Genome sequence analysis of archival Malayan pangolin (*Manis javanica*) tissues reveals the presence of *Paraburkholderia fungorum* sequences (#81047)

First submission

Guidance from your Editor

Please submit by **16 Feb 2023** for the benefit of the authors (and your token reward) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

8 Figure file(s)
2 Table file(s)
3 Other file(s)

Custom checks

Vertebrate animal usage checks

- Have you checked the authors ethical approval statement?
- Were the experiments necessary and ethical?
- Have you checked our animal research policies?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Genome sequence analysis of archival Malayan pangolin (*Manis javanica*) tissues reveals the presence of *Paraburkholderia fungorum* sequences

Ka Yun Tan [Equal first author, 1], **Siwei Deng** [Equal first author, 2], **Tze King Tan** [3], **Ranjeev Hari** [1], **Rofina Yasmin Othman** [4], **Kum Thong Wong** [5], **Taznim Begam Mohd Mohidin** [1], **Siew Woh Choo** [Corresp. 2, 6, 7]

[1] Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia

[2] College of Science and Technology, Wenzhou-Kean University, Wenzhou, Zhejiang, China

[3] Cancer Science Institute of Singapore, National University of Singapore, Singapore, Singapore

[4] Centre for Research in Biotechnology for Agriculture (CEBAR), University of Malaya, Kuala Lumpur, Malaysia

[5] Department of Pathology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

[6] Zhejiang Bioinformatics International Science and Technology Cooperation Center, Wenzhou, Zhejiang, China

[7] Wenzhou Municipal Key Laboratory for Applied Biomedical and Biopharmaceutical Informatics, Wenzhou, Zhejiang, China

Corresponding Author: Siew Woh Choo
Email address: cwoh@wku.edu.cn

Background. Malayan pangolin (*Manis javanica*) is a placental mammal and *Critically Endangered* ~~species~~. Previous efforts have attempted to breed pangolins in captivity, but with little success because they mainly died of infections. In our previous pangolin genome sequencing data analysis, we found a considerable amount of bacterial DNA in a pregnant female Malayan pangolin (named “UM3”), which was likely infected by *Paraburkholderia fungorum* – an agent of biodegradation and bioremediation in agriculture.

Methodology. Here, we further confirmed and characterized this bacterial species using PCR, histological staining, whole-genome sequencing, and bioinformatics approaches. PCR assays with in-house designed primer sets and 16S universal primers showed clear positive bands in the tissues of UM3 such as cerebrum, cerebellum, lung, and blood, suggesting that UM3 might have developed septicaemia. Histological staining showed the presence of Gram-negative rod-shaped bacteria in the pangolin brain and lungs, indicating the colonization of the bacteria in these two organs. In addition, PCR screening of ~~the~~ UM3’s fetus tissues revealed the presence of *P. fungorum* in the gastrocnemius muscle, but not in other tissues that we examined. We ~~have~~ also sequenced and reconstructed the genome of pangolin *P. fungorum* ~~with~~ a genome size of 7.7 Mbps.

Conclusion. Our study first ~~time~~ report~~s~~ the presence of *P. fungorum* in a pregnant mammalian pangolin species and a fetus. Here, we raise the concern that *P. fungorum* may potentially infect humans, especially YOPI (young, old, pregnant, and immunocompromised) people. Therefore, ~~it should be cautious~~ in using this bacterial species as biodegradation or bioremediation agents in agriculture. Furthermore, our study implicates the importance of improving the current conservation and breeding strategies of this threatened and immunologically fragile mammal.

Genome sequence analysis of archival Malayan pangolin (*Manis Javanica*) tissues reveals the presence of *Paraburkholderia fungorum* sequences

Ka Yun Tan[1*], Siwei Deng[2*], Tze King Tan[3], Ranjeev Hari[1], Rofina Yasmin Othman[4], Kum Thong Wong[5], Taznim Begam Mohd Mohidin[1], Siew Woh Choo[2,6,7#]

[1]Institute of Biological Sciences, Faculty of Science, Universiti Malaya, Jalan Profesor Diraja Ungku Aziz, Wilayah Persekutuan Kuala Lumpur, Malaysia
[2]College of Science and Technology, Wenzhou-Kean University, 88 Daxue Road, Ouhai, Wenzhou, Zhejiang Province, China
[3]Cancer Science Institute of Singapore, National University of Singapore, 117599 Singapore.
[4]Centre for Research in Biotechnology for Agriculture (CEBAR), Level 3, Research Management & Innovation Complex, University of Malaya, 50603 Kuala Lumpur, Malaysia
[5]Department of Pathology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, W. Persekutuan Kuala Lumpur, Malaysia
[6]Zhejiang Bioinformatics International Science and Technology Cooperation Center, Ouhai, Wenzhou, Zhejiang Province, 325060 China
[7]Wenzhou Municipal Key Laboratory for Applied Biomedical and Biopharmaceutical Informatics, Ouhai, Wenzhou, Zhejiang Province, 325060 China

*= Both contribute equally.

[#] =Corresponding Author:
Choo Siew Woh
Email: cwoh@wku.edu.cn

Abstract

Background. Malayan pangolin (*Manis javanica*) is a placental mammal and *Critically Endangered* species. Previous efforts have attempted to breed pangolins in captivity, but with little success because they mainly died of infections. In our previous pangolin genome sequencing data analysis, we found a considerable amount of bacterial DNA in a pregnant female Malayan pangolin (named "UM3"), which was likely infected by *Paraburkholderia fungorum* – an agent of biodegradation and bioremediation in agriculture.

Methodology. Here, we further confirmed and characterized this bacterial species using PCR, histological staining, whole-genome sequencing, and bioinformatics approaches. PCR assays with in-house designed primer sets and 16S universal primers showed clear positive bands in the tissues of UM3 such as cerebrum, cerebellum, lung, and blood, suggesting that UM3 might have developed septicaemia. Histological staining showed the presence of Gram-negative rod-shaped bacteria in the pangolin brain and lungs, indicating the colonization of the bacteria in these two organs. In addition, PCR screening of the UM3's fetus tissues revealed the presence of *P. fungorum* in the gastrocnemius muscle, but not in other tissues that we examined. We have also sequenced and reconstructed the genome of pangolin *P. fungorum* with a genome size of 7.7 Mbps.

Conclusion. Our study first-time reports the presence of *P. fungorum* in a pregnant mammalian pangolin species and a fetus. Here, we raise the concern that *P. fungorum* may potentially infect humans, especially YOPI (young, old, pregnant, and immunocompromised) people. Therefore, it should be cautious in using this bacterial species as biodegradation or bioremediation agents in agriculture. Furthermore, our study implicates the importance of improving the current conservation and breeding strategies of this threatened and immunologically fragile mammal.

Keywords

pangolin conservation; bacteria genome analysis; microbial infection; fetal infection

Introduction

Pangolins are unique terrestrial mammals with special physical traits such as scaly, no teeth, poor vision, and strong-smelling sense (1). Pangolins are extremely difficult to maintain and breed in

captivity (2, 3) mainly because they frequently die of infections such as gastrointestinal disease, pneumonia, skin disease, and parasitic infections (3).

