DATABASE
The Journal of Biological Databases and Curation

Original article

DemaDb: an integrated dematiaceous fungal genomes database

Chee Sian Kuan[1], Su Mei Yew[1], Chai Ling Chan[1], Yue Fen Toh[1], Kok Wei Lee[2], Wei-Hien Cheong[2], Wai-Yan Yee[2], Chee-Choong Hoh[2], Soon-Joo Yap[2] and Kee Peng Ng[1,*]

[1]Department of Medical Microbiology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia and [2]Codon Genomics SB, No.26, Jalan Dutamas 7, Taman Dutamas, Balakong, 43200 Seri Kembangan, Selangor Darul Ehsan, Malaysia

*Corresponding author: Tel: +60 3 79492418; Fax: +60 3 79494881; Email: kpng@ummc.edu.my

Abstract

Many species of dematiaceous fungi are associated with allergic reactions and potentially fatal diseases in human, especially in tropical climates. Over the past 10 years, we have isolated more than 400 dematiaceous fungi from various clinical samples. In this study, DemaDb, an integrated database was designed to support the integration and analysis of dematiaceous fungal genomes. A total of 92 072 putative genes and 6527 pathways that identified in eight dematiaceous fungi (*Bipolaris papendorfii* UM 226, *Daldinia eschscholtzii* UM 1400, *D. eschscholtzii* UM 1020, *Pyrenochaeta unguis-hominis* UM 256, *Ochroconis mirabilis* UM 578, *Cladosporium sphaerospermum* UM 843, *Herpotrichiellaceae* sp. UM 238 and *Pleosporales* sp. UM 1110) were deposited in DemaDb. DemaDb includes functional annotations for all predicted gene models in all genomes, such as Gene Ontology, EuKaryotic Orthologous Groups, Kyoto Encyclopedia of Genes and Genomes (KEGG), Pfam and InterProScan. All predicted protein models were further functionally annotated to Carbohydrate-Active enzymes, peptidases, secondary metabolites and virulence factors. DemaDb Genome Browser enables users to browse and visualize entire genomes with annotation data including gene prediction, structure, orientation and custom feature tracks. The Pathway Browser based on the KEGG pathway database allows users to look into molecular interaction and reaction networks for all KEGG annotated genes. The availability of downloadable files containing assembly, nucleic acid, as well as protein data allows the direct retrieval for further downstream works. DemaDb is a useful resource for fungal research community especially those involved in genome-scale analysis, functional genomics, genetics and disease studies of dematiaceous fungi.

Database URL: http://fungaldb.um.edu.my

Introduction

The kingdom fungi is made up of large eukaryotic organisms consisting of more than 100 000 species, including unicellular yeasts and multicellular fungi known as moulds and mushrooms (1). Dematiaceous fungi (brown-pigmented) are a large and heterogeneous group of moulds that produce melanin pigment in the cell wall of fungal hyphae or the conidia (2). Dematiaceous fungi occupy a plethora of niches, being found in soil, wood, as well as associated with plants as endophytes, saprophytes, parasites or plant pathogens (3–5). Until 2008, >130 species from 70 genera of dematiaceous fungi have been implicated in a wide range of human diseases, such as eumycetoma, chromoblastomycosis and phaeohyphomycosis (6). *Alternaria* spp., *Bipolaris* spp., *Cladophialophora bantiana*, *Curvularia* spp., *Exophiala* spp., *Fonsecaea pedrosoi*, *Madurella* spp., *Scedosporium prolificans*, *Neoscytalidium dimidiatum* and *Wangiella dermatitidis* are among the most important human pathogens commonly found in the tropical and subtropical climates (2). Additional reported cases worldwide further expand the existing long list of potential pathogens.

From 2008 to 2015, we have isolated a total of 437 dematiaceous fungi in the Mycology Unit of University Malaya Medical Centre (UMMC), Malaysia. These clinical isolates were recovered from superficial skin samples, nails, subcutaneous tissues, and nasopharyngeal secretion, blood, and tissue biopsies (7). Among these isolates, one strain of *Pyrenochaeta unguis-hominis* and two strains of *Nigrospora oryzae* demonstrated potential multidrug resistance features. In addition, several of the isolates are rare human pathogens or non-reported human pathogens, such as *Bipolaris papendorfii* (8), *Daldinia eschscholtzii* (9), *Pyrenochaeta* sp. (10), *Ochroconis* sp. (11) and *Cladosporium sphaerospermum* (12). Phylogenetic relationship of these dematiaceous fungi has been described by Yew *et al.* (7). The internal transcribed spacer (ITS)-based phylogenetic analysis resolved them into four distinct classes of Dothideomycetes, Sordariomycetes, Eurotiomycetes and one unclassified cluster (7).

The rapid advancement of Next-Generation Sequencing technologies has led to the sequencing of many fungal genomes, paving the way to decipher their biology and the underlying mechanisms of fungal pathogenicity and multidrug resistance. Several web-based analyses are also available for annotation of genes predicted from high-throughput genomic data to gain insight into the fungal living system machinery (13). However, the true challenge is to integrate the multiple sources of genomics data into useful information (14). In this work, we design the DemaDb that enables mycologists to access easily and analyse the genomics data using a genome browser. Currently, a total of eight genomes (*B. papendorfii*, *D. eschscholtzii*, *P. unguis-hominis*, *C. sphaerospermum*, *Ochroconis*
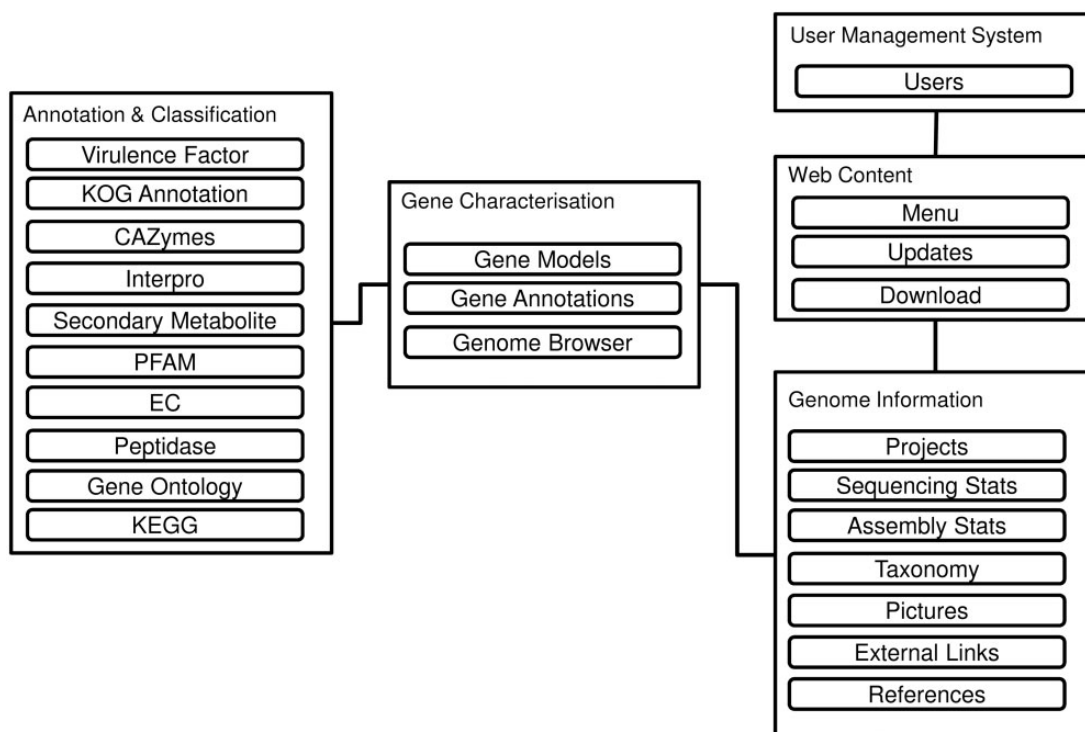
mirabilis, *Herpotrichiellaceae* sp. and *Pleosporales* sp.) have been integrated into the DemaDb. Considering that dematiaceous fungal genomes will be generated from future projects, it is essential to manage and integrate the data generated from different analyses in a more organized manner. The current version of DemaDb is freely available at fungaldb.um.edu.my and an improved version with additional genomes is forthcoming.

Database Organization

Figure 1 reveals the database schema of the DemaDb. DemaDb is built using a typical LAMP (Linux, Apache, MySQL and PHP) stack as the back-end components, along with Javascript and CSS (Cascading Style Sheet) for the front-end Web User Interface. The data are processed and stored in the MySQL relational database. The genome information is stored under 'Genome Information' table that records all the detailed information such as taxonomy classification, sequencing statistics, assembly statistics, pictures, references and external links. The information of gene models, annotations and genome browser links for each fungal genome is stored under 'Gene characterization' tables. 'Annotation and Classification' table is used to store various detail annotation, including Pfam, InterProScan, Gene Ontology (GO), EuKaryotic Orthologous Groups (KOG), Kyoto Encyclopedia of Genes and Genomes (KEGG), Carbohydrate-Active enzymes (CAZymes), peptidases, secondary metabolites and virulence factors.

Usage and Utility of Basic Data

A total of eight dematiaceous fungi that isolated from various samples (skin scraping, blood and nasopharyngeal secretion) were collected from University of Malaya between 2011 and 2014 (Table 1). The genomes of these rare and potential human pathogens were then sequenced and deposited in DemaDb. DemaDb is a web database that built based on a relational database system that contains dematiaceous fungal genomic profiles. The genome in DemaDb has its individual profile page with strain characteristics and genomic details. The information included on this page comprises the strain details, colonial characteristics on Sabouraud Dextrose Agar, microscopic morphology, taxonomic classification, assembly statistics and gene models (Figure 2A). The individual fungal profiles are linked to the NCBI taxonomic browser and The Catalogue of Life (15) to provide information on taxonomic hierarchy, distribution and ecological environment. Users can also explore the genome sequencing statistics in a single-genome context to gain insight into each sequencing platforms, read type, library size, genome size and sequencing

Figure 1. The database schema of DemaDb.

Table 1. Dematiaceous fungal genomes in the FungalDB

Fungal isolate	Isolation source	Isolation year
Bipolaris papendorfii UM 226	Skin scraping	2011
Cladosporium sphaerospermum UM 843	Blood	2008
Daldinia eschscholtzii UM 1400	Skin scraping	2012
Daldinia eschscholtzii UM 1020	Blood	2010
Herpotrichiellaceae sp. UM 238	Skin scraping	2011
Ochroconis mirabilis UM 578	Skin scraping	2012
Pleosporales sp. UM 1110	Nasopharyngeal secretion	2010
Pyrenochaeta unguis-hominis UM 256	Skin scraping	2011

coverage (Figure 2B). An overview of genome statistics (Table 2) and gene models (Table 3) of all genomes provides a basic comparative genomic analysis.
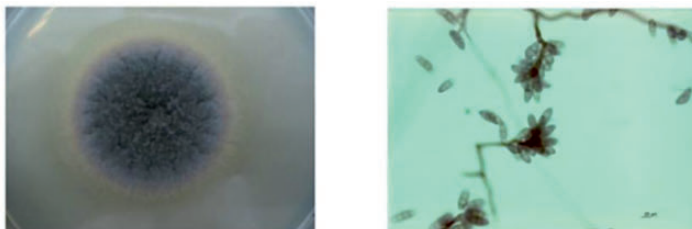
Uniform Functional Annotation

Application of the same annotation pipeline to all genomes is necessary for data integration in the DemaDb comparative genomic framework. All the raw data were pre-processed, assembled and functionally annotated using our pipeline. The compilation pipeline is provided in Figure 3. Protein-coding gene models were predicted from repeat-masked genome using GeneMark-ES version 2.3e (16). The

annotation of protein-coding gene models was completed using BLAST (Basic Local Alignment Search Tool) alignments of fungal genomes against NCBI non-redundant (nr) protein and SwissProt databases. Individual rRNAs and tRNAs were identified using RNAmmer v1.2 (17) and tRNAscan-SE v1.3.1 (18), respectively. All putative proteins were then functionally annotated. Pfam protein families database (19) and InterProScan 5 (20) were used to identify functional domains and sites in all predicted protein models. GO and KEGG metabolic pathways matches were carried out using local BLAST2GO tools (21). All the predicted proteins were also ascribed to 21 different functional groups based on KOG (22) for additional functional interpretation. The CAZymes was annotated by submitting the predicted protein models to the databases of automated Carbohydrate-active enzyme ANnotation (dbCAN) (23). The peptidases were identified by mapping all protein models against MEROPS database (24). Genomic mapping of fungal secondary metabolite clusters was performed using web-based SMURF (Secondary Metabolite Unknown Regions Finder) (www.jcvi.org/smurf/) (25). The putative virulence factor was predicted using PHI-base (The Pathogen-Host Interaction Database) (26).

The number of genes listed in the categories of KEGG, EC, GO, KOG, Pfam, Interpro, CAZyme, secondary metabolite, peptidase and virulence factor is available for comparison either among genomes in the DemaDb or other genomes outside the DemaDb (Tables 4). Each category is linked to a

A Morphological

The surface of the colonies was initially grayish white to olive green and becomes to black. The colony becomes mature in 5 days growing on Sabouraud Dextrose Agar (SDA) and a raised grayish periphery was observed. Colonies are effuce and velvety. Conidiophores are simple or branched, flexuose with zigzag rachis bearing conidium in a sympodial pattern. Conidia are curved and kidney-shaped, which usually broadest at the second cell from the base.

References

1. Giridharan, P., Verekar, S.A., Gohil, A.R., Mishra, P.D., Khanna, A. and Deshmukh, S.K., 2014. Antiproliferative activity of hamigerone and radicinol isolated from *Bipolaris papendorfii*. Biomed Res Int 2014, 890904.

2. Li, P.P. and Cao, Z.Y., 2013. First report of *Bipolaris papendorfii* causing corn leaf spot in China. Plant disease 97, 1506.

Taxonomic Classification

Kingdom:	Fungi
Phylum:	Ascomycota
Class:	Dothideomycetes
Order:	Pleosporales
Family:	Pleosporaceae
Genus:	*Bipolaris*
Species:	*Bipolaris papendorfii*

Strain Details

Stock Number	1154
Morphological Identity	*Bipolaris* sp.
Molecular Identity	*Bipolaris papendorfii*
Isolation Location	Skin scraping
Collection Date	24/2/2011
Isolation Country	Malaysia
Accession	JX966599

Assembly

Assembly size(bp):	33,397,470
Number of contigs:	374
Number of contigs (>=200bp):	374
Total contigs size(bp):	-
Contigs N50(bp):	146,099
Contigs GC content(%):	50.65
Number of scaffolds (>=1000bp):	-
Total scaffolds size(bp):	-
Scaffolds N50(bp):	-
Scaffolds GC content(%):	-

Gene Models

Number of predicted genes (>=99bp):	11,015
Average gene length:	1,425
Average number of exons per gene:	2.95
rRNAs:	18
tRNAs:	128

B

Organism	Platform	Read type	Library size(bp)	Number of reads	Read length(bp)	Total read size(bp)	Estimated genome size(Mbp)	Sequencing coverage(X)
Bipolaris	Illumina HiSeq 2000	Small Insert	500	33,655,556	90	3,029,000,040	32.4 - 34.8	87
Bipolaris	PacBio	Single End	20,000	405,257	2,742	1,111,215,948	32.4 - 34.8	34

Figure 2. The layout of Project page for *B. papendorfii* UM 226 genome. (**A**) Morphological features (colonial characteristic and microscopic morphology) of *B. papendorfii* UM 226 are shown on the left, whereas taxonomic classification, strain details, assembly statistics and gene models are shown on the right. (**B**) In the Sequencing Stats tab, an overview of sequencing technology platform, read type, library size (bp), number of reads, read length (bp), total read size (bp), estimated genome size (Mbp) and sequencing coverage (×) are displayed.

Table 2. Comparative