



Comprehensive genome analysis of a pangolin-associated *Paraburkholderia fungorum* provides new insights into its secretion systems and virulence

Ka Yun Tan^{1,2}, Avirup Dutta^{2,5}, Tze King Tan^{2,6}, Ranjeev Hari², Rofina Y. Othman^{1,3} and Siew Woh Choo⁴

¹Institute of Biological Sciences, Faculty of Science, Universiti Malaya, Kuala Lumpur, Malaysia

²Genome Informatics Research Laboratory, Centre for Research in Biotechnology for Agriculture (CEBAR), High Impact Research Building, Universiti Malaya, Kuala Lumpur, Malaysia

³Centre for Research in Biotechnology for Agriculture (CEBAR), Level 3, Research Management & Innovation Complex, Universiti Malaya, Copenhagen, Kuala Lumpur, Malaysia

⁴College of Science and Technology, Wenzhou-Kean University, Wenzhou, Zhejiang Province, China

⁵Current affiliation: The Novo Nordisk Foundation Center for Basic Metabolic Research, Human Genomics and Metagenomics in Metabolism, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

⁶Current affiliation: Cancer Science Institute of Singapore, National University of Singapore, Singapore

ABSTRACT

Background. *Paraburkholderia fungorum* (*P. fungorum*) is a Gram-negative environmental species that has been commonly used as a beneficial microorganism in agriculture as an agent for biocontrol and bioremediation. Its use in agriculture is controversial as many people believe that it could harm human health; however, there is no clear evidence to support.

Methodology. The pangolin *P. fungorum* (pangolin Pf) genome has a genomic size of approximately 7.7 Mbps with N50 of 69,666 bps. Our study showed that pangolin Pf is a *Paraburkholderia fungorum* supported by evidence from the core genome SNP-based phylogenetic analysis and the ANI analysis. Functional analysis has shown that the presence of a considerably large number of genes related to stress response, virulence, disease, and defence. Interestingly, we identified different types of secretion systems in the genome of pangolin Pf, which are highly specialized and responsible for a bacterium's response to its environment and in physiological processes such as survival, adhesion, and adaptation. The pangolin Pf also shared some common virulence genes with the known pathogenic member of the Burkholderiales. These genes play important roles in adhesion, motility, and invasion.

Conclusion. This study may provide better insights into the functions, secretion systems and virulence of this pangolin-associated bacterial strain. The addition of this genome sequence is also important for future comparative analysis and functional work of *P. fungorum*.

Submitted 7 April 2020

Accepted 25 July 2020

Published 3 September 2020

Corresponding authors

Rofina Y. Othman,

yasmin@um.edu.my

Siew Woh Choo, cwoh@wku.edu.cn

Academic editor

Joseph Gillespie

Additional Information and
Declarations can be found on
page 16

DOI 10.7717/peerj.9733

© Copyright
2020 Tan et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Bioinformatics, Taxonomy

Keywords Secretion system, Taxonomy, Burkholderia, RAST

INTRODUCTION

The genus *Burkholderia* is a group of Gram-negative, motile, obligate aerobic rod-shaped bacteria. Some members of the Burkholderiales species are environmentally important and some are pathogenic (Landers et al., 2000). The genus consists of more than 120 species spreading across soil and water environments that can be utilized in bioremediation and bioconversion application (Morya, Salvachúa & Thakur, 2020), while some being opportunistic pathogens of plants or humans (Estrada-de los Santos et al., 2013). The most famous among the pathogenic members of *Burkholderia* are *B. mallei* and *B. pseudomallei*. The *B. mallei* can cause glanders in horses, donkeys, and humans (Whitlock, Estes & Torres, 2007), whereas the *B. pseudomallei* can cause melioidosis in humans and animals (Limmathurotsakul et al., 2016). Another group of *Burkholderia*, *Burkholderia cepacia complex* (Bcc) is a group comprising at least 20 species which are known to be opportunistic pathogens in immunocompromised patients, especially those suffering cystic fibrosis (Rojas-Rojas et al., 2019; Leitão et al., 2017).

Paraburkholderia has been split from the *Burkholderia* genera (Sawana, Adeolu & Gupta, 2014; Beukes et al., 2017; Estrada-de Los Santos et al., 2018). It gained considerable importance for their abilities to fix nitrogen, promote plant growth, and degrade recalcitrant chemical compounds. It has been sustainable as alternatives to chemical fertilizers in agriculture, and also bioremediation of the impacted environment (Andreolli et al., 2011; Kaur, Selvakumar & Ganeshamurthy, 2017). *P. fungorum* can be isolated from diverse ecological niches including samples from human, animals, and plants (Coenye et al., 2001). Although previous studies reported the isolation of *P. fungorum* in synovial tissue of an overweight but healthy patient (Loong et al., 2019), and has also been shown to cause septicemia (Gerrits et al., 2005), and infectious granuloma (Zhang et al., 2014), no clear evidence for its pathogenicity has been reported. However, there are debates on its suitability as an agent of biodegradation and bioremediation because many people believe that it could affect human health (Jones, Dodd & Webb, 2001; Nally et al., 2018).

In a previous genome sequencing project, we sequenced the genomes of the infected cerebellum and cerebrum of a female pregnant Malayan pangolin (*Manis javanica*) and its infected fetus (Tan et al., 2016b; Choo et al., 2016b). A bacterium genome was identified in the pangolin genome data and identified as a *Paraburkholderia* species (hereto referred to as pangolin Pf). In this study, the pangolin Pf genome was further analysed to advance our understanding of its functions and virulence.

MATERIALS & METHODS

Genome datasets

The genome sequence of pangolin Pf (accession numbers: CP028829–CP028832) can be downloaded from the GenBank database at the National Center for Biotechnology Information (NCBI) website. It was assembled from the previous pangolin genome data (accession numbers: SRR3949798, SRR3949787, SRR3949728, SRR3949765, and SRR11935558) generated from the genome sequencing of an infected female Malayan pangolin (Tan et al., 2016b; Choo et al., 2020). Briefly, the raw reads generated from three

P. fungorum-infected tissues (cerebrum, cerebellum and fetus) of the Malayan pangolin (manuscript in preparation) were pooled and assembled by mapping the reads to the reference genome, *Paraburkholderia fungorum* ATCC BAA-463 (Accession numbers: CP010024–CP010027) downloaded from NCBI using CLC Assembly Cell software with default parameters (D’Cruze *et al.*, 2011). The mapped reads were assembled into a draft genome sequence for downstream analyses (Table 1).

Genome annotation

The pangolin Pf genome sequence was submitted to the Rapid Annotation using Subsystem Technology (RAST) server for gene prediction and functional annotation (Aziz *et al.*, 2008). To ensure the uniformity in the annotations for comparative analysis, the genome sequences of 17 closely related species (Table 1), were retrieved from NCBI (<http://www.ncbi.nlm.nih.gov>) (Pruitt, Tatusova & Maglott, 2007) and annotated by the RAST server.

The core-genome SNP-based phylogenetic inference

The single nucleotide polymorphism (SNP) refers to a single-base variation in a DNA sequence (Vignal *et al.*, 2002). The SNPs are the most abundant genetic variation, in a region that can be used to assess bacterial diversity and for taxonomic classification. The core-genome SNP-based phylogenetic analyses have successfully been used to infer the taxonomic relationships of many bacterial species (Ang *et al.*, 2016; Choo *et al.*, 2016a; Choo *et al.*, 2014; Zheng *et al.*, 2017). PanSeq was used to align all genome sequences and SNPs were identified in the conserved genomic regions among all species (Laing *et al.*, 2010). All SNPs were extracted and concatenated into a long sequence for each genome. Phylogenetic tree based on the core-genome SNP was constructed using the genome sequences of the pangolin Pf and 17 selected closely related species (Table 1). The tree was constructed using MEGA-X (Kumar *et al.*, 2018). Neighbour-joining tree was inferred using the Kimura’s two parameter model and nodal support was estimated using 1,000 replicates.

Whole-genome Average Nucleotide Identity (ANI) analysis

To evaluate the genetic relatedness between the genome of pangolin Pf and its closely related species, the ANI values were calculated based on the method as previously described (Ang *et al.*, 2016; Rodriguez & Konstantinidis, 2014; Goris *et al.*, 2007; Tan *et al.*, 2016a). Two-way BLAST was chosen and only the forward and reversed-matched orthologs were used in the calculations. For robustness, the BLAST match was set for at least 50% identity at the nucleotide and amino acid level and a sequence coverage of at least 70%.

Genome Island and prophage prediction

All putative genomic islands (GIs) in the pangolin Pf genome were predicted using IslandViewer4, which is based on unique features in codon usage, dinucleotide sequence composition and the presence of mobile element genes (Bertelli *et al.*, 2017; Langille & Brinkman, 2009). The RAST provided GenBank file was submitted into IslandViewer4 web server for GI prediction.

The identification of putative prophages in the pangolin Pf genome was performed using the PHAST (PhAge Search Tool) with default thresholds (Zhou *et al.*, 2011). PHAST uses

Table 1 Genome statistics. The table shows the number of contigs, protein-coding genes, RNA, GC content, and virulence factor (VF) of pangolin Pf and members of the Paraburkholderiales and Burkholderiales.