Burkholderia genus belongs to β-Proteobacteria – a Gram-negative, aerobic, rod-shaped bacteria associated with lethal human diseases (4). They are widely used in agriculture because they can fix nitrogen, promote plant growth, and degrade recalcitrant chemical compounds (4). They have been further categorized into *Burkholderia* and *Paraburkholderia* – the former being an animal and plant pathogen, and the latter being environmental and plant beneficial species (5-7). Among the well-known members of *Burkholderia* spp. are *B. pseudomallei* and *B. mallei* (have been used as bioweapons in wars), as well as *Paraburkholderia fungorum* which is a soil bacterium usually isolated from diverse ecological niches (8, 9). *P. fungorum* is commonly used in agriculture as a biodegradation and bioremediation agent (10, 11). However, there have been reports of the isolation of *P. fungorum* from mouse nose and cystic fibrosis patients (8). Therefore, there are debates on the suitability of *P. fungorum* in agricultural use because some people believe that it would affect human health, although no clear evidence has been reported (12).

Here, we report a case of *P. fungorum* infection in a pregnant female pangolin (named "UM3") and its fetus, supported by evidence from PCR assays, histological analysis, whole-genome analysis, and phylogenetic analysis. Our PCR and whole-genome sequencing results also showed the presence of this bacterial species in its fetal muscle, suggesting that *P. fungorum* may also have the capability to colonize in the fetus of UM3.

Methods

Ethics statement

Veterinary officers conducted all procedures involving animals and experts at the Department of Wildlife and National Parks (DWNP), Malaysia, following internationally recognized guidelines and approved by the University of Malaya Institutional Animal Care and Use Committee (UM IACUC) [reference number of the approval: DRTU/11/10/2013/RH (R)].

Biological samples

The samples of a female pregnant pangolin (called "UM3") weighing 2.73 kg were provided by the DWNP. All tissues of UM3 (including cerebellum, cerebrum, lungs, thymus, liver, blood,

heart, and its fetus's gastrocnemius muscle) were harvested by DWNP veterinary officers before being flash-frozen using liquid nitrogen and stored at −80 °C.

Genomic DNA extraction

The fetal muscle tissue was used for genomic DNA extraction. The extraction was done using Qiagen Genomic Tips 20/G kits according to the manufacturer's protocol.

Library preparation and sequencing

Libraries for genomic DNA obtained from fetal muscle were prepared with a fragment length of approximately 300 bp and sequenced using Illumina HiSeq 2000 following vendor'ssequencing protocol.

Discovery of bacterial sequences in the pangolin genome

The pangolin genome was previously sequenced and assembled by our group (13). The tissue-specific genome assemblies were generated by CLC Assembly Cell using sequencing data from pangolin brain (cerebrum and cerebellum), liver, and lung samples. The sequencing reads were searched against a bacterial nucleotide sequence database using BLASTN (14). We screened the bacterial identity using two criteria: 90% sequence identity and 90% sequence coverage and 97% sequence identity and 97% sequence coverage.

Average nucleotide identity (ANI) and average amino acid identity (AAI) analyses

The average nucleotide identity (ANI) values between species were calculated based on the method previously described (15-18). We used two-way BLAST and only used the forward and reversed-matched orthologs in the calculations. For robustness, the BLAST hits were filtered for at least 50% identity at the nucleotide and amino acid level and a sequence coverage of at least 70%.

The protein sequences of 18 genomes belonging to the *Paraburkholderia* and *Burkholderia* genera were annotated using Rapid Annotation Search Tool (RAST) (19). The RAST-predicted protein sequences for each assembly were retrieved, and average amino acid identity (AAI) values were calculated using the AAI calculator (15, 16).

PCR assays

To further validate the presence of the bacterial sequences in the pangolin, the frozen tissue samples of pangolin UM3 were examined. gDNA extracted from nine adult tissues (cerebrum, cerebellum, liver, lungs, heart, spleen, thymus, skin, and blood) and the fetal tissues were screened using polymerase chain reaction (PCR) assays. Three different target genomic regions that showed top hits to the bacteria identified from the previous BLASTN results were selected to design and synthesize ~~the~~ PCR primers ~~provided by FirstBase company~~. ~~The~~ bacterial universal 16S primers and three in-house designed primer sets (Table S1), targeting bacterial 16S rRNA, *Burkholderia*-specific transposase genomic region, OI25_7129 hypothetical protein genomic region, and *P. fungorum*-specific DNA polymerase genomic region ~~has been used~~ respectively.

All PCR assays were performed using ~~the following described protocol. The~~ total reaction volume of 50 µL ~~contained~~ 160 ng purified organ gDNA, 0.3 mol of each primer, deoxynucleotides triphosphates (dNTP, 400 µM each), 1.0 U Taq DNA polymerase and a supplied buffer ~~were used~~. The PCR was performed as follows: one cycle (94 ºC for two minutes) for initial denaturation; 35 cycles (98 ºC for 10 ~~secs~~; 68 ºC for three minutes) for annealing and DNA amplification. The PCR products were purified by standard methods and directly sequenced with the same primers using BigDye© Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) for validation.

Tissue preparation and histological staining

We examined the histology of the cerebellum and lungs. Each of the thawed organs was excised into two sets of smaller tissue pieces and fixed ~~with~~ 10% formalin at 12°C for a week, followed by embedding in paraffin wax to produce paraffin blocks. For histology, the tissue blocks were sectioned on a rotary microtome (Leica RM2235, Leica Biosystems) with a blade of 3 µm. To prevent cross contamination, the blades were cleaned with ~~100~~% ethanol between sections. Subsequently, the slices were dewaxed using graded alcohol. ~~These tissue~~ slices were separately counterstained using hematoxilin/eosin (HE, Sigma) for tissue abnormality such as inflammation, and Brown-Hopps Gram stains for bacterial presence, as described by (20). Slices were ~~observed~~ under a light microscope with a Leica DF300 camera (Leica).

Results

Presence of bacterial sequences in ~~the~~ UM3

In our previous pangolin genome sequencing project, we sequenced the genomes of pangolin cerebellum, cerebrum and liver using Illumina HiSeq2000 platform (13). During an *in-silico* bacterial sequence screening of the contig sequences of the tissue-specific assemblies using BLAST, we found many exogenous DNA sequences. The bacterial sequences were found in the assemblies of the cerebrum and cerebellum, but not in the liver assembly (Table S2). Specifically, in the assembled cerebral genome, there were 6,730 contigs mapped to bacterial genomes, where 6,635 of them (98.58%) had best matches with *P. fungorum*. Similarly, in the cerebellum-specific genome, 3,533 contigs mapped to bacterial genomes, among which 3,452 (97.7%) were from *P. fungorum*. These results indicate that the cerebrum and cerebellum tissues were predominantly colonised or infected by *P. fungorum* although they should be sterile.

PCR screening and Sanger sequencing across different tissues of UM3

~~To examine which tissues of UM3 were infected, we screened different tissues using PCR assays with *P. fungorum*-specific primers, *Burkholderia*-specific primers, and 16S primers (Table S1).~~ The genomic DNA of nine different tissues (cerebrum, cerebellum, liver, blood, kidney, thymus, spleen, lung, and heart) were extracted for testing. Our results showed clear positive bands in the lung, cerebrum, cerebellum, and blood (Fig. 1)~~, which was also observed using Sanger sequencing (Fig. S1)~~, whereas no clear bands were observed in other tissues (Fig. 1). The positive band in the blood tissue indicates that UM3 might have developed septicaemia. Therefore, we observed the presence of *P. fungorum* sequences across UM3's lung, cerebrum, cerebellum, and blood.

