenicity determinants and re-modelling the core effector repositories.

Declarations

Acknowledgment

The authors are highly thankful to M/S Xcelris Labs Limited, Ahmedabad-380015, India, for their service and help provided during the present study.

Author contributions

All the authors have equally contributed to this article:

MF- Design of the study, writing of the manuscript, pathogen isolation and inoculation, and data analysis;

MDS- Idea and design of the study, writing of the manuscript, pathogen identification and data generation;

BAP- Design of the study, writing of the manuscript and data analysis;

TAS- Pathogen purification and identification;

KZM- Microscopy and data analysis;

AN- Pathogen isolation and pathogenicity;

SH- Pathogen purification, inoculation and pathogenicity;

MA- Design of the study and writing of the manuscript.

Conflict of interest

The authors assure that the manuscript has not been submitted for publication elsewhere. All the authors have contributed equally to this article, and all have agreed to submit it for publication in this journal. All the authors declare that there are no conflicts of interest to disclose.

Ethical approval

This research is about the whole genomic analysis in fungus *(Wilsonomyces carpophilus)*. Human participants/animals were not involved in this study.

Informed consent

Not applicable in this study.

References

1. FAO (2019) New standards to curb the global spread of plant pests and diseases. Web page of the Food Agriculture Organization of the United Nations
2. Hahn M (2014) The rising threat of fungicide resistance in plant pathogenic fungi: *Botrytis* as a case study.Journal of Chemical Biology; 7; 133 – 41.
3. Verma S, Gazara RK, Nizam S, Parween S, Chattopadhyay D, Verma PK (2016) Draft genome sequencing and secretome analysis of fungal phytopathogen *Ascochyta rabiei* provides insight into the necrotrophic effector repertoire. Sci Rep 6:1–14
4. Murray GM, Brennan JP (2009) Estimating disease losses to the Australian wheat industry. Australas Plant Pathol 38:558–570
5. Nabi A, Padder B, Dar M, Ahmad M (2018) Morpho-cultural, pathological and molecular variability in *Thyrostroma carpophilum* causing shot hole of stone fruits in India. Eur J Plant Pathol 151:613–627
6. Ivanová H, Kalocaiová M, Bolvansky M (2012) Shot-hole disease on *Prunus persica*-the morphology and biology of *Stigmina carpophila*. Folia Oecol 39:21
7. Väcäroju C, Zalä C, Cristea S (2008) Research on the morphology and biology of the *Stigmina carpophila* fungus. scientific conferences with international participation durable agriculture-Agriculture of future, 4th edition. Craiova (Romania): Faculty of Agriculture, University of Craiova,
8. Ye S, Jia H, Cai G, Tian C, Ma R, Morphology (2020) DNA Phylogeny, and Pathogenicity of *Wilsonomyces carpophilus* Isolate Causing Shot-Hole Disease of *Prunus divaricata* and *Prunus armeniaca* in Wild-Fruit Forest of Western Tianshan Mountains, China. Forests ; 11; 319
9. Marin-Felix Y, Groenewald J, Cai L, Chen Q, Marincowitz S, Barnes I et al (2017) Genera of phytopathogenic fungi: GOPHY 1. Stud Mycol 86:99–216
10. Adaskaveg J, Ogawa J, Butler E (1990) Morphology and ontogeny of conidia in *Wilsonomyces carpophilus*, causal pathogen of shot hole disease of *Prunus* species. Mycotax 37:275–290
11. Ahmadpour A, Ghosta Y, Javan-Nikkhah M, Fatahi R, Ghazanfari K (2009) Isolation and pathogenicity tests of Iranian cultures of the shot hole pathogen of *Prunus* species, *Wilsonomyces carpophilus*. Australasian Plant Disease Notes 4:133–134
12. Liu Z, Zurn JD, Kariyawasam G, Faris JD, Shi G, Hansen J et al (2017) Inverse gene-for-gene interactions contribute additively to tan spot susceptibility in wheat. Theoretical Appl Genet 130:1267–1276
13. Rasool RS, Padder BA, Alam A, Shah MD, Masoodi KZ, Khan NA et al (2020) *Thyrostroma carpophilum* insertional mutagenesis: A step towards understanding its pathogenicity mechanism.Journal of Microbiological Methods;105885