Strain name	Genome size	No. contigs	GC (%)	CDS	RNA	tRNAs	Subsystems	VF
<i>P. fungorum</i> Pf	7.7	3	62.2	6,746	82	18	513	186
<i>P. fungorum</i> ATCC BAA-463	8.6	3	61.8	8,010	81	18	527	189
<i>P. phytofirmans</i> PsJN	8.1	2	62.3	7,562	81	19	517	192
<i>P. xenovorans</i> LB400	9.7	3	62.6	9,196	80	18	564	187
<i>P. phenoliruptrix</i> BR3459a	6.8	2	63.6	6,327	80	18	507	175
<i>P. phymatum</i> STM815	6.1	2	62.7	5,864	79	18	491	130
<i>B. mallei</i> ATCC 23344	5.8	2	68.5	5,738	67	18	501	282
<i>B. multivorans</i> ATCC 17616	6.8	3	66.8	6,562	80	18	524	172
<i>B. cepacia</i> GG4	6.4	2	66.7	6,119	66	18	529	181
<i>B. thailandensis</i> MSMB59	6.7	2	67.7	6,496	69	–	505	273
<i>B. pseudomallei</i> K96243	7.2	2	68.1	7,053	71	18	525	281
<i>B. oklahomensis</i> EO147	7.3	2	66.9	7,122	69	18	517	274
<i>B. cenocepacia</i> AU 1054	7.2	3	66.9	6,821	85	18	542	208
<i>B. pyrrocinia</i> DSM 10685	7.8	3	66.5	7,337	85	18	541	222
<i>B. vietnamiensis</i> LMG 10929	6.8	3	66.9	6,325	84	18	532	178
<i>B. ubonensis</i> MSMB22	7.1	3	67.3	6,531	89	18	529	201
<i>B. glumae</i> LMG 2196 ATCC 33617	6.4	2	68.5	5,832	81	18	493	233
<i>B. gladioli</i> ATCC 10248	8.5	2	67.9	7,511	81	18	521	199

a web server to perform a series of database comparisons and phage feature identification analyses, locating and annotating prophage sequences.

Comparative virulence gene analysis

To identify the potential virulence factors, the RAST-predicted protein-coding genes in the genome of pangolin Pf were BLAST searched against the Virulence Factors Database (VFDB) ([Chen et al., 2016](#); [Chen et al., 2005](#)) using the BLASTP of the BLAST software package ([Altschul et al., 1997](#)). In-house developed Perl scripts were used to screen out the genes that were identified as orthologous to virulence genes with at least 40% sequence identity and at least 40% sequence coverage in query and subject from the BLAST search. The filtered results were then further processed using in-house developed R scripts for clustering and constructing a graphical heat map with dendrograms, sorted according to similarities across the strains and genes.

RESULTS

Genome summary

The genome sequence of pangolin Pf was assembled from the raw sequencing reads generated from the infected tissues (cerebrum and cerebellum) of a pregnant female pangolin and its infected fetal muscle (accession ID: [SRR3949798](#), [SRR3949787](#), [SRR3949728](#), [SRR3949765](#), and [SRR11935558](#)). The assembled genome of pangolin Pf has a genomic size of 7.7 Mbp (7,746,865 bp), covering 86% of the reference genome *P.*

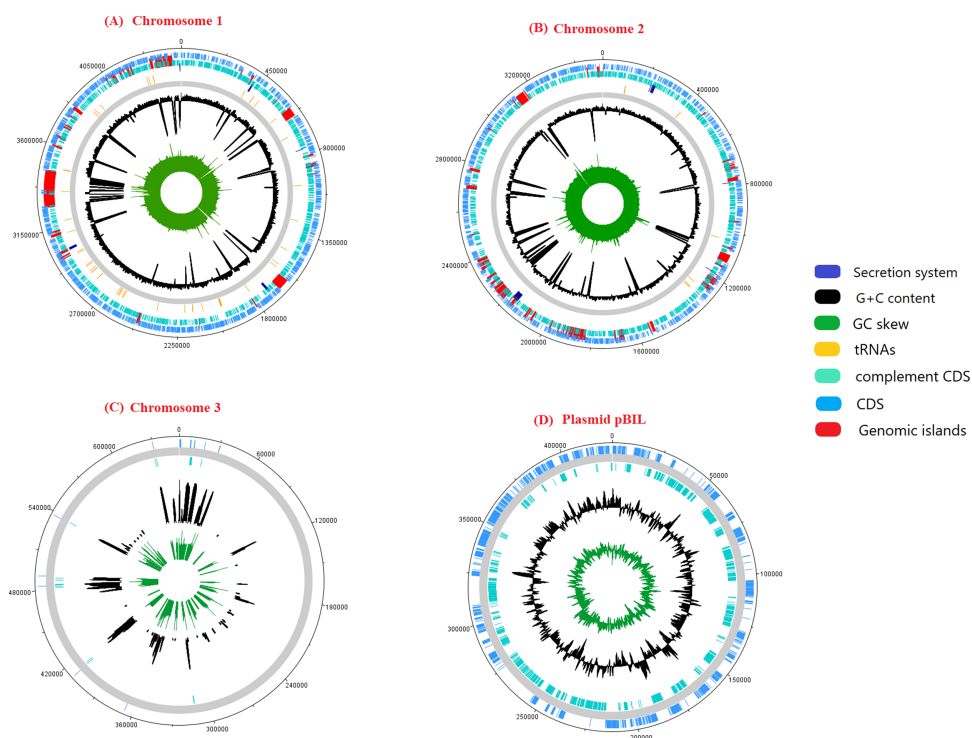


Figure 1 A schematic representation of the genome pangolin Pf. The ring represents chromosomes and plasmid of pangolin Pf, representatively (dark blue, secretion system; red, genomic island; yellow, tRNA; black, G+C content; green, GC skew). The outermost layer and second layer represent the forward and reverse coding region (CDS).

Full-size [DOI: 10.7717/peerj.9733/fig-1](https://doi.org/10.7717/peerj.9733/fig-1)

fungorum ATCC BAA 463 and with a GC content of 62.2%. The pangolin Pf genome had a total number of 222 contigs with an N50 metric of 69,666 bp (Table 1). The genome was plotted using a circular representation by including genomic features such as CDS, tRNA, GC content, and GIs (Fig. 1). When mapping the read sequences back to the assembled genome sequence, they covered 99% of the genome with an average read depth of 29X, suggesting the high quality of the assembled genome sequence and its suitability for downstream analyses.

Confirmation of the taxonomic position of the pangolin Pf

We previously discovered the identity of the pangolin Pf based on the preliminary analyses using two bacterial classification markers such as *rec A* and 16S rRNA genes (Fig. S1). The single marker gene approach may not be able to clearly differentiate all closely related members of Burkholderiales (Jin et al., 2020). Therefore, we used a more robust and reliable whole-genome method or the core-genome SNP approach to further confirm the identity of the pangolin Pf (Fig. 2). The core-genome SNP-based phylogenetic tree classified the pangolin Pf as *P. fungorum* species.

To further validate the taxonomic position of pangolin Pf, the average nucleotide identity (ANI) was calculated based on the method proposed by Goris and his colleagues to measure the relatedness between the members of the Burkholderiales and Paraburkholderiales (Goris

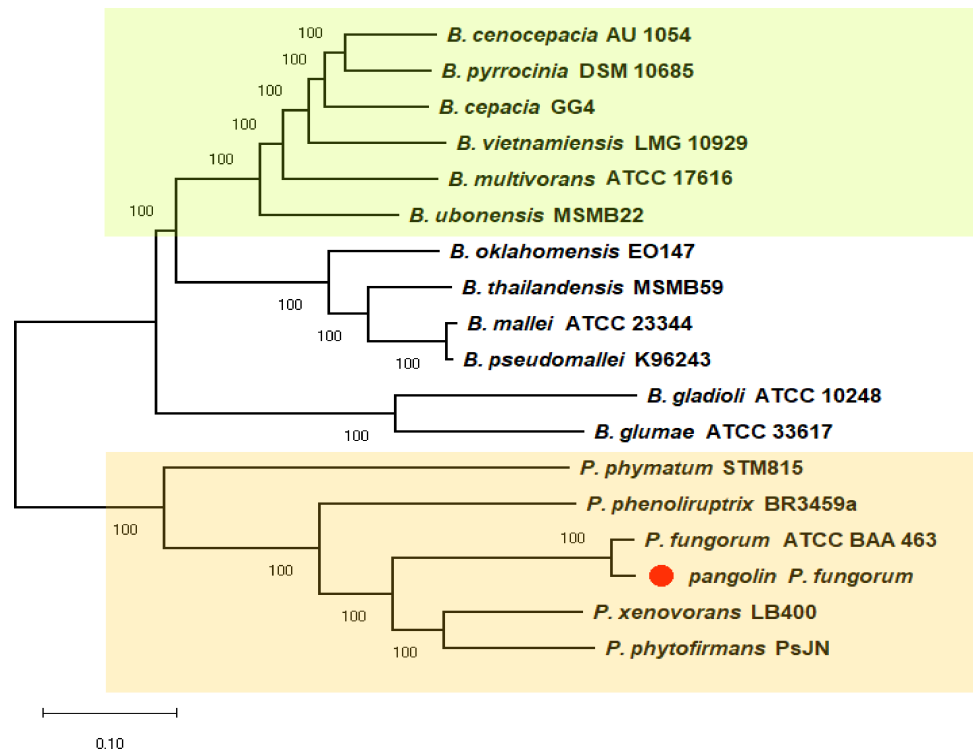


Figure 2 A core-genome SNP-based phylogenetic tree of pangolin Pf and closely related members of the Burkholderiales. The phylogenetic tree was generated using the Neighbour-joining algorithm method. Bootstrap numbers were generated in 1,000 replicates. The pangolin Pf genome sequence was aligned with the genome sequences of 17 other *Burkholderia* and *Paraburkholderia* species and the SNPs located in the core genome (conserved genomic regions among all species) were extracted for alignment and tree reconstruction. The green color part represents the Bcc group while the orange color part represents the Paraburkholderia group.