Histological examinations

To further confirm the presence of *P. fungorum* in UM3, the lung and cerebellum tissues were dissected and stained using Brown-Hopps Gram stains. Our staining revealed the presence of gram-negative and rod-shaped bacteria with a size of approximately 6-7 microns, supporting that the lungs and cerebellum were invaded by *P. fungorum* (Fig. 2A, B, D, and E). Histopathology screening was also performed using hematoxylin/eosin (H&E) staining to confirm the ~~pathology potentially due to~~ *P. fungorum* infection in the pangolin lung and brain

(cerebrum and cerebellum). The histological presentation of *Paraburkholderia* infection observed from lung tissues in other mammals is an abscess composed of cellular debris, numerous degenerate neutrophils, and macrophages that contain abundant intracytoplasmic basophilic material composed of rod-shaped bacteria (21). However, our analyses showed no significant pathological signs in the dissected organs (Fig. 2C and F). It may indicate the bacteria might colonise pangolins without any overt disease pathological presentations. However, we cannot rule out the possibility that the colonisation of *P. fungorum* in pangolins might be at the early stage where disease symptoms had not manifested.

Presence of *P. fungorum* in other adult pangolins

To examine whether the presence of *P. fungorum* in the UM3 pangolin organs was an isolated case, we randomly selected and screened the presence of *P. fungorum* in the blood of four individual pangolins (26T, 2T9, 12T, and 2T2) that seized in the same batch (about 40 pangolins) as UM3 in operation by the Department of Wildlife and National Parks Peninsular Malaysia (PERHILITAN) using the same sets of primers (Table S1) (22). As a control, we also tested the blood of two individual pangolins (UM1 and UM2) seized in a separate operation. Of the seven blood samples, four ~~pangolin samples~~ (2T9, 12T, 2T2, and UM3) showed positive PCR bands, and the presence of bacterial sequence was validated using Sanger sequencing (Fig. S1), again indicating that these pangolins' blood was infected by *P. fungorum* (Fig. 3). The pangolins showing positive PCR bands were all from batch 1, supporting our view that *P. fungorum* in UM3 is not an isolated case – perhaps they might be transmitted within the batch 1 pangolins.

***~~P.~~ fungorum* may have the capability to colonize fetus**

Since UM3 was pregnant, we wondered whether its fetus was also infected by *P. fungorum.* To examine this, we harvested and screened tissues of its fetus (including ~~blood cord,~~ lungs, intestine, kidney, liver, and brain) by performing PCR with the same sets of primers. We found that the fetal gastrocnemius muscle showed clear positive PCR bands (Fig. 4), and the presence of bacterial sequence was validated using Sanger sequencing (Fig. S1), but no significant bands were observed in other tissues (cerebrum, cerebellum, kidney, lung, ~~blood cord,~~ intestine and liver) (Fig. 4). Furthermore, the fetal muscle genome was also sequenced using the Illumina HiSeq 2000 platform with a 20X sequencing coverage (after removing the pangolin sequences)

and found a substantial amount of *P. fungorum* DNA sequences. Altogether, our data have confirmed the existence of *P. fungorum* in the fetal muscle and suggest the possibility of transplacental infection or an ascending infection pathway from the cervix.

~~Reconstruction~~ of *P. fungorum* genomes

To further analyse the genomes of *P. fungorum*, we ~~reconstructed~~ genomes using three different strategies: (i) ~~Mapped~~ reads from ~~the~~ UM3's cerebrum and cerebellum whole-genome data onto *P. fungorum* ATCC BAA-463 reference genome (Accession number: CP010024-27) and ~~assembled~~ them into contigs; (ii) ~~S~~equence~~d~~ the DNA extracted from ~~the infected~~ UM3's fetal muscle and ~~mapped~~ these reads onto *P. fungorum* reference genome and assemble~~d~~ them into contigs, and (iii) ~~Mapped~~ reads from ~~the~~ UM3's cerebrum, cerebellum, and fetal muscle whole-genome data onto *P. fungorum* reference genome and ~~assembled~~ them into contigs. The assembled genome sequences and 17 other *Burkholderia* and *Paraburkholderia* assemblies were uploaded to PanSeq server to identify Single Nucleotide Polymorphisms (SNPs) in common genomic regions (23). To identify the taxonomic position of these genomes, the core-genome SNPs were used to reconstruct a robust phylogenetic tree. Our results showed that our genomes were closest to *P. fungorum* ATCC BAA-463 (Fig. 5). Furthermore, the evolutionary relationships between the assembled genome (iii) and *Burkholderia* / *Paraburkholderia* spp. have been constructed based on concatenated sequences for 27 conserved proteins identified by ProteinOrtho and aligned with ClustalW, which further ~~validating the closest taxonomic relationship~~ to *P. fungorum* ATCC BAA-463 (Fig. 6). ~~The proposed taxonomic position of *P. fungorum* was very similar to the analysis of the phylogenetic study of *Burkholderiales* and *Paraburkholderiales* by Sawana *et al.* (5).~~

We also compared the pangolin *P. fungorum* genomes with the genomes of other *Burkholderiales* and *Paraburkholderiales* ~~spp.~~ by calculating ANI (15) and AAI values. Both ANI and AAI indicated that the pangolin *P. fungorum* was closely related to the reference *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. 7). As anticipated, the pangolin *P. fungorum* present in UM3's brain and fetal muscle ~~was~~ almost identical, with at least 99.9% for both AAI and ANI values, providing strong evidence that they

were from the same source. Altogether, the ANI, AAI and the core-genome SNP-based phylogenetic analyses are consistent and confirm the presence of *P. fungorum* in UM3.

Discussions

Here, we report a case of infection in a pregnant placental mammal, Malayan pangolin, called "UM3" and its fetus. The presence of *P. fungorum* in UM3 was confirmed by the evidence from PCR assays, histological examinations, whole-genome sequencing, phylogenetic analysis, ANI and AAI analyses. We identified *P. fungorum* in the cerebrum, cerebellum, lung, and blood of the pregnant UM3, but not in other tissues that we examined. Gram-negative and rod-shaped bacteria with a size of approximately 6 – 7 microns in the lungs and cerebellum provide strong evidence to support the invasion of *P. fungorum* in these supposedly sterile mammalian tissues. Wiersinga *et al.* suggested the lung is the primary target organ for infectious *Burkholderial* spp. such as *B. pseudomallei* and *B. mallei* (24). Moreover, it has been reported that *P. fungorum* was found in human blood with septicaemia and the bacterium was being transported in the circulatory system to other host organs (8, 25). Therefore, *P. fungorum* might initiate a systemic infection through the lungs of UM3 and spread to other critical organs, including the brain and blood.