14. Aryal S (2015) Potato Dextrose Agar (PDA)—Principle, uses, composition, procedure and colony characteristics. Retrieved on ; 18; 2017
15. Tuite J (1969)
16. Sukumar j, Ramalingam A (1981) Detached leaf technique to study leaf spot and other foliar diseases of mulberry.Indian Phytopathology
17. Kanchana-Udomkan C, Taylor P, Mongkolporn (2004) Development of a bioassay to study anthracnose infection of *Capsicum chinense* Jacq. fruit caused by *Colletotrichum capsici*. Thai Journal of Agricultural Science
18. Zerbino D, Birney E, Velvet Algorithms for de novo short read assembly using de Bruijn graphs..Genome Research 200818;821
19. Boetzer M, Pirovano W (2014) SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. BMC Bioinformatics 15:1–9
20. English AC, Richards S, Han Y, Wang M, Vee V, Qu J et al (2012) ; 7; e47768
21. Iqbal U, Mukhtar T (2014) ; 2014
22. Li Y, Yang R-H, Jiang L, Hu X-D, Wu Z-J, Yao Y-J (2017) rRNA pseudogenes in filamentous ascomycetes as revealed by genome data, vol G3. Genes, Genomes, Genetics, pp 2695–2703
23. Haas BJ, Zeng Q, Pearson MD, Cuomo CA, Wortman J (2011) Approaches to fungal genome annotation. Mycology 2:118–141
24. Toh YF, Yew SM, Chan CL, Na SL, Lee KW, Hoh C-C et al (2016) ; 11; e0162095
25. Aragona M, Minio A, Ferrarini A, Valente MT, Bagnaresi P, Orrù L et al (2014) De novo genome assembly of the soil-borne fungus and tomato pathogen *Pyrenochaeta lycopersici*. BMC Genomics 15:313
26. Samuel G (1927) On the shot-hole disease caused by *Clasterosporium carpophilum* and on the'shot-hole'effect. Ann Botany 41:375–404
27. Wilson EE (1937) The shot-hole disease of stone-fruit trees.
28. Newman TE, Derbyshire MC (2020) The evolutionary and molecular features of broad host-range necrotrophy in plant pathogenic fungi.Frontiers in plant science;1772
29. Hall C, Welch J, Kowbel DJ, Glass NL (2010) Evolution and diversity of a fungal self/nonself recognition locus. PLoS ONE 5:e14055
30. Breen S, Solomon PS, Bedon F, Vincent D (2015) Surveying the potential of secreted antimicrobial peptides to enhance plant disease resistance. Front Plant Sci 6:900
31. Shin J, Kim J-E, Lee Y-W, Son H (2018) Fungal cytochrome P450s and the P450 complement (CYPome) of Fusarium graminearum. Toxins 10:112
32. Doidy J, Grace E, Kühn C, Simon-Plas F, Casieri L, Wipf D (2012) Sugar transporters in plants and in their interactions with fungi. Trends Plant Sci 17:413–422
33. Formela-Luboińska M, Chadzinikolau T, Drzewiecka K, Jeleń H, Bocianowski J, Kęsy J et al (2020) The role of sugars in the regulation of the level of endogenous signaling molecules during defense

response of yellow lupine to *Fusarium oxysporum*. Int J Mol Sci 21:4133

34. Turrà D, Segorbe D, Di Pietro A (2014) Protein kinases in plant-pathogenic fungi: conserved regulators of infection. Annu Rev Phytopathol 52:267–288
35. Neckers L, Tatu U (2008) Molecular chaperones in pathogen virulence: emerging new targets for therapy.Cell Host & Microbe; 4; 519 – 27.
36. Edens WA, Goins TQ, Dooley D, Henson JM (1999) Purification and characterization of a secreted laccase of *Gaeumannomyces graminis* var. tritici. Appl Environ Microbiol 65:3071–3074