Full-size DOI: 10.7717/peerj.9733/fig-2

et al., 2007). Using the nucleotide sequence of pangolin Pf as the reference genome, the ANI was calculated by performing a pairwise comparison between the reference genome with the genome of other members of Burkholderiales and Paraburkholderiales (Table 1). ANI results indicated that our pangolin Pf strain is closely related to *P. fungorum* ATCC BAA-463 with an ANI value of 98.49% (passing the threshold of 97% to define a species), whereas other species showed ANI values below the threshold (Fig. 3). Taken together, the results of our ANI and the core-genome SNP-based phylogenetic analyses are consistent and confirm that pangolin Pf is indeed *P. fungorum*.

Functional analysis

RAST predicted 6,746 protein-coding genes and, 82 RNAs (Table S1). As anticipated, a large number of genes were enriched in basic functions such as carbohydrates (746 genes), amino acid and derivative metabolism (718 genes), cofactors, vitamins, prosthetic groups and pigments (401 genes), fatty acids, lipids and isoprenoids (318 genes), protein metabolism (300 genes), membrane transport (213 genes), and RNA metabolism (193 genes), which are important for bacterial survival (Fig. 4). We also observed 196 genes related to stress

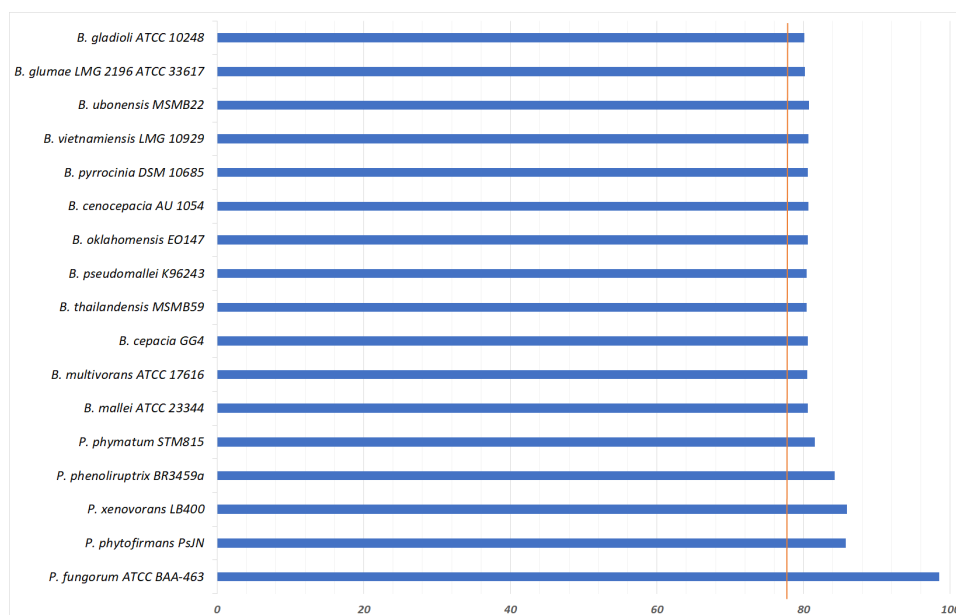


Figure 3 Average Nucleotide Identity statistics.

Full-size DOI: [10.7717/peerj.9733/fig-3](https://doi.org/10.7717/peerj.9733/fig-3)

responses that may play an important role in bacterial adaptation to various environments (Table 2) (Koh & Roe, 1995; Koh & Roe, 1996; May et al., 1989).

Besides that, we identified 158 putative genes related to virulence, disease and defense including resistance to antibiotics and toxic compounds (132 genes), the host cell invasion and intracellular resistance (14 genes) and the ribosomally synthesized antibacterial peptides (12 genes), probably helping bacteria to inhibit the growth of similar or closely related bacterial strain (bacteriocins).

Toxin-antitoxin (TA) systems

The genome of pangolin Pf has two putative toxin-antitoxin (TA) systems; *higAB* and *ygiUT*. The *ygiT* is the transcriptional repressor while *ygiU* has been predicted to be a cyanide hydratase induced upon biofilm formation, acting as a global regulator controlling biofilm formation by inducing motility (Shah et al., 2006). The TA gene has been suggested to play a role in mediating growth arrest during stress situations (Butt et al., 2013; Fivian-Hughes & Davis, 2010). Both *higAB* and *ygiUT* were located closely with each other, forming the toxin-antitoxin system.

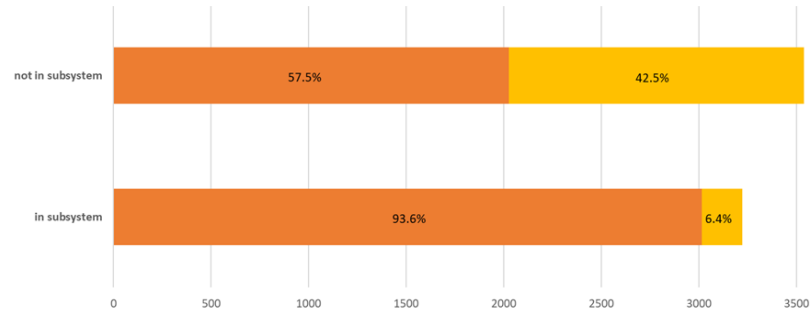
Urease

The pangolin Pf genome has several putative genes important for the urea decomposition and urease subunits such as Urea ABC transporter *urtBCDE*, Urease accessory proteins *ureDEFG*, Urease alpha, beta, and gamma subunits in the genome of pangolin Pf. The Urea ABC transporter *urtBCDE* involves in the uptake of urea and, in response to nitrogen limitation (Beckers et al., 2004). Urease, a nickel containing metalloenzyme, is a virulence factor enabling bacteria to survive by hydrolysing urea as the sole nitrogen source in nutrient

a. Subsystem Coverage



b. Subsystem Coverage Breakdown



c. Subsystem Category Distribution

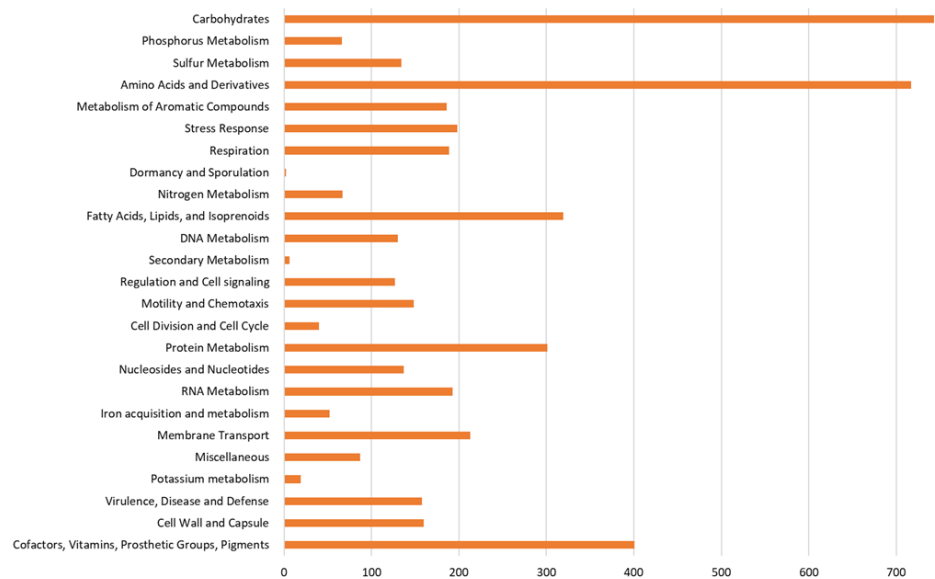


Figure 4 RAST annotation summary. (A) The subsystem coverage; (B) the subsystem coverage breakdown; (C) RAST annotation result showing the distribution of annotation across defined structural and functional subsystem roles in percentage. RAST uses a subsystem approach for annotation as assigning similar functional or structural roles into a group. For pangolin Pf, 48% of annotated genes belong to an identified functional role, or subsystem. The coverage breakdown shows the percentage of hypothetical and non-hypothetical annotations for genes assigned to subsystems and those for which a known functional role was unassigned.