Paraburkholderial species have been used as bioremediation or biodegradation medium of polycyclic aromatic hydrocarbon (PAHs) contaminated soil (10) and oxidized halo-benzene contaminated water (26, 27). The use of these opportunistic pathogenic species in agriculture potentially increases the risk of *Paraburkholderia* infection in humans by artificially introducing these bacteria into the contaminated soil and water. Notably, there are some reported clinical cases such as a nine-year-old female with *P. fungorum* causing septicaemia (9), a 66-year-old woman with *P. fungorum* LMG 16307 observed in the cerebrospinal fluid (24), and *P. fungorum* V02 10158 cultured from a pregnant woman's vaginal secretion (24). Our study showed that *P. fungorum* could cause septicemia and colonize the brain and lungs, as well as fetus, supporting the pathogenicity of *P. fungorum.* Our study may raise an alert on the use of *P. fungorum* in agriculture. We cannot rule out the possibility that *P. fungorum* may potentially target YOPI (young, old, pregnant, and immunocompromised) people.

For the first time, our data reported that *P. fungorum* could colonize the pangolin brain. A possibility is that the bacterial species might invade the pangolin blood-brain barrier (BBB).

Another possibility is that it might invade the brain through the olfactory nerve. The second possibility is more likely ~~because of the~~ three main reasons. It has been demonstrated that other *Burkholderia* species such as *B. pseudomallei* can invade the nerves of the nasal cavity by colonizing the thin respiratory epithelium and rapidly ~~migrates~~ along the underlying trigeminal nerve to penetrate the cranial cavity, thus leading to direct brain infection without going through BBB (28). Therefore, it is possible that the genetically related *P. fungorum* may also invade the pangolin brain via olfactory nerve cells. ~~The pangolins are known to have presumably~~ weak immunity due to the loss of the interferon epsilon (IFNE) gene, which is exclusively expressed in other mammalian epithelial cells and is important for skin and mucosal immunity. The weakened mucosal immunity of pangolin may make the invasion of *P. fungorum* into the olfactory epithelium easier.

Our PCR assays showed the presence of *P. fungorum* in pangolin fetal gastrocnemius muscle, but not in other fetal tissues (e.g., ~~blood cord~~, lung, and brain). Our results suggest that this bacterial species can infect ~~fetus in a pregnant female mammalian species~~; however, its underlying mechanism remains unknown. Notably, ~~it has been reported that~~ *B. pseudomallei* ~~causes~~ infectious disease in a pregnant woman, resulting in intrauterine infection with a subsequent spontaneous abortion (29). Therefore, it is possible that pangolin *P. fungorum* colonized the fetal muscle via a transplacental invasion as previously shown in *B. pseudomallei* in goats (30). However, it is unlikely since we could not detect *P. fungorum* in the ~~fetal~~ cord blood. Another possibility is that *P. fungorum* invaded the fetus through the urinary tract. This mechanism is similar to the invasive Group B *Streptococcus* bacteria that are able to infect the perinatal space in humans (31). Another possible mechanism is that the *P. fungorum* may be an invasive bacterial species that can penetrate the mucosa-protected cervix of the female pangolin and bypass the amnion of the uterus and the fetal skin before arriving in the leg muscle. Notably, some invasive pathogens such as Group B *Streptococcus* (31), *Listeria monocytogenes* (32) and *Mycoplasma hominis* (33) are also known to use this route to infect fetus. If it is true, this could be the first indirect evidence to show that ~~the~~ *P. fungorum* can be an invasive bacterial species and deserve more studies.

Our study ~~implicates~~ the importance of improving the management of these endangered pangolins. We should provide careful treatments and extensive medical care to pangolins~~,~~ ~~especially in the captive pangolin~~ conservation programmes. It is important to provide a hygienic

environment (as well as hygienic food and water) when keeping pangolins in captivity in order to minimize the risk of infection and stress. Regular monitoring of possible infections (e.g., blood screening) may also be an important measure in the rescue and conservation of ~~the threatened~~ pangolins in captivity.

Conclusion

Our study provides a first insight into the discovery of *Paraburkholderia* ~~genus~~ *fungorum* ~~species~~ in a mammal species, ~~Pangolin~~. We believe pangolin can be a reference for humans, particularly immunocompromised people, in studying *P. fungorum* infection. Our study also raises the concern about using *P. fungorum* as a biodegradation or bioremediation agent in agriculture.

Declaration of Competing Interest

The authors have no conflicts of interest.

Acknowledgements

CSW conceived and coordinated this project. TKY, RH and TTK performed experiments. TKY, TTK and RH analyzed data. KTW performed staining and histology experiments. CSW, TKY, SD, TTK, TBM Mohidin and RF wrote manuscript, ~~and~~ revised by TBM Mohidin and RF. This project ~~is~~ under SWC and RF ~~supervision~~. All authors read and approved the final manuscript.

Funding Statement

This work was funded by the high-level talent recruitment programme for academic and research platform construction (Reference Number: 5000105) from Wenzhou-Kean University. Furthermore, TKY was supported by the Centre for Research in Biotechnology for Agriculture (CEBAR) grant IRU-MRUN (RU023-2015).

Data Availability Statement

The genome datasets analyzed for this study can be found in the GenBank database with Accession number: CP028829- CP028832.

References

1. Ganguly S. Pangolin-zoological characteristics and its uniqueness in mammalian group. J Entomol Zool Stud. 2013;1(1):1-2.
2. Yang CW, Chen S, Chang CY, Lin MF, Block E, Lorentsen R, et al. History and dietary husbandry of pangolins in captivity. Zoo Biology: Published in affiliation with the American Zoo and Aquarium Association. 2007;26(3):223-30.
3. Hua L, Gong S, Wang F, Li W, Ge Y, Li X, et al. Captive breeding of pangolins: current status, problems and future prospects. ZooKeys. 2015(507):99.
4. Dias GM, de Sousa Pires A, Grilo VS, Castro MR, de Figueiredo Vilela L, Neves BC. Comparative genomics of Paraburkholderia kururiensis and its potential in bioremediation, biofertilization, and biocontrol of plant pathogens. MicrobiologyOpen. 2019;8(8):e00801.
5. Sawana A, Adeolu M, Gupta RS. 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. Front Genet. 2014;5:429.
6. Beukes CW, Palmer M, Manyaka P, Chan WY, Avontuur JR, van Zyl E, et al. Genome data provides high support for generic boundaries in Burkholderia sensu lato. Frontiers in microbiology. 2017;8:1154.
7. Estrada-de Los Santos P, Palmer M, Chávez-Ramírez B, Beukes C, Steenkamp ET, Briscoe L, et al. 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. Genes. 2018;9(8):389.
8. Coenye T, Laevens S, Willems A, Ohlen M, Hannant W, Govan JR, et al. Burkholderia fungorum sp. nov. and Burkholderia caledonica sp. nov., two new species isolated from the environment, animals and human clinical samples. Int J Syst Evol Microbiol. 2001;51(Pt 3):1099-107.
9. Gerrits GP, Klaassen C, Coenye T, Vandamme P, Meis JF. Burkholderia fungorum septicemia. Emerging infectious diseases. 2005;11(7):1115.
10. Andreolli M, Lampis S, Zenaro E, Salkinoja-Salonen M, Vallini G. Burkholderia fungorum DBT1: a promising bacterial strain for bioremediation of PAHs-contaminated soils. FEMS microbiology letters. 2011;319(1):11-8.
11. Liu X-x, Hu X, Cao Y, Pang W-j, Huang J-y, Guo P, et al. Biodegradation of phenanthrene and heavy metal removal by acid-tolerant Burkholderia fungorum FM-2. Frontiers in microbiology. 2019;10:408.
12. Jones A, Dodd M, Webb A. Burkholderia cepacia: current clinical issues, environmental controversies and ethical dilemmas. European Respiratory Journal. 2001;17(2):295-301.
13. Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, Wee WY, et al. Pangolin genomes and the evolution of mammalian scales and immunity. Genome research. 2016;26(10):1312-22.
14. Mount DW. Using the Basic Local Alignment Search Tool (BLAST). CSH protocols. 2007;2007:pdb top17.
15. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol. 2007;57(Pt 1):81-91.