Figures

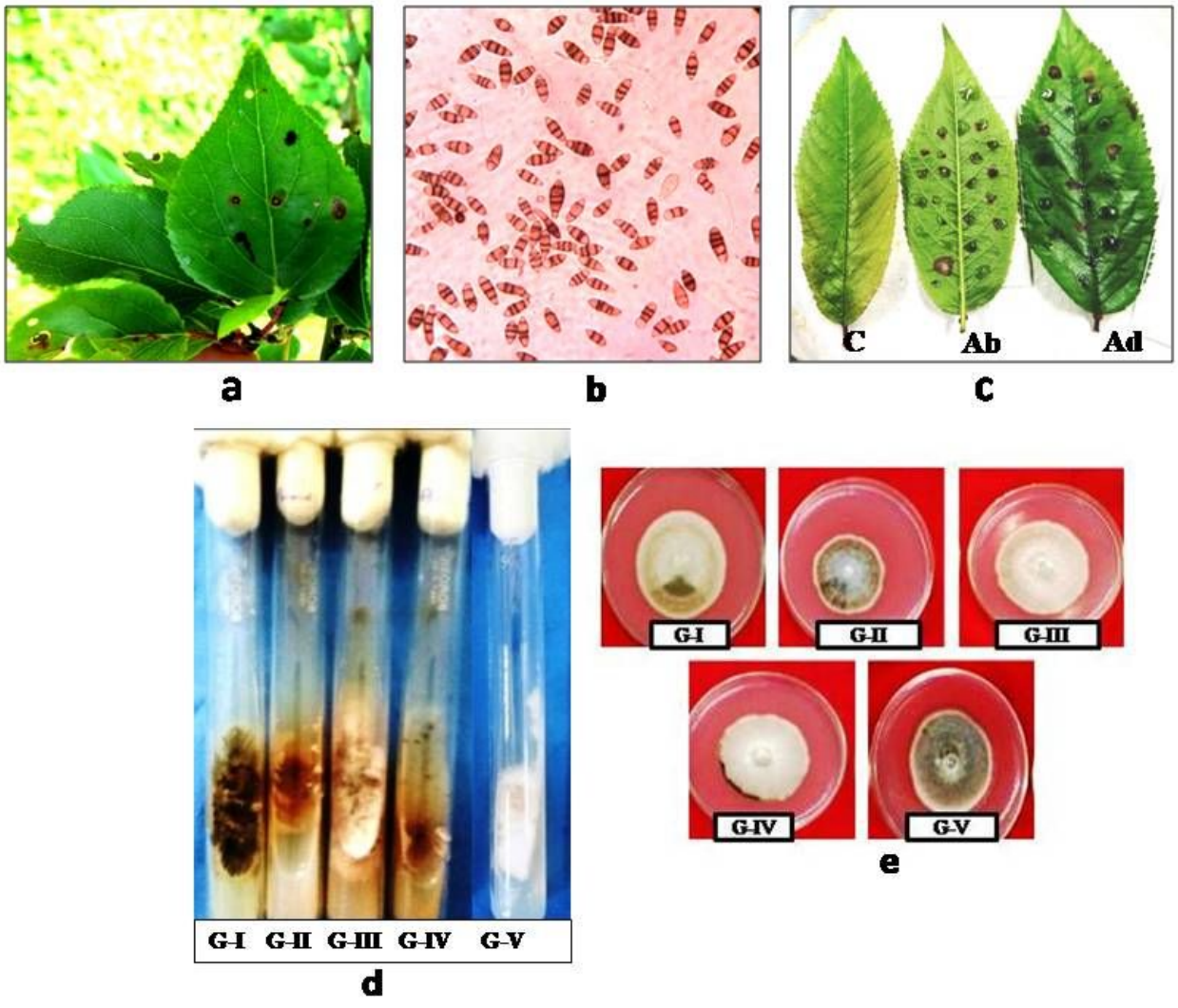

Figure 1

(a) Shot hole infected leaves (b) Conidia under 400X (c) In vitro Pathogenicity test on leaves, where, Ad represents adaxial surface of the leaf, Ab represents abaxial surface of the leaf and C represents control

(inoculated with water) (d) Pathogen culture raised on Asthana Hawker's medium (e) Pathogen culture raised on PDA (G stands for Group).