Full-size DOI: [10.7717/peerj.9733/fig-4](https://doi.org/10.7717/peerj.9733/fig-4)

Table 2 A list of RAST-predicted stress response genes in pangolin Pf. Stress response genes were predicted and categorized into heat shock, cold shock, detoxification, osmotic stress, oxidative stress, and periplasmic stress.

Type of Stress Response	Predicted genes
Heat shock	Chaperone protein DnaJ, DnaK, GroEL & GroES
	YggX a Probable Fe(2+)-trafficking protein
	Glutathione synthetase
	Heat shock protein GrpE, YegD, YciM precursor
	Hsp20 family protein
	Heat-inducible transcription repressor HrcA
	Nucleoside 5-triphosphatase RdgB, RpoH, tmRNA-binding protein SmpB, RpoE
Cold shock	Translation elongation factor LepA
	CspDG
Detoxification	DedA protein, Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)
	Various polyols ABC transporter, ATP-binding component, periplasmic substrate-binding protein and permease component 1 & 2
	Betaine aldehyde dehydrogenase (EC 1.2.1.8)
	Choline dehydrogenase (EC 1.1.99.1)
	Choline-sulfatase (EC 3.1.6.6)
	GbcA Glycine betaine demethylase subunit A
	High-affinity choline uptake protein BetT
	HTH-type transcriptional regulator BetI
	L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1), ProW (TC 3.A.1.12.1) & transporter protein ProX (TC 3.A.1.12.1)
	Sarcosine oxidase alpha, beta, delta & gamma subunits (EC 1.5.3.1)
	Integral membrane protein YggT)
	Osmotically inducible proteins C & Y
	<i>proVWX</i> operon
Osmotic stress	Genes related to Glutathione Biosynthesis and gamma-glutamyl cycle
	Glutathione Non-redox reactions;
	Cluster containing Glutathione synthetase; Glutaredoxins
	Folate-dependent protein for Fe/S cluster synthesis/repair in oxidative stress;
	NTP pyrophosphohydrolases including oxidative damage repair enzymes
Oxidative stress	Genes for Paraquat-inducible protein A & B

(continued on next page)

Table 2 (continued)

Type of Stress Response	Predicted genes
Periplasmic stress	HtrA protease/chaperone protein
	Outer membrane protein H precursor; sensor protease DegQ, serine protease
	Sigma factor RpoE negative regulatory protein RseA & RseB precursor
	Survival protein SurA precursor (Peptidyl-prolyl cis-trans isomerase SurA) (EC 5.2.1.8)
	<i>RpoE</i> , <i>RseA</i> and <i>RseB</i> of the <i>rpoE-rseABC</i> operon

limiting conditions to ammonia (Lin et al., 2012). For example, in other Gram-negative bacteria it has been shown that urease enables survival of the bacteria in strong acidic conditions of the stomach by neutralizing gastric acid with the released ammonia and thus playing a major role in the pathogenesis of gastroduodenal diseases (Cussac, Ferrero & Labigne, 1992).

Stress response

There were 196 putative stress response genes identified in the pangolin Pf genome (Table 2). They were mostly related to oxidative stress, osmotic stress, detoxification, heat shock, cold shock, various polyols ABC transporter, ATP-binding component, periplasmic substrate-binding protein, and permease component. A large number of these stress response genes are associated with osmotic stress such as the ones related to Choline and Betaine Uptake and Betaine Biosynthesis. The *proVWX* operon has been reported to be remarkably stimulated in response to hyperosmotic stress encoding for a high affinity transport system for glycine-betaine (Kudva et al., 2016; Price et al., 2008). The glycine-betaine have a role as a cryoprotectant that help bacteria to stabilize their cell membranes and adapt to cold environments (Subramanian et al., 2011).

Several oxidative stress-related genes were also identified in the pangolin Pf genome. For instance, the presence of oxidative stress-related genes for Paraquat-inducible (Pqi) protein A and B (O'Grady & Sokol, 2011). These genes are the members of the *soxRS* regulon (Koh & Roe, 1995) and their expression can increase during the carbon or phosphate starvation in the bacterial stationary phase (Koh & Roe, 1996).

The pangolin Pf also had genes related to periplasmic stress such as *rpoE*, *rseA*, *rseB*, and *htrA* (Table 2). RpoE is an important transcription factor, which functions as effector molecules responding to extracytoplasmic stimuli and is essential for *Burkholderia* to cope with thermal stress (Vanaporn et al., 2008). Besides that, *htrA* is a virulence factor in *B. cenocepacia* important for the growth under the exposure to stress and survival *in vivo* (Flannagan et al., 2007).

Membrane transport

The pangolin Pf genome has putative genes encoding ABC transporters related to membrane transport such as Phosphonate ABC transporter phosphate-binding periplasmic component; Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA, Oligopeptide transport system permease protein OppC & OppB, ATP-binding

protein OppD & OppF; Branched-chain amino acid transport ATP-binding protein LivF & LivG, permease protein LivH & LivM); Dipeptide transport system permease protein DppB & DppC, ATP-binding protein DppD & DppF; Hopanoid-associated RND transporter, HpnN; Phosphonate ABC transporter permease protein phnE. The OppABC functions in the recycling of cell wall peptides (*Goodell & Higgins, 1987*). While the DppBCDF operon encodes an ABC transporter responsible for the utilization of di/tripeptides in *Pseudomonas aeruginosa* (*Pletzer et al., 2014*). The LivH and LivM permeases and the LivG and LivF ATPases are essential, to mediate the transport of these branched-chain amino acids into the cytoplasm (*Adams et al., 1990*). We also identified genes related to the resistance to antibiotics and toxic compounds such as the RND efflux system, the inner membrane transporter CmeB, outer membrane lipoprotein CmeC & NodT family; Cobalt/zinc/cadmium efflux RND transporter, membrane fusion protein, CzcB & CzcC family; Membrane fusion protein of RND family multidrug efflux pump, as well as the genes related to Ton and Tol transport systems in the pangolin Pf genome. The CmeABC gene is crucial to maintain a high-level resistance to fluoroquinolones and contributes significantly to the emergence of fluoroquinolone-resistant mutants (*Yan et al., 2006*). The RND multidrug efflux system is mainly responsible for the intrinsic multidrug resistance in Gram-negative bacteria (*Gugliera et al., 2006*).

Secretion systems

In Gram-negative bacteria, the components of secretion systems are highly specialized macromolecule nanomachines secrete substrates across bacterial inner and outer membranes (*Abby et al., 2016*). The secreted substrates are responsible for a bacterium's response to its environment and in physiological processes such as survival, adhesion, adaptation, and pathogenicity (*Costa et al., 2015; Duangurai, Indrawattana & Pumirat, 2018*). In the pangolin Pf genome, we identified several putative secretion systems which will be discussed below (*Table 3*).

Type II secretion system (T2SS)

Many Gram-negative bacteria use this secretion system as the pathway to translocate proteins from the periplasm across the outer membrane. This system is normally used to secrete a variety of toxins and enzymes by pathogenic bacteria (*Cianciotto & White, 2017*). This system has been suggested to be the key, in making it possible for *Burkholderia Cepacia Complex* to function as opportunistic pathogens (*Somvanshi et al., 2010*). We identified several genes encoding the general secretion pathway proteins (general secretion pathway protein CDEFGHIJKLMN) in the pangolin Pf genome, increasing the possibility that the bacteria species could function as an opportunistic pathogen.

Type III secretion system (T3SS)

This secretion system plays an important role in the secretion of virulence factors by Gram-negative pathogens and the translocation of "effector" proteins into eukaryotic host cells (*Saier, 2006*). There are three classes in T3SSs in which class 1 and class 2 are predicted to mediate interactions with plants, however, class 3 (*bsa* locus) has been implicated in animal pathogenesis (*Teh et al., 2014*). Interestingly, we found the presence of BsaX

Table 3 Secretion systems predicted in the RAST subsystem of the pangolin Pf genome.