16. Rodriguez-R LM, Konstantinidis KT. Bypassing cultivation to identify bacterial species. Microbe. 2014;9(3):111-8.
17. Ang MY, Dutta A, Wee WY, Dymock D, Paterson IC, Choo SW. Comparative genome analysis of Fusobacterium nucleatum. Genome biology and evolution. 2016;8(9):2928-38.
18. Tan SY, Tan IKP, Tan MF, Dutta A, Choo SW. Evolutionary study of Yersinia genomes deciphers emergence of human pathogenic species. Scientific reports. 2016;6:36116.
19. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC genomics. 2008;9:75.
20. Brown RC, Hopps HC. Staining of bacteria in tissue sections; a reliable Gram stain method. American Journal of Clinical Pathology. 1973;60:234-40.
21. Glaros TG, Blancett CD, Bell TM, Natesan M, Ulrich RG. Serum biomarkers of Burkholderia mallei infection elucidated by proteomic imaging of skin and lung abscesses. Clinical proteomics. 2015;12(1):1-14.
22. Choo SW, Zhou J, Tian X, Zhang S, Qiang S, O'Brien SJ, et al. Are pangolins scapegoats of the COVID‐19 outbreak‐CoV transmission and pathology evidence? Conservation Letters. 2020:e12754.
23. Laing C, Buchanan C, Taboada EN, Zhang Y, Kropinski A, Villegas A, et al. Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions. BMC bioinformatics. 2010;11:461.
24. Wiersinga WJ, van der Poll T, White NJ, Day NP, Peacock SJ. Melioidosis: insights into the pathogenicity of Burkholderia pseudomallei. Nat Rev Microbiol. 2006;4(4):272-82.
25. Gerrits GP, Klaassen C, Coenye T, Vandamme P, Meis JF. Burkholderia fungorum septicemia. Emerging Infectious Diseases. 2005;11(7):1115-7.
26. Strunk N, Engesser KH. Degradation of fluorobenzene and its central metabolites 3-fluorocatechol and 2-fluoromuconate by Burkholderia fungorum FLU100. Applied microbiology and biotechnology. 2013;97(12):5605-14.
27. Dobslaw D, Engesser KH. Degradation of toluene by ortho cleavage enzymes in Burkholderia fungorum FLU100. Microbial biotechnology. 2015;8(1):143-54.
28. John JAS, Ekberg JA, Dando SJ, Meedeniya AC, Horton RE, Batzloff M, et al. Burkholderia pseudomallei penetrates the brain via destruction of the olfactory and trigeminal nerves: implications for the pathogenesis of neurological melioidosis. MBio. 2014;5(2).
29. Chang CY, Lau NLJ, Currie BJ, Podin Y. Disseminated melioidosis in early pregnancy-an unproven cause of foetal loss. BMC infectious diseases. 2020;20(1):1-4.
30. Choy JL, Mayo M, Janmaat A, Currie BJ. Animal melioidosis in Australia. Acta tropica. 2000;74(2-3):153-8.
31. Cagno CK, Pettit JM, Weiss BD. Prevention of perinatal group B streptococcal disease: updated CDC guideline. American family physician. 2012;86(1):59-65.
32. Gibbs RS, Duff P. Progress in pathogenesis and management of clinical intraamniotic infection. American journal of obstetrics and gynecology. 1991;164(5 Pt 1):1317-26.
33. Eschenbach DA. Ureaplasma urealyticum and premature birth. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 1993;17 Suppl 1:S100-6.

Legend of Figures

Figure 1. *P. fungorum* screening of different tissues of UM3 using PCR assays. Nine sets of pangolin tissues and four sets of independent primer sets were used, including PCR results for (A) Target A primer set, (B) Target B primer set, (C) Target C primer set, and (D) the bacterial universal 16S primer set. N = Negative control, X = Cerebrum, Y = Cerebellum, Liv = Liver, B = Blood, H = Heart, T = Thymus, L = Lung, S = Spleen, and K = Kidney. Target A = *Burkholderia*-specific transposase genomic region; Target B = OI25_7129 hypothetical protein genomic region; Target C = *P. fungorum*-specific DNA polymerase genomic region.

Figure 2. Histological staining. Staining results at a magnification of 100X on brain tissue (A, B and C) and lung tissue (D, E and F). (A, B, D, and E) Red arrows point to the locations of Gram-negative rod-shaped bacteria. (C and F) H&E histological staining.

Figure 3. PCR assays using the blood of seven other individual adult pangolins. UM1 and UM2 were seized in an operation (first batch), whereas the pangolins UM3, 26T, 2T9, 12T and 2T2 were seized together in a separate operation and date (second batch). N = negative control; UM3 = positive control; Target A = *Burkholderia*-specific transposase genomic region; Target B = OI25_7129 hypothetical protein genomic region; Target C = *P. fungorum*-specific DNA polymerase genomic region.

Figure 4. *P. fungorum* screening on the fetal tissue using PCR assays. Target A = *Burkholderia*-specific transposase genomic region; Target B = OI25_7129 hypothetical protein genomic region; Target C = *P. fungorum*-specific DNA polymerase genomic region. (-ve = negative control; +ve = positive control; X = Cerebrum; Y = Cerebellum ; GL = Intestine; K = Kidney; T = Blood cord; L = Liver; G = Lungs; FB= Gastrocnemius muscle)

Figure 5. Phylogenetic tree generated using core-genome SNPs. *P. fungorum* assemblies generated from fetus-specific sequencing data, brain (cerebrum and cerebellum) -specific sequencing data and pooled sequencing data (cerebrum, cerebellum, and fetal muscle) were aligned with the genome sequences of 17 other *Burkholderial* species and the core-genome SNPs were extracted for alignment and tree reconstruction.

Figure 6. Phylogenetic tree generated using conserved protein. The conserved protein-based phylogenetic tree was generated using the Neighbour-joining (NJ) algorithm and 1,000 bootstrapping replications.

Figure 7. AAI and ANI analyses. The red line indicates the threshold of 95%, above which indicates the same species.

Supplementary Figure Legends

Supplementary Figure S1. Sanger sequencing performed on the blood and fetal with four independent primer sets. Each row represents a pangolin individual. (F = forward strand; R= reverse strand).

Supplementary Table Legends

Supplementary Table S1. PCR primers used in *P. fungorum* screening. Three sets of in-house designed primers and one set of universal 16S RNA primers were used in screening.