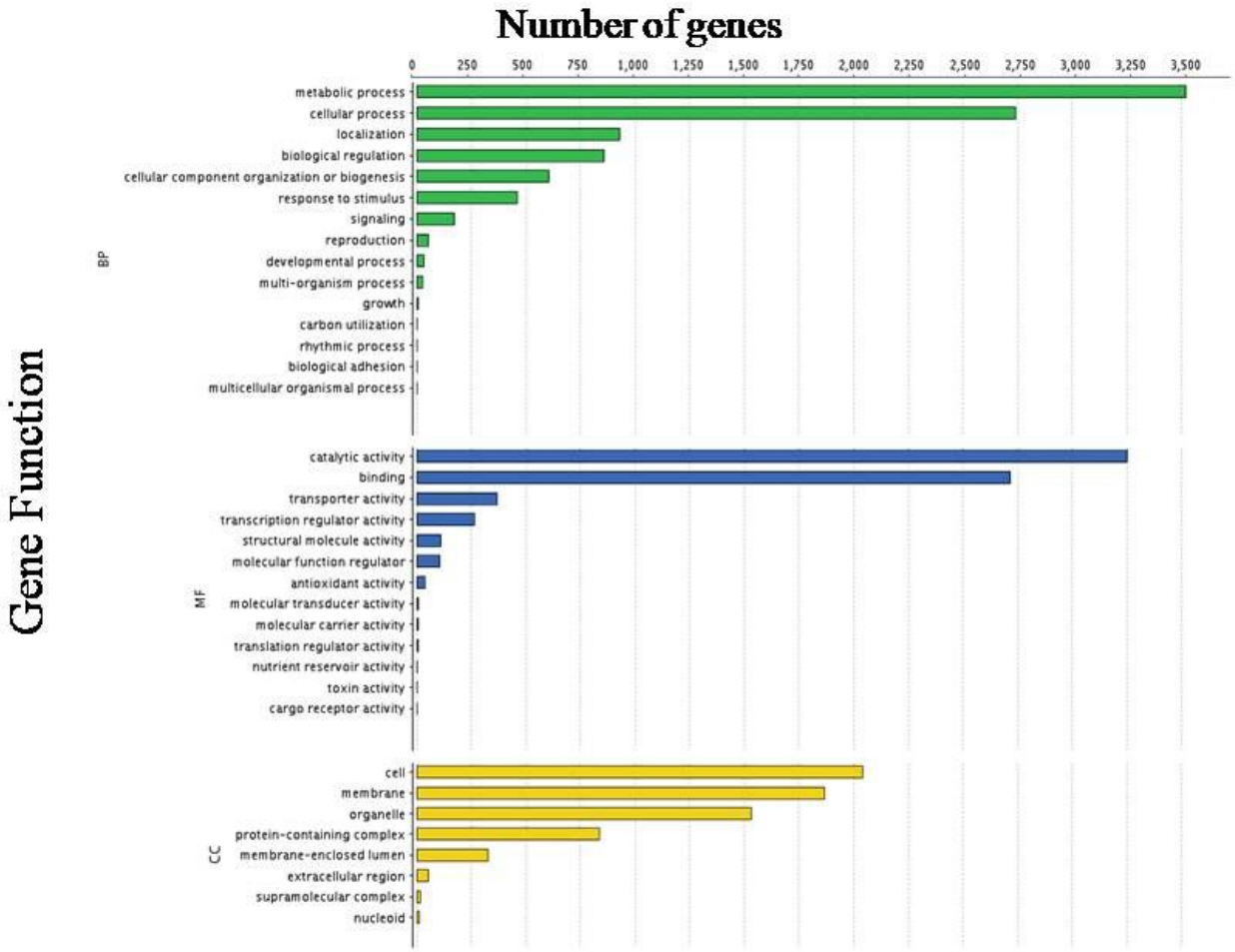

Figure 2

Gene annotation for all the three GO domains i.e. Biological Processes (BP), Molecular Function (MF) and Cellular Components (CC). The bars represent the number of genes involved in the biological, cellular and molecular functions of the *W. carpophilus* genome.