Type of secretion system	Predicted genes
Type I Secretion System (T1SS)	Type I secretion outer membrane protein, TolC precursor
Type II Secretion System (T2SS)	Type I secretion system, outer membrane component LapE Secretion pathway protein CDEFGHIJKLMN
Type III Secretion System (T3SS)	Type III secretion bridge between inner and outer membrane lipoprotein (YscJ,HrcJ,EscJ, PscJ) YscU, SpaS, EscU, HrcU, SsaU, homologous to flagellar export components YscT, HrcT, SpaR, EscT, EpaR1, homologous to flagellar export components YscS, homologous to flagellar export components YscR, SpaR, HrcR, EscR, homologous to flagellar export components YscQ, homologous to flagellar export components Type III secretion cytoplasmic protein (YscL)
Type IV Secretion System (T4SS)	Type II/IV secretion system ATP hydrolase TadA/VirB11/CpaF, TadA subfamily Type II/IV secretion system ATPase TadZ/CpaE, associated with Flp pilus assembly
Type V Secretion System (T5SS)	Channel-forming transporter/cytolysins activator of TpsB family; ShlA/HecA/FhaA family Proteins ImpA, ImpB, ImpC, ImpD; Protein of avirulence locus ImpE, ImpF, ImpG/VasA, ImpH/VasB, ImpI/VasC; Type VI secretion lipoprotein/VasD, ImpJ/VasE;
Type VI Secretion System (T6SS)	Outer membrane protein ImpK/VasF, OmpA/MotB domain; Protein phosphatase ImpM; VgrG protein; ClpB protein; IcmF-related protein; Sigma-54 dependent transcriptional regulator/VasH

in the pangolin Pf genome. The *bsa* (or *Burkholderia* Secretion Apparatus) found in *B. pseudomallei* plays an important role in helping the pathogen to survive and replicate in mammalian cells (Haraga et al., 2008). It also plays important roles in invasion, endosome escape and net intracellular replication in cultured cells and in virulence in murine and Syrian hamster models of melioidosis (Burtnick et al., 2008; Stevens et al., 2004; Stevens et al., 2002; Warawa & Woods, 2005).

Type V secretion system (T5SS)

This secretion system encompasses the auto-transporting and two partner systems (Henderson et al., 2004). In the pangolin Pf, a channel forming β -barrel transporter proteins belonging to the TpsB family (Jacob-Dubuisson, Locht & Antoine, 2001) was identified in a genomic island, suggesting that it might be acquired from other source. TpsB plays

an important role as channel for the translocation of the exoproteins across the outer membrane and a specific receptor for TpsA signal secretion.

Type VI secretion system (T6SS)

It appears in a phage-tail-spike-like injectisome that has a potential to introduce effector proteins directly into the cytoplasm of host cells (Filloux, 2013). This secretion system has emerged as virulence factor that may take part in the pathogenic bacterial-host interactions or promote commensal or mutualistic relationships between bacteria and eukaryotes or to mediate cooperative or competitive interactions between bacteria (Jani & Cotter, 2010). For instance, in Gram-negative bacteria, it has been used for the purpose of interbacterial competition in Bcc member (Spiewak et al., 2019), and endure the innate immune response in host (Lennings, West & Schwarz, 2019). We also identified several regulatory genes that form the major requirement for a functional T6SS, such as structural proteins (VasH, VasK, VasF, VasA), effector protein (VgrG), and functional domains (chaperone ClpB, ImpA, ImpB, ImpC, ImpD, ImpF, ImpG, ImpH, ImpJ) in the pangolin Pf genome. The T6SS is comprised of the proteins encoded by *imp* locus and its counterparts which appear to have an important role in pathogen-symbiont host interactions (Records, 2011).

Comparative virulence gene analysis

To have a better understanding and more complete insights into the virulence of pangolin Pf, we performed a comparative virulence gene analysis of the genomes with known pathogenic *Burkholderia* sp. and *Paraburkholderia* sp... Our analysis showed that pangolin Pf shared some common virulence genes with these pathogenic *Burkholderia* strains such as *epsI*, *pilQ*, *bsaX*, *escS*, and *hrcS* (Fig. 5 and Table S2). Most of these predicted genes were associated with secretion systems especially Type III, and VI. The pangolin Pf also showed the presence of Type IV pilli (pilABCDERT) which are involved in adhesion (Piepenbrink & Sundberg, 2016). We found another set of common genes (eg. *bsaX*, *escS*, and *hrcS*) among the strains related to flagella that are involved in motility and invasion. Type III Secretion System is the main secretion system that mediates the secretion of effector molecules directly into host cells in *B. pseudomallei* and *B. mallei* (Beeckman & Vanrompay, 2010). Besides that, we found a T6SS cluster (TssA-TssM) in the pangolin Pf genome, probably playing an important role in bacterial competition (Spiewak et al., 2019).

Genomic Island and prophage analysis

We wondered whether the pangolin Pf had acquired the horizontally transferred Genomic Islands (GIs) over the evolutionary time. GIs are clusters of genes that are inserted into a bacterial genome during a single horizontal gene transfer event, playing important roles in microbial evolution, virulence, drug resistance and/or adaptation to different environments (Zhang et al., 2014). To examine this, we predicted the GIs in the pangolin Pf genome using the IslandViewer4 (Bertelli et al., 2017). We identified putative GIs predicted in the pangolin Pf genome such as harboured genes related to Molybdenum cofactor biosynthesis (MoaD and MoaE), Molybdenum transport proteins (*modABCE*) and Molybdopterin biosynthesis protein *moeA* (Table S3). The molybdopterin biosynthesis pathway is important for the molybdopterin cofactor syntheses, which are required for a

Figure 5 (...continued)

format - heatmap. The selected strains are shown as a horizontal line at the end of the map (y -axis) while the occurring virulence genes across the selected strains shown as a vertical line on the top of the map (x -axis). Red represents the presence of the virulence gene, whereas black represents the absence of virulence gene.

variety of molybdoenzymes such as nitrate reductase that which plays an important role in the denitrification process in bacteria under oxygen limiting conditions (*Leimkühler & Iobbi-Nivol, 2016*). The nitrate reductase and the Molybdopterin biosynthetic pathway have also been associated with bacterial virulence (*Andreae, Titball & Butler, 2014; Filiatrault et al., 2013; Fritz et al., 2002*). In addition, we identified T6SS genes in the GIs of pangolin Pf, suggesting that these genes might have originated from other sources.

There are three incomplete prophages predicted in the pangolin Pf genome, in region 1 with about 7.8 kb in length with 10 coding regions; region 2, 8.2 kb in length with 10 coding regions; and region 3 with 8.8 kb in length and with 6 coding regions (*Fig. S2; Table S4*), suggesting that pangolin Pf might have acquired these prophages a long time ago. No intact prophages or recently evolved complete prophages were detected in the genome. We cannot rule out the possibility that the secretion systems of pangolin Pf might provide a strong defense to the bacterium in preventing the integration of phages into its genome.

DISCUSSION

Here we present the genome sequence and comprehensive analysis of a *Paraburkholderia* species which was discovered from normally sterile tissue of a female pregnant Malayan pangolin. Our analyses showed that the pangolin Pf was supported by evidence from single-gene phylogenetic analyses, the reliable and robust core-genome SNP-based phylogenetic analysis and the ANI analysis. Genome analysis revealed genes involved in stress responses such as cold shock, detoxification, heat shock, and osmotic stress, and oxidative stress, and periplasmic stress. Since *P. fungorum* is considered an environmental bacterial species (*Compant et al., 2008*), these genes may play important roles to help the bacterial species to adapt and survive in a variety of environments.

Interestingly, the pangolin Pf genome also showed a number of genes related to virulence and defence mechanisms such as genes associated with maintenance of cell wall integrity, resistance to antibiotics and toxic compounds, invasion and intracellular resistance, which might make it more resistant to drugs. Moreover, we found that the genome of pangolin Pf consists of the secretion system proteins that may involve in pathogenicity. For instance, T3SSs are important for virulence in many Gram-negative pathogens (*Coburn, Sekirov & Finlay, 2007*). Furthermore, the T6SS has recently emerged as a critical virulence factor for the members of the Burkholderiales including the well-known pathogens *B. pseudomallei*, *B. mallei*, and *B. thailandensis* (*Schwarz et al., 2014*). We cannot rule out the possibility that the presence of the putative secretion systems in the pangolin Pf genome could render this environmental bacterium to a pathogen.

The *Paraburkholderia fungorum* that we studied here was found in an infected female pangolin, which is a placental mammal. Since we found this bacterium in the supposedly sterile tissues of pangolin, it is possible that the *P. fungorum* may have also the ability to colonize and infect humans. Moreover, we found that *P. fungorum* has a set of common virulence genes with the known pathogenic *Burkholderia* spp., raising the possibility that it may have the potential of being an opportunistic pathogen and cause human diseases especially this bacterium has also previously been isolated in the cerebrospinal fluid (Coenye *et al.*, 2001), and the synovial tissue of humans (Loong *et al.*, 2019). Our study raises caution about the use of *P. fungorum* for biocontrol and bioremediation in agriculture. However, more experimental studies are needed to further validate its pathogenic potential.

CONCLUSION

In this study, we report the first comprehensive whole-genome analysis of pangolin Pf that was isolated from a Malayan pangolin. Here our study has provided new insights into the functions, secretion systems and virulence of *P. fungorum*. Besides that, the addition of the genome sequence provides a useful resource for future comparative analysis and functional work of *P. fungorum*.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was funded by the high-level talent recruitment programme for academic and research platform construction (Reference Number: 5000105) from Wenzhou-Kean University and the UM High Impact Research (HIR) grant (Grant Number: UM.C/625/HIR/MOHE/CHAN-08) from the Ministry of Higher Education of Malaysia. Ka Yun Tan was partially supported by the Centre for Research in Biotechnology for Agriculture (CEBAR) grant IRU-MRUN (RU023-2015). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Wenzhou-Kean University: The high-level talent recruitment programme for academic and research platform construction (Reference Number: 5000105).