Supplementary Table S2. Identification of *P. fungorum* sequences in different pangolin tissues. We found 6,635 contigs (total length = 6,818,896 bp) in cerebrum-specific genomic data and 3,533 contigs (total length = 1,109,334 bp) in cerebellum-specific genomic data that have at least 90% identity and 90% coverage compared to *P. fungorum* ATCC BAA-463 genome.

Figure 1

P. fungorum screening of different tissues of UM3 using PCR assays.

Nine sets of pangolin tissues and four sets of independent primer sets were used, including PCR results for (A) Target A primer set, (B) Target B primer set, (C) Target C primer set, and (D) the bacterial universal 16S primer set. N = Negative control, X = Cerebrum, Y = Cerebellum, Liv = Liver, B = Blood, H = Heart, T = Thymus, L = Lung, S = Spleen, and K = Kidney. Target A = *Burkholderia*-specific transposase genomic region; Target B = OI25_7129 hypothetical protein genomic region; Target C = *P. fungorum*-specific DNA polymerase genomic region.

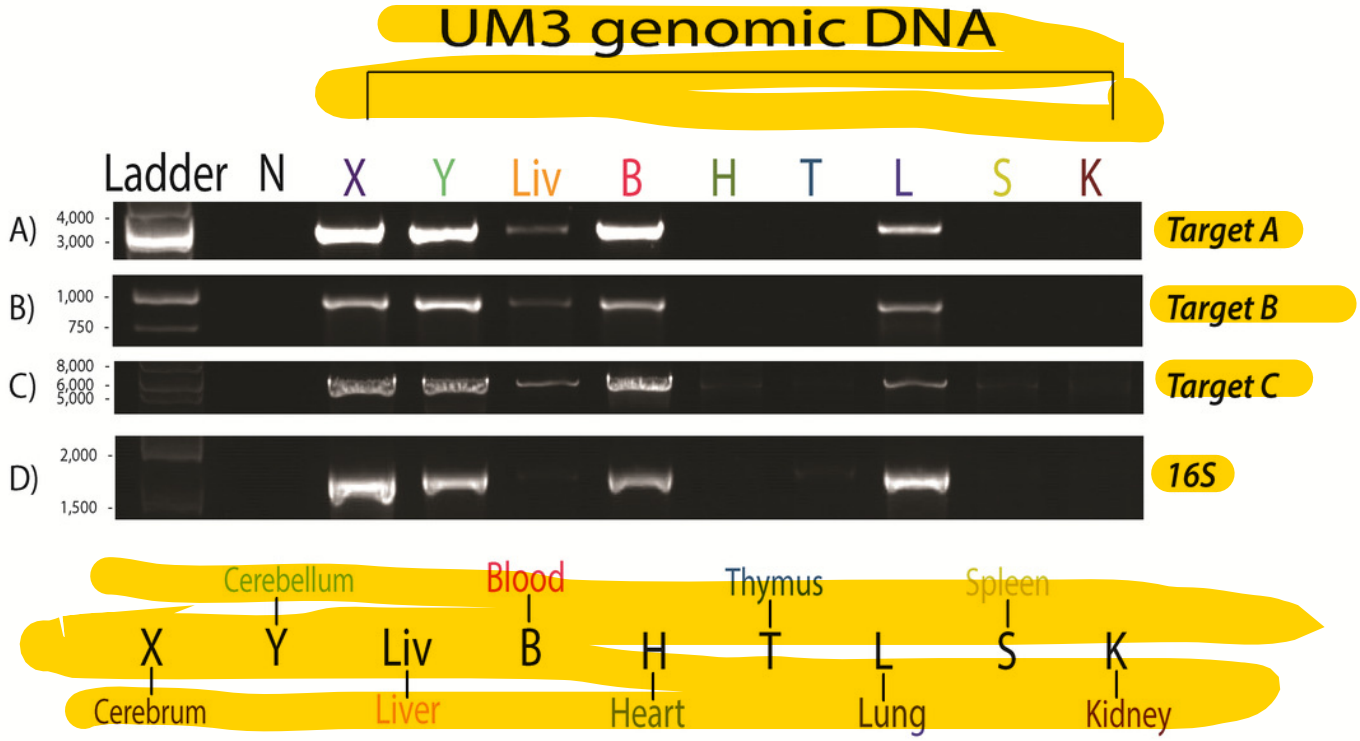

Figure 2

Histological staining. Staining results at a magnification of 100X on brain tissue (A, B and C) and lung tissue (D, E and F).

(A, B, D, and E) Red arrows point to the locations of Gram-negative rod-shaped bacteria. (C and F) H&E histological staining.

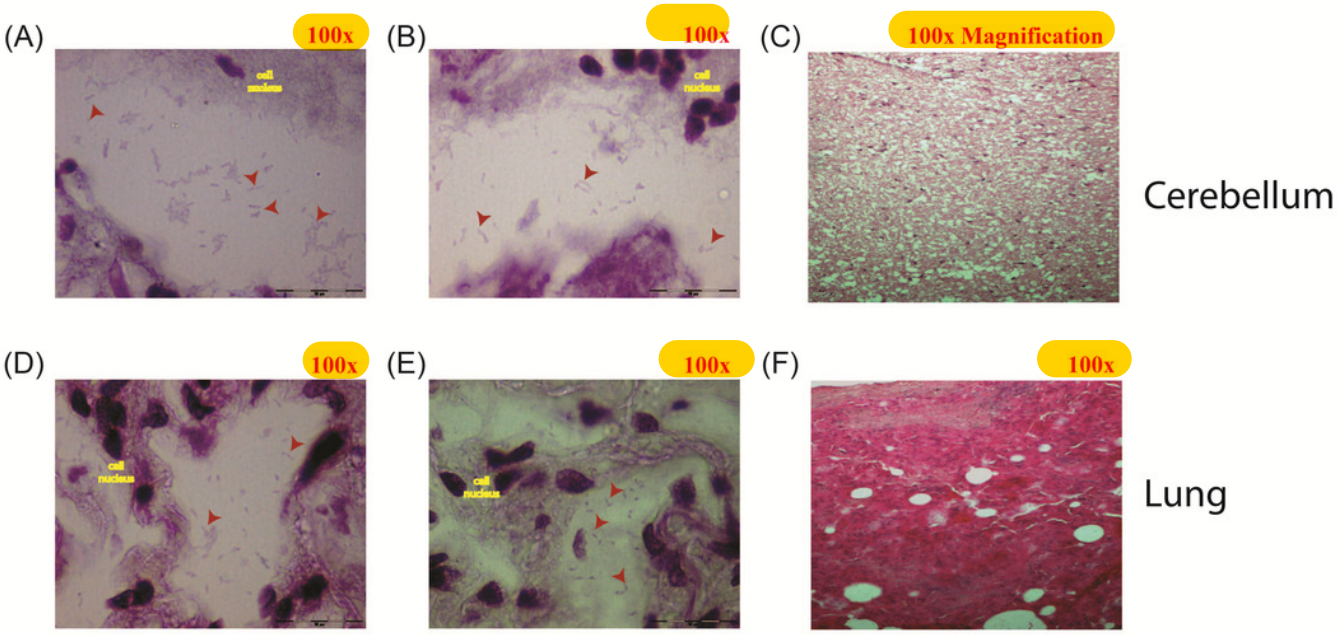
(A)
100x
cell nucleus
(B)
100x
cell nucleus
(C)
100x Magnification
Cerebellum
(D)
100x
cell nucleus
(E)
100x
cell nucleus
(F)
100x
Lung

Figure 3

PCR assays using the blood of seven other individual adult pangolins.