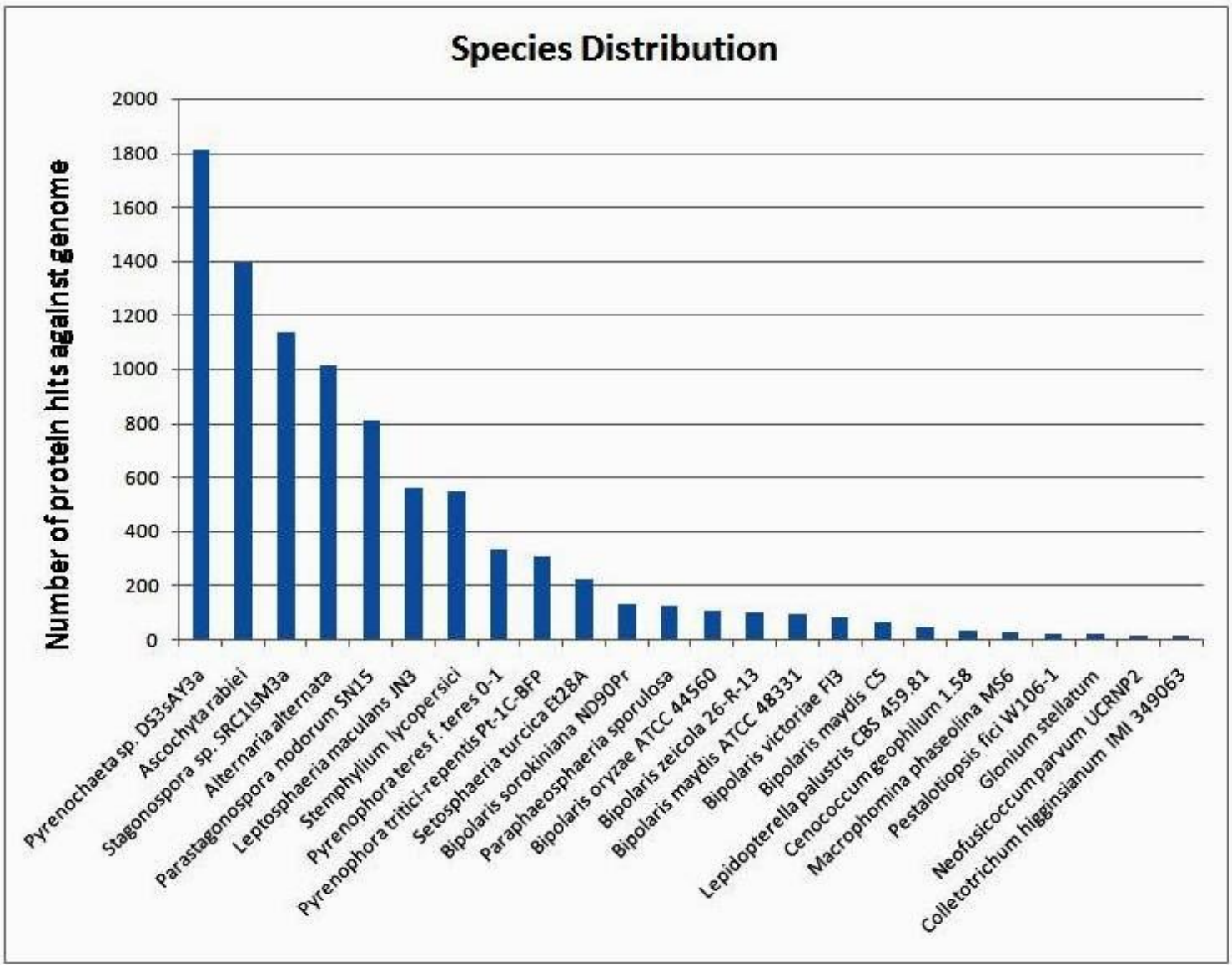

Figure 3

Similarity in genes with other fungal species. The bars represent no. of genes similar to other fungal species with maximum similarity with *Pyrenocheata sp.*

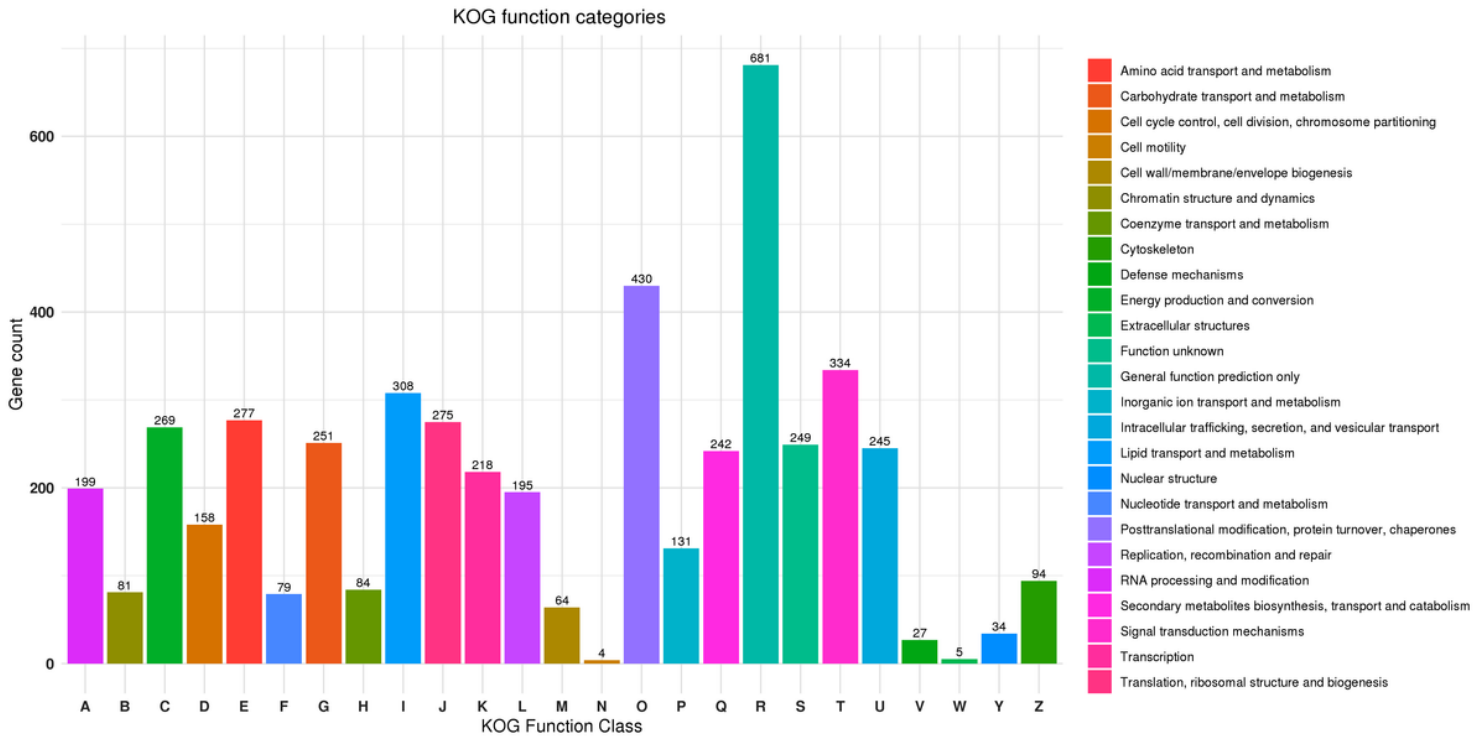

Figure 4

KOG functional annotation of genes, the height of the bars represents the number of genes involved in a particular function and color represents that function.

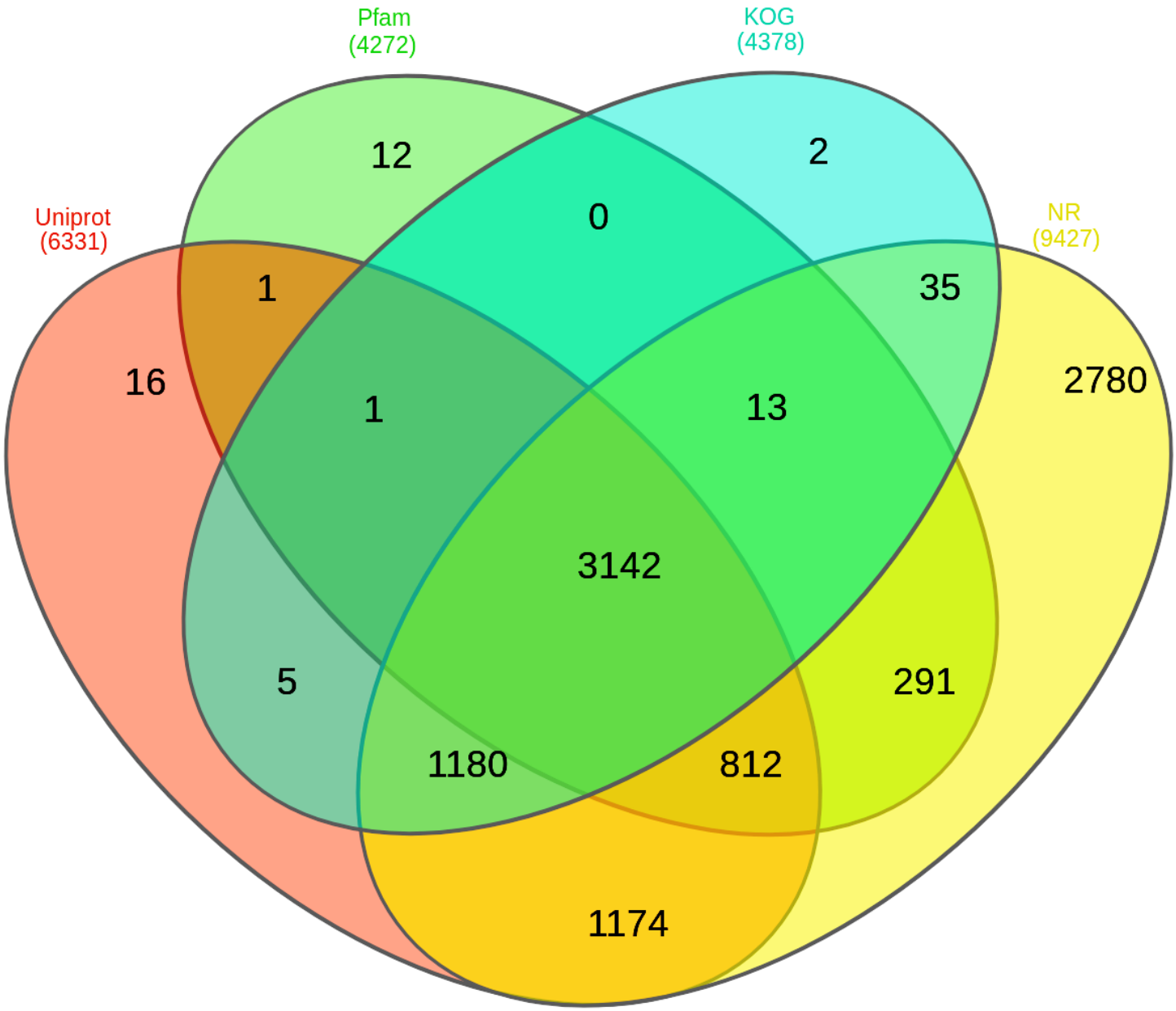

Figure 5

Venn diagram represents the number of genes that are common between four data bases. Out of 10901 genes only 3142 genes are common to each data base