UM High Impact Research (HIR): UM.C/625/HIR/MOHE/CHAN-08.

The Ministry of Higher Education of Malaysia.

The Centre for Research in Biotechnology for Agriculture (CEBAR): IRU-MRUN (RU023-2015).

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Ka Yun Tan conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Avirup Dutta performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Tze King Tan performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Ranjeev Hari performed the experiments, prepared figures and/or tables, and approved the final draft.
- Rofina Y. Othman analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Siew Woh Choo conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

Data is available at GenBank: [CP028829–CP028832](https://www.ncbi.nlm.nih.gov/GenBank/CP028829-CP028832).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.9733#supplemental-information>.

REFERENCES

- Abby SS, Cury J, Guglielmini J, Néron B, Touchon M, Rocha EP. 2016. Identification of protein secretion systems in bacterial genomes. *Scientific Reports* 6(1):1–14 DOI 10.1038/s41598-016-0001-8.
- Adams MD, Wagner LM, Graddis TJ, Landick R, Antonucci TK, Gibson AL, Oxender DL. 1990. Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of *Escherichia coli*. *Journal of Biological Chemistry* 265(20):11436–11443.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25(17):3389–3402 DOI 10.1093/nar/25.17.3389.
- Andreae CA, Titball RW, Butler CS. 2014. Influence of the molybdenum co-factor biosynthesis on anaerobic respiration, biofilm formation and motility in *Burkholderia thailandensis*. *Research in Microbiology* 165(1):41–49 DOI 10.1016/j.resmic.2013.10.009.
- Andreolli M, Lampis S, Zenaro E, Salkinoja-Salonen M, Vallini G. 2011. *Burkholderia fungorum* DBT1: a promising bacterial strain for bioremediation of PAHs-contaminated soils. *FEMS Microbiology Letters* 319(1):11–18 DOI 10.1111/j.1574-6968.2011.02259.x.

- Ang MY, Dutta A, Wee WY, Dymock D, Paterson IC, Choo SW. 2016. Comparative genome analysis of *Fusobacterium nucleatum*. *Genome Biology and Evolution* 8(9):2928–2938 DOI 10.1093/gbe/evw199.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75 DOI 10.1186/1471-2164-9-75.
- Beckers G, Bendt AK, Kramer R, Burkovski A. 2004. Molecular identification of the urea uptake system and transcriptional analysis of urea transporter- and urease-encoding genes in *Corynebacterium glutamicum*. *Journal of Bacteriology* 186(22):7645–7652 DOI 10.1128/JB.186.22.7645-7652.2004.
- Beeckman DS, Vanrompay DC. 2010. Bacterial secretion systems with an emphasis on the chlamydial Type III secretion system. *Current Issues in Molecular Biology* 12(1):17–41.
- Bertelli C, Laird MR, Williams KP, Group SFURC, Lau BY, Hoad G, Winsor GL, Brinkman FS. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research* 45(W1):W30–W35 DOI 10.1093/nar/gkx343.
- Beukes CW, Palmer M, Manyaka P, Chan WY, Avontuur JR, Zyl Evan, Huntemann M, Clum A, Pillay M, Palaniappan K. 2017. Genome data provides high support for generic boundaries in *Burkholderia sensu lato*. *Frontiers in Microbiology* 8:1154 DOI 10.3389/fmicb.2017.01154.
- Burtneck MN, Brett PJ, Nair V, Warawa JM, Woods DE, Gherardini FC. 2008. *Burkholderia pseudomallei* type III secretion system mutants exhibit delayed vacuolar escape phenotypes in RAW 264.7 murine macrophages. *Infection and Immunity* 76(7):2991–3000 DOI 10.1128/IAI.00263-08.
- Butt A, Mller C, Harmer N, Titball RW. 2013. Identification of type II toxin–antitoxin modules in *Burkholderia pseudomallei*. *FEMS Microbiology Letters* 338(1):86–94 DOI 10.1111/1574-6968.12032.
- Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. 2005. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Research* 33(Database issue):D325–D328 DOI 10.1093/nar/gki177.
- Chen L, Zheng D, Liu B, Yang J, Jin Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Research* 44(D1):D694–D697 DOI 10.1093/nar/gkv1239.
- Choo SW, Wee WY, Ngeow YF, Mitchell W, Tan JL, Wong GJ, Zhao Y, Xiao J. 2014. Genomic reconnaissance of clinical isolates of emerging human pathogen *Mycobacterium abscessus* reveals high evolutionary potential. *Scientific Reports* 4:4061.
- Choo SW, Dutta A, Wong GJ, Wee WY, Ang MY, Siow CC. 2016a. Comparative genomic analysis reveals a possible novel non-tuberculous *Mycobacterium* species with high pathogenic potential. *PLOS ONE* 11(4):e0150413 DOI 10.1371/journal.pone.0150413.

- Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, Wee WY, Yurchenko AA, Kliver S, Tamazian G, Antunes A. 2016b.** Pangolin genomes and the evolution of mammalian scales and immunity. *Genome Research* **26(10)**:1312–1322 DOI [10.1101/gr.203521.115](https://doi.org/10.1101/gr.203521.115).
- Choo SW, Zhou J, Tian X, Zhang S, Qiang S, J. O'Brien S, Tan KY, Platto S, Koepfli K-P, Antunes A, Sitam FT. 2020.** Are pangolins scapegoats of the COVID-19 outbreak-CoV transmission and pathology evidence? *Conservation Letters* In press.
- Cianciotto NP, White RC. 2017.** Expanding role of type II secretion in bacterial pathogenesis and beyond. *Infection and Immunity* **85(5)**:e00014–17.
- Coburn B, Sekirov I, Finlay BB. 2007.** Type III secretion systems and disease. *Clinical Microbiology Reviews* **20(4)**:535–549 DOI [10.1128/CMR.00013-07](https://doi.org/10.1128/CMR.00013-07).
- Coenye T, Laevens S, Willems A, Ohlen M, Hannant W, Govan JR, Gillis M, Falsen E, Vandamme P. 2001.** Burkholderia fungorum sp. nov. and Burkholderia caledonica sp. nov. two new species isolated from the environment, animals and human clinical samples. *International Journal of Systematic and Evolutionary Microbiology* **51(Pt 3)**:1099–2107 DOI [10.1099/00207713-51-3-1099](https://doi.org/10.1099/00207713-51-3-1099).
- Compant S, Nowak J, Coenye T, Clment C, Ait Barka E. 2008.** Diversity and occurrence of Burkholderia spp. in the natural environment. *FEMS Microbiology Reviews* **32(4)**:607–626 DOI [10.1111/j.1574-6976.2008.00113.x](https://doi.org/10.1111/j.1574-6976.2008.00113.x).
- Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. 2015.** Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology* **13(6)**:343–359 DOI [10.1038/nrmicro3456](https://doi.org/10.1038/nrmicro3456).
- Cussac V, Ferrero RL, Labigne A. 1992.** Expression of Helicobacter pylori urease genes in Escherichia coli grown under nitrogen-limiting conditions. *Journal of Bacteriology* **174(8)**:2466–2473 DOI [10.1128/JB.174.8.2466-2473.1992](https://doi.org/10.1128/JB.174.8.2466-2473.1992).
- D'Cruze T, Gong L, Treerat P, Ramm G, Boyce JD, Prescott M, Adler B, Devenish RJ. 2011.** Role for the Burkholderia pseudomallei type three secretion system cluster 1 bpscN gene in virulence. *Infection and Immunity* **79(9)**:3659–3664 DOI [10.1128/IAI.01351-10](https://doi.org/10.1128/IAI.01351-10).
- Duangurai T, Indrawattana N, Pumirat P. 2018.** Burkholderia pseudomallei Adaptation for Survival in Stressful Conditions. *BioMed Research International* **2018**:3039106.
- Estrada-de Los Santos P, Palmer M, Chvez-Ramrez B, Beukes C, Steenkamp ET, Briscoe L, Khan N, Maluk M, Lafos M, Humm E. 2018.** Whole genome analyses suggests that Burkholderia sensu lato contains two additional novel genera (Mycetohabitans gen. nov. and Trinickia gen. nov.): implications for the evolution of diazotrophy and nodulation in the Burkholderiaceae. *Gene* **9(8)**:389 DOI [10.3390/genes9080389](https://doi.org/10.3390/genes9080389).
- Estrada-delos Santos P, Vinuesa P, Martinez-Aguilar L, Hirsch AM, Caballero-Mellado J. 2013.** Phylogenetic analysis of burkholderia species by multilocus sequence analysis. *Current Microbiology* **67(1)**:51–60 DOI [10.1007/s00284-013-0330-9](https://doi.org/10.1007/s00284-013-0330-9).
- Filiatrault MJ, Tomblin G, Wagner VE, Van Alst N, Rumbaugh K, Sokol P, Schwingel J, Iglewski BH. 2013.** Pseudomonas aeruginosa PA1006, which plays a role in