UM1 and UM2 were seized in an operation (first batch), whereas the pangolins UM3, 26T, 2T9, 12T and 2T2 were seized together in a separate operation and date (second batch). N = negative control; UM3 = positive control; Target A = *Burkholderia*-specific transposase genomic region; Target B = OI25_7129 hypothetical protein genomic region; Target C = *P. fungorum*-specific DNA polymerase genomic region.

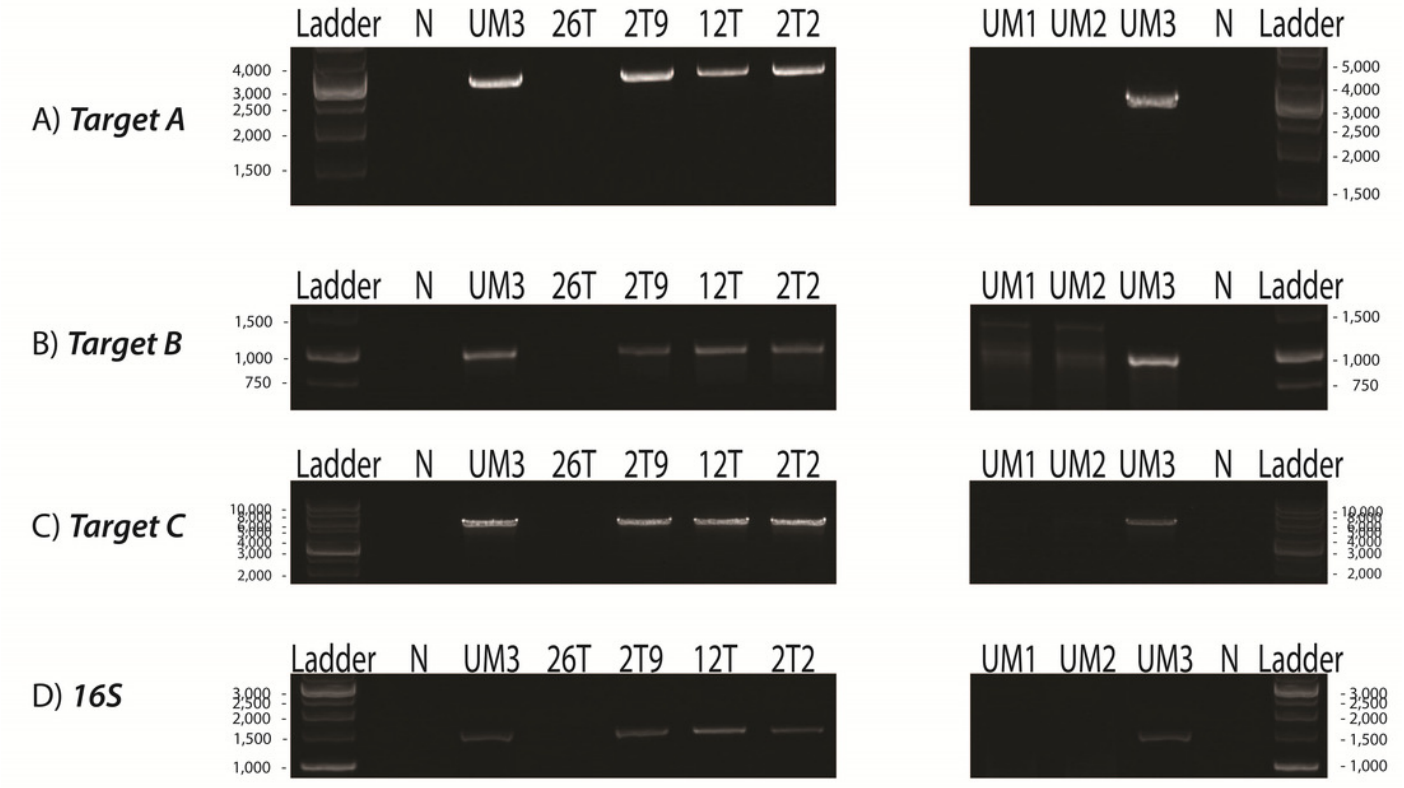

Figure 4

P. fungorum screening on the fetal tissue using PCR assays.

Target A = *Burkholderia*-specific transposase genomic region; Target B = OI25_7129 hypothetical protein genomic region; Target C = *P. fungorum*-specific DNA polymerase genomic region. (-ve = negative control; +ve = positive control; X = Cerebrum; Y = Cerebellum ; GL = Intestine; K = Kidney; T = Blood cord; L = Liver; G = Lungs; FB= Gastrocnemius muscle)

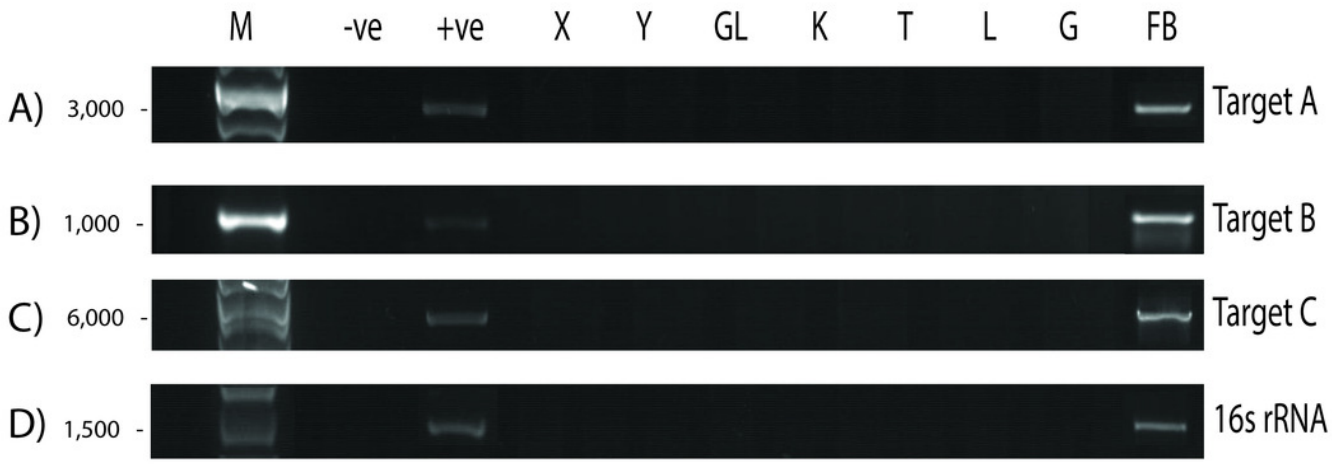

Figure 5

Phylogenetic tree generated using core-genome SNPs.

P. fungorum assemblies generated from fetus-specific sequencing data, brain (cerebrum and cerebellum) -specific sequencing data and pooled sequencing data (cerebrum, cerebellum, and fetal muscle) were aligned with the genome sequences of 17 other *Burkholderial* species and the core-genome SNPs were extracted for alignment and tree reconstruction.