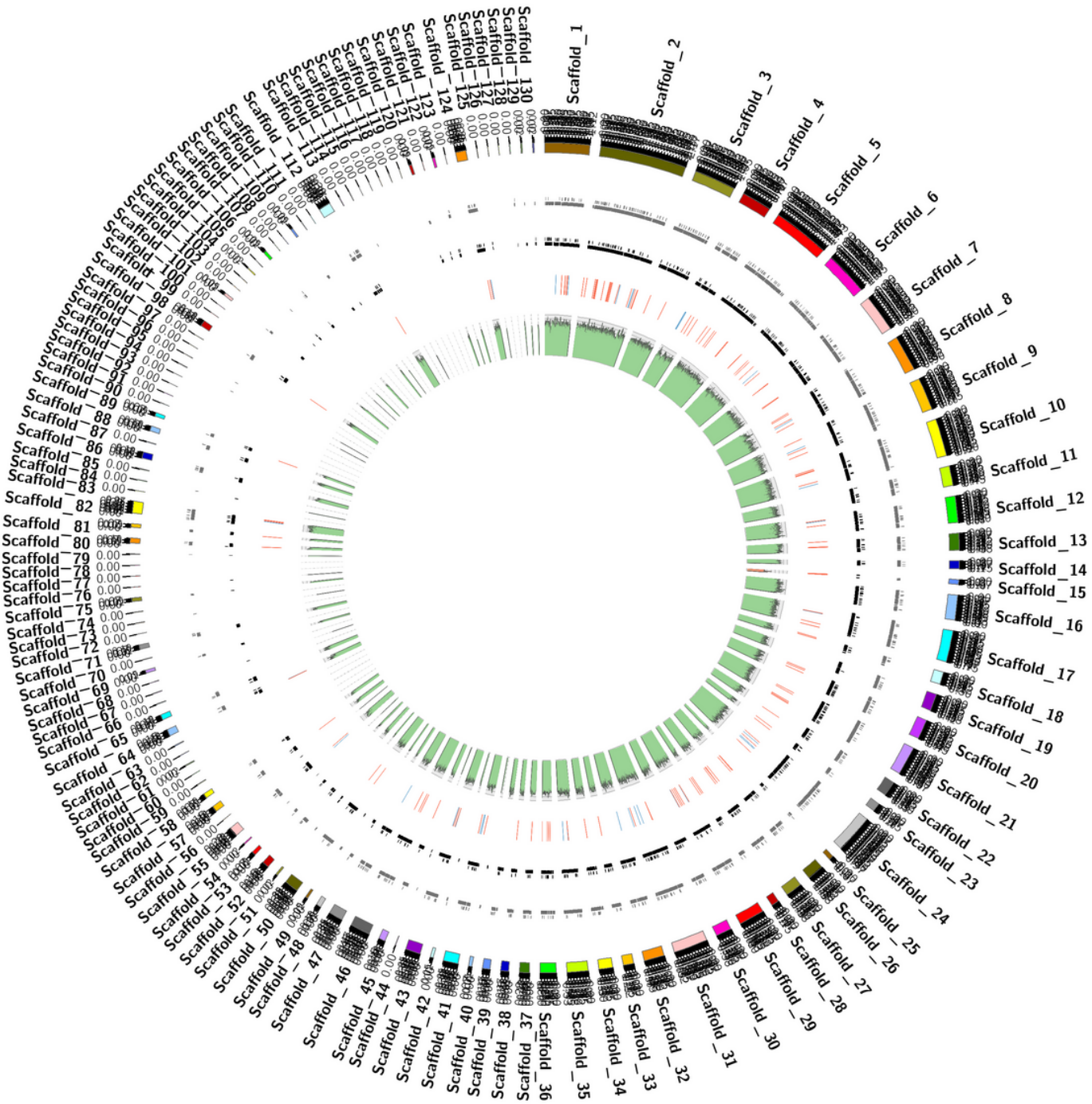

Figure 6

Schematic representation assembled genome. Circle A and Circle B represents gene predicted in the positive frame. Circle C represents gene predicted in the negative frame, Circle D represents tRNA (red tiles) and rRNA (blue tiles). The inner E circle represents distribution of GC content. The red shade in the E circle represents regions of low GC contents (less than 35%).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplimentarymaterial.pdf