- molybdenum homeostasis, is required for nitrate utilization, biofilm formation, and virulence. *PLoS ONE* **8**(2):e55594 DOI [10.1371/journal.pone.0055594](https://doi.org/10.1371/journal.pone.0055594).
- Filloux A. 2013.** The rise of the Type VI secretion system. *F1000prime Reports* **5**:52 DOI [10.12703/P5-52](https://doi.org/10.12703/P5-52).
- Fivian-Hughes AS, Davis EO. 2010.** Analyzing the regulatory role of the HigA antitoxin within *Mycobacterium tuberculosis*. *Journal of Bacteriology* **192**(17):4348–4356 DOI [10.1128/JB.00454-10](https://doi.org/10.1128/JB.00454-10).
- Flannagan RS, Aubert D, Kooi C, Sokol PA, Valvano MA. 2007.** *Burkholderia cenocepacia* requires a periplasmic HtrA protease for growth under thermal and osmotic stress and for survival in vivo. *Infection and Immunity* **75**(4):1679–1689 DOI [10.1128/IAI.01581-06](https://doi.org/10.1128/IAI.01581-06).
- Fritz C, Maass S, Kreft A, Bange FC. 2002.** Dependence of *Mycobacterium bovis* BCG on anaerobic nitrate reductase for persistence is tissue specific. *Infection and Immunity* **70**(1):286–291 DOI [10.1128/IAI.70.1.286-291.2002](https://doi.org/10.1128/IAI.70.1.286-291.2002).
- Gerrits GP, Klaassen C, Coenye T, Vandamme P, Meis JF. 2005.** *Burkholderia fungorum* septicemia. *Emerging Infectious Diseases* **11**(7):1115 DOI [10.3201/eid1107.041290](https://doi.org/10.3201/eid1107.041290).
- Goodell EW, Higgins CF. 1987.** Uptake of cell wall peptides by *Salmonella typhimurium* and *Escherichia coli*. *Journal of Bacteriology* **169**(8):3861–3865 DOI [10.1128/JB.169.8.3861-3865.1987](https://doi.org/10.1128/JB.169.8.3861-3865.1987).
- Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007.** DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology* **57**(Pt 1):81–91 DOI [10.1099/ijs.0.64483-0](https://doi.org/10.1099/ijs.0.64483-0).
- Gugliera P, Pasca MR, De Rossi E, Buroni S, Arrigo P, Manina G, Riccardi G. 2006.** Efflux pump genes of the resistance-nodulation-division family in *Burkholderia cenocepacia* genome. *BMC Microbiology* **6**(1):66 DOI [10.1186/1471-2180-6-66](https://doi.org/10.1186/1471-2180-6-66).
- Haraga A, West TE, Brittnacher MJ, Skerrett SJ, Miller SI. 2008.** *Burkholderia thailandensis* as a Model System for the Study of the Virulence-Associated Type III Secretion System of *Burkholderia pseudomallei*. *Infection and Immunity* **76**(11):5402–5411 DOI [10.1128/IAI.00626-08](https://doi.org/10.1128/IAI.00626-08).
- Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D. 2004.** Type V protein secretion pathway: the autotransporter story. *Microbiology and Molecular Biology Reviews* **68**(4):692–744 DOI [10.1128/MMBR.68.4.692-744.2004](https://doi.org/10.1128/MMBR.68.4.692-744.2004).
- Jacob-Dubuisson F, Locht C, Antoine R. 2001.** Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. *Molecular Microbiology* **40**(2):306–313 DOI [10.1046/j.1365-2958.2001.02278.x](https://doi.org/10.1046/j.1365-2958.2001.02278.x).
- Jani AJ, Cotter PA. 2010.** Type VI secretion: not just for pathogenesis anymore. *Cell Host & Microbe* **8**(1):2–6 DOI [10.1016/j.chom.2010.06.012](https://doi.org/10.1016/j.chom.2010.06.012).
- Jin Y, Zhou J, Zhou J, Hu M, Zhang Q, Kong N, Ren H, Liang L, Yue J. 2020.** Genome-based classification of *Burkholderia cepacia* complex provides new insight into its taxonomic status. *Biology Direct* **15**(1):1–14 DOI [10.1186/s13062-019-0257-6](https://doi.org/10.1186/s13062-019-0257-6).
- Jones A, Dodd M, Webb A. 2001.** *Burkholderia cepacia*: current clinical issues, environmental controversies and ethical dilemmas. *The European Respiratory Journal* **17**(2):295–301 DOI [10.1183/09031936.01.17202950](https://doi.org/10.1183/09031936.01.17202950).

- Kaur C, Selvakumar G, Ganeshamurthy AN. 2017.** Burkholderia to Paraburkholderia: The Journey of a Plant-Beneficial-Environmental Bacterium. In: Shukla P, ed. *Recent advances in Applied Microbiology*. Singapore: Springer, 213–228.
- Koh YS, Roe JH. 1995.** Isolation of a novel paraquat-inducible (pqi) gene regulated by the soxRS locus in Escherichia coli. *Journal of Bacteriology* **177**(10):2673–2678 DOI [10.1128/JB.177.10.2673-2678.1995](https://doi.org/10.1128/JB.177.10.2673-2678.1995).
- Koh YS, Roe JH. 1996.** Dual regulation of the paraquat-inducible gene pqi-5 by SoxS and RpoS in Escherichia coli. *Molecular Microbiology* **22**(1):53–61 DOI [10.1111/j.1365-2958.1996.tb02655.x](https://doi.org/10.1111/j.1365-2958.1996.tb02655.x).
- Kudva TI, Crnick AN, Plummer JP, Zhang Q, Nicholson LT, Bannantine PJ, Bellaire HB. 2016.** *Virulence mechanisms of bacterial pathogens, Fifth Edition*. Washington, D.C.: ASM Press.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018.** MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**(6):1547–1549 DOI [10.1093/molbev/msy096](https://doi.org/10.1093/molbev/msy096).
- Laing C, Buchanan C, Taboada EN, Zhang Y, Kropinski A, Villegas A, Thomas JE, Gannon VP. 2010.** Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions. *BMC Bioinformatics* **11**:461 DOI [10.1186/1471-2105-11-461](https://doi.org/10.1186/1471-2105-11-461).
- Landers P, Kerr KG, Rowbotham TJ, Tipper JL, Keig PM, Ingham E, Denton M. 2000.** Survival and growth of Burkholderia cepacia within the free-living amoeba Acanthamoeba polyphaga. *European Journal of Clinical Microbiology and Infectious Diseases* **19**(2):121–123 DOI [10.1007/s100960050442](https://doi.org/10.1007/s100960050442).
- Langille MG, Brinkman FS. 2009.** IslandViewer: an integrated interface for computational identification and visualization of genomic islands. *Bioinformatics* **25**(5):664–665 DOI [10.1093/bioinformatics/btp030](https://doi.org/10.1093/bioinformatics/btp030).
- Leimkühler S, Iobbi-Nivol C. 2016.** Bacterial molybdoenzymes: old enzymes for new purposes. *FEMS Microbiology Reviews* **40**(1):1–18 DOI [10.1093/femsre/fuv043](https://doi.org/10.1093/femsre/fuv043).
- Leitão JH, Feliciano JR, Sousa SA, Pita T, Guerreiro SI. 2017.** Burkholderia cepacia complex infections among cystic fibrosis patients: perspectives and challenges. *Progress in Understanding Cystic Fibrosis*.
- Lennings J, West TE, Schwarz S. 2019.** The Burkholderia type VI secretion system 5: composition, regulation and role in virulence. *Frontiers in Microbiology* **9**:3339 DOI [10.3389/fmicb.2018.03339](https://doi.org/10.3389/fmicb.2018.03339).
- Limmathurotsakul D, Golding N, Dance DA, Messina JP, Pigott DM, Moyes CL, Rolim DB, Bertherat E, Day NP, Peacock SJ. 2016.** Predicted global distribution of Burkholderia pseudomallei and burden of melioidosis. In: *Nature microbiology*. vol. 1. London: InTechOpen, 15008 DOI [10.1038/nmicrobiol.2015.8](https://doi.org/10.1038/nmicrobiol.2015.8).
- Lin W, Mathys V, Ang EL, Koh VH, Martinez Gomez JM, Ang ML, Zainul Rahim SZ, Tan MP, Pethe K, Alonso S. 2012.** Urease activity represents an alternative pathway for Mycobacterium tuberculosis nitrogen metabolism. *Infection and Immunity* **80**(8):2771–2779 DOI [10.1128/IAI.06195-11](https://doi.org/10.1128/IAI.06195-11).

- Loong SK, Tan K-K, Zulkifle N-I, AbuBakar S. 2019.** Draft genome of Paraburkholderia fungorum sequence type 868 recovered from human synovial tissues. *Data in Brief* 25:104159 DOI 10.1016/j.dib.2019.104159.
- May G, Faatz E, Lucht JM, Haardt M, Bolliger M, Bremer E. 1989.** Characterization of the osmoregulated Escherichia coli proU promoter and identification of ProV as a membrane-associated protein. *Molecular Microbiology* 3(11):1521–1531 DOI 10.1111/j.1365-2958.1989.tb00138.x.
- Morya R, Salvachúa D, Thakur IS. 2020.** Burkholderia: an untapped but promising bacterial genus for the conversion of aromatic compounds. *Trends in Biotechnology* DOI 10.1016/j.tibtech.2020.02.008.
- Nally E, Groah SL, Prez-Losada M, Caldovic L, Ljungberg I, Chandel NJ, Sprague B, Hsieh MH, Pohl HG. 2018.** Identification of Burkholderia fungorum in the urine of an individual with spinal cord injury and augmentation cystoplasty using 16S sequencing: copathogen or innocent bystander? *Spinal Cord Series and Cases* 4(1):85 DOI 10.1038/s41394-018-0115-2.
- O’Grady EP, Sokol PA. 2011.** Burkholderia cenocepacia differential gene expression during host–pathogen interactions and adaptation to the host environment. *Frontiers in Cellular and Infection Microbiology* 1:15.
- Piepenbrink KH, Sundberg EJ. 2016.** Motility and adhesion through type IV pili in Gram-positive bacteria. *Biochemical Society Transactions* 44(6):1659–1666 DOI 10.1042/BST20160221.
- Pletzer D, Lafon C, Braun Y, Kohler T, Page MG, Mourez M, Weingart H. 2014.** High-throughput screening of dipeptide utilization mediated by the ABC transporter Dpp-BCDF and its substrate-binding proteins DppA1-A5 in Pseudomonas aeruginosa. *PLOS ONE* 9(10):e111311 DOI 10.1371/journal.pone.0111311.
- Price CT, Bukka A, Cynamon M, Graham JE. 2008.** Glycine betaine uptake by the ProX-VWZ ABC transporter contributes to the ability of Mycobacterium tuberculosis to initiate growth in human macrophages. *Journal of Bacteriology* 190(11):3955–3961 DOI 10.1128/JB.01476-07.
- Pruitt KD, Tatusova T, Maglott DR. 2007.** NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research* 35(Database issue):D61–D65 DOI 10.1093/nar/gkl842.
- Records AR. 2011.** The type VI secretion system: a multipurpose delivery system with a phage-like machinery. *Molecular Plant-Microbe Interactions* 24(7):751–757 DOI 10.1094/MPMI-11-10-0262.
- Rodriguez RLM, Konstantinidis KT. 2014.** Bypassing cultivation to identify bacterial species. *Microbe* 9(3):111–118.
- Rojas-Rojas FU, Lopez-Sanchez D, Meza-Radilla G, Mendez-Canarios A, Ibarra J, Estrada-de PLS. 2019.** The controversial Burkholderia cepacia complex, a group of plant growth promoting species and plant, animals and human pathogens. *Revista Argentina de Microbiologia* 51(1):84–92 DOI 10.1016/j.ram.2018.01.002.
- Saier M. 2006.** Protein secretion systems in Gram-negative bacteria. *Microbe-American Society for Microbiology* 1(9):414.

- Sawana A, Adeolu M, Gupta RS. 2014.** Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. *Frontiers in Genetics* 5:429.
- Schwarz S, Singh P, Robertson JD, LeRoux M, Skerrett SJ, Goodlett DR, West TE, Mougous JD. 2014.** VgrG-5 is a *Burkholderia* type VI secretion system-exported protein required for multinucleated giant cell formation and virulence. *Infection and Immunity* 82(4):1445–1452 DOI 10.1128/IAI.01368-13.
- Shah D, Zhang Z, Khodursky A, Kaldalu N, Kurg K, Lewis K. 2006.** Persisters: a distinct physiological state of *E. coli*. *BMC Microbiology* 6:53 DOI 10.1186/1471-2180-6-53.
- Somvanshi VS, Viswanathan P, Jacobs JL, Mulks MH, Sundin GW, Ciche TA. 2010.** The Type 2 Secretion Pseudopilin, gspJ, is required for multihost pathogenicity of *Burkholderia cenocepacia* AU1054. *Infection and Immunity* 78(10):4110–4121 DOI 10.1128/IAI.00558-10.
- Spiewak HL, Shastri S, Zhang L, Schwager S, Eberl L, Vergunst AC, Thomas MS. 2019.** *Burkholderia cenocepacia* utilizes a type VI secretion system for bacterial competition. *MicrobiologyOpen* 8(7):e00774.
- Stevens MP, Wood MW, Taylor LA, Monaghan P, Hawes P, Jones PW, Wallis TS, Galyov EE. 2002.** An Inv/Mxi-Spa-like type III protein secretion system in *Burkholderia pseudomallei* modulates intracellular behaviour of the pathogen. *Molecular Microbiology* 46(3):649–659 DOI 10.1046/j.1365-2958.2002.03190.x.
- Stevens MP, Haque A, Atkins T, Hill J, Wood MW, Easton A, Nelson M, Underwood-Fowler C, Titball RW, Bancroft GJ, Galyov EE. 2004.** Attenuated virulence and protective efficacy of a *Burkholderia pseudomallei* bsa type III secretion mutant in murine models of melioidosis. *Microbiology* 150(Pt 8):2669–2676 DOI 10.1099/mic.0.27146-0.
- Subramanian P, Joe MM, Yim W, Hong B, Tipayno SC, Saravanan VS, Yoo J, Chung J, Sultana T, Sa T. 2011.** Psychrotolerance mechanisms in cold-adapted bacteria and their perspectives as plant growth-promoting bacteria in temperate agriculture. *Korean Journal of Soil Science and Fertilizer* 44(4):625–636.
- Tan SY, Tan IKP, Tan MF, Dutta A, Choo SW. 2016a.** Evolutionary study of *Yersinia* genomes deciphers emergence of human pathogenic species. *Scientific Reports* 6:36116 DOI 10.1038/srep36116.
- Tan TK, Tan KY, Hari R, Mohamed Yusoff A, Wong GJ, Siow CC, Mutha NV, Rayko M, Komissarov A. 2016b.** PGD: a pangolin genome hub for the research community. Database (Oxford) DOI 10.1093/database/baw063.
- Teh BE, French CT, Chen YH, Chen IGJ, Wu TH, Sagullo E, Chiou PY, Teitell MA, Miller JF, Gan YH. 2014.** Type three secretion system-mediated escape of *Burkholderia pseudomallei* into the host cytosol is critical for the activation of NF kappa B. *BMC Microbiology* 14:115 DOI 10.1186/1471-2180-14-115.
- Vanaporn M, Vattanaviboon P, Thongboonkerd V, Korbsrisate S. 2008.** The *rpoE* operon regulates heat stress response in *Burkholderia pseudomallei*. *FEMS Microbiology Letters* 284(2):191–196 DOI 10.1111/j.1574-6968.2008.01216.x.

- Vignal A, Milan D, SanCristobal M, Eggen A. 2002.** A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* **34(3)**:275–305 DOI [10.1186/1297-9686-34-3-275](https://doi.org/10.1186/1297-9686-34-3-275).
- Warawa J, Woods DE. 2005.** Type III secretion system cluster 3 is required for maximal virulence of *Burkholderia pseudomallei* in a hamster infection model. *FEMS Microbiology Letters* **242(1)**:101–108 DOI [10.1016/j.femsle.2004.10.045](https://doi.org/10.1016/j.femsle.2004.10.045).
- Whitlock GC, Estes DM, Torres AG. 2007.** Glanders: off to the races with *Burkholderia mallei*. *FEMS Microbiology Letters* **277(2)**:115–122 DOI [10.1111/j.1574-6968.2007.00949.x](https://doi.org/10.1111/j.1574-6968.2007.00949.x).
- Yan M, Sahin O, Lin J, Zhang Q. 2006.** Role of the CmeABC efflux pump in the emergence of fluoroquinolone-resistant *Campylobacter* under selection pressure. *Journal of Antimicrobial Chemotherapy* **58(6)**:1154–1159 DOI [10.1093/jac/dkl412](https://doi.org/10.1093/jac/dkl412).
- Zhang R, Ran Y, Dai Y, Yang H, Zhang H, Lu Y, Wang L. 2014.** Infectious granuloma caused by *Burkholderia fungorum* confirmed by laser-capture microdissection and polymerase chain reaction. *British Journal of Dermatology* **171(5)**:1261–1263 DOI [10.1111/bjd.13094](https://doi.org/10.1111/bjd.13094).
- Zheng W, Tan MF, Old LA, Paterson IC, Jakubovics NS, Choo SW. 2017.** Distinct biological potential of *Streptococcus gordonii* and *Streptococcus sanguinis* revealed by comparative genome analysis. *Scientific Reports* **7(1)**:2949 DOI [10.1038/s41598-017-02399-4](https://doi.org/10.1038/s41598-017-02399-4).
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011.** PHAST: a fast phage search tool. *Nucleic Acids Research* **39(Web Server issue)**:W347–W352 DOI [10.1093/nar/gkr485](https://doi.org/10.1093/nar/gkr485).