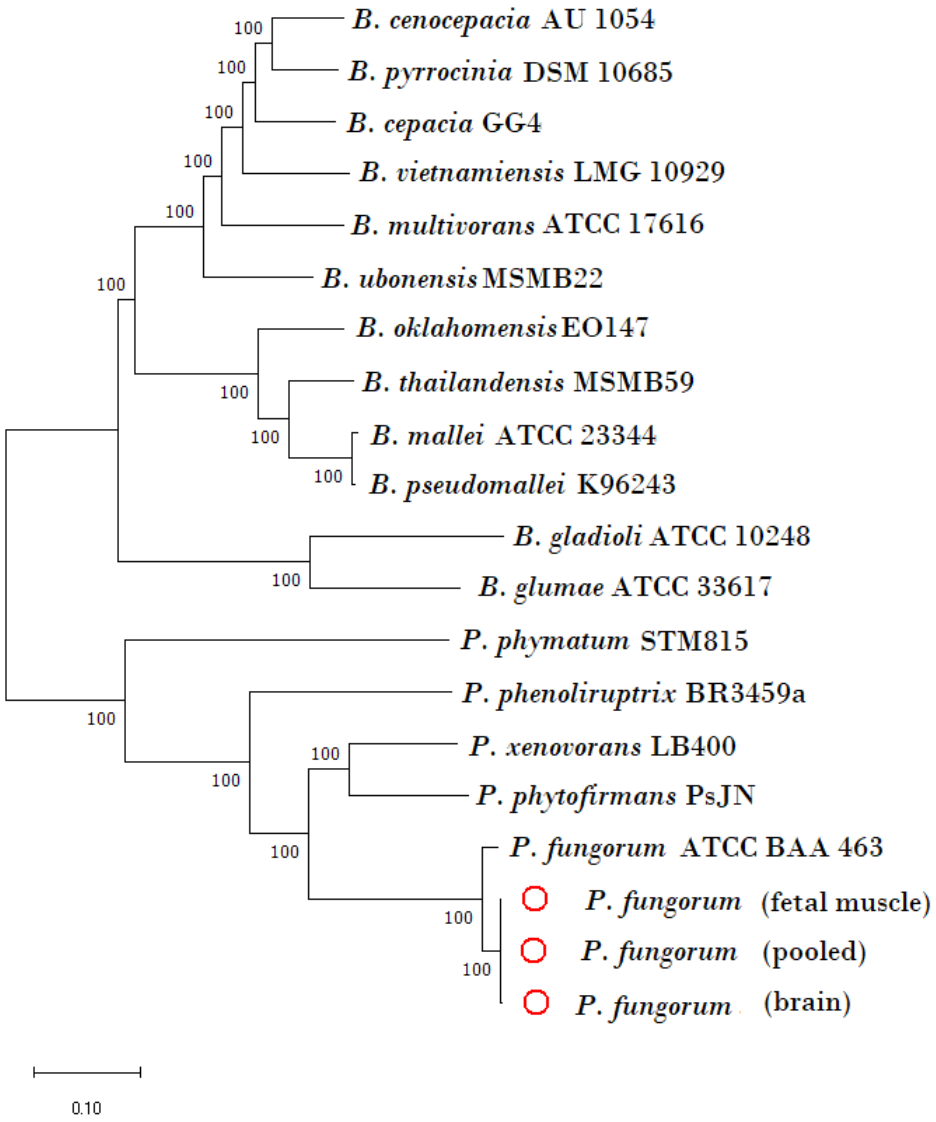
B. cenocepacia AU 1054
B. pyrrocinia DSM 10685
B. cepacia GG4
B. vietnamiensis LMG 10929
B. multivorans ATCC 17616
B. ubonensis MSMB22
B. oklahomensis EO147
B. thailandensis MSMB59
B. mallei ATCC 23344
B. pseudomallei K96243
B. gladioli ATCC 10248
B. glumae ATCC 33617
P. phymatum STM815
P. phenoliruptrix BR3459a
P. xenovorans LB400
P. phytofirmans PsJN
P. fungorum ATCC BAA 463
P. fungorum (fetal muscle)
P. fungorum (pooled)
P. fungorum (brain)
100
0.10

Figure 6

Phylogenetic tree generated using conserved protein.

The conserved protein-based phylogenetic tree was generated using the Neighbour-joining (NJ) algorithm and 1,000 bootstrapping replications.

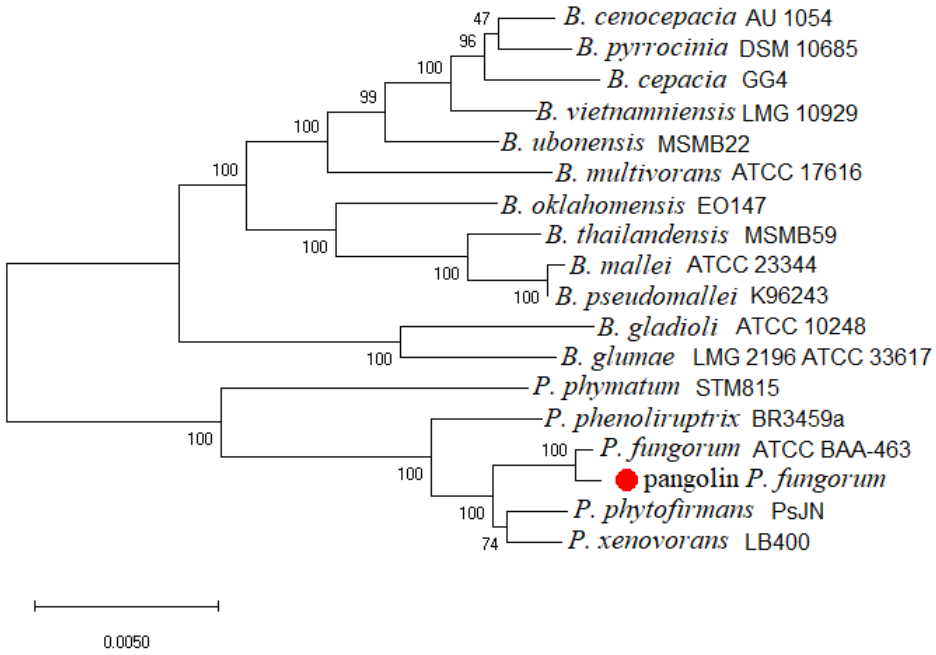
47
96
100
99
100
100
100
100
100
100
100
100
100
74
B. cenocepacia AU 1054
B. pyrrocinia DSM 10685
B. cepacia GG4
B. vietnamniensis LMG 10929
B. ubonensis MSMB22
B. multivorans ATCC 17616
B. oklahomensis EO147
B. thailandensis MSMB59
B. mallei ATCC 23344
B. pseudomallei K96243
B. gladioli ATCC 10248
B. glumae LMG 2196 ATCC 33617
P. phymatum STM815
P. phenoliruptrix BR3459a
P. fungorum ATCC BAA-463
pangolin P. fungorum
P. phytofirmans PsJN
P. xenovorans LB400
0.0050

Figure 7

AAI and ANI analyses.

The red line indicates the threshold of 95%, above which indicates the same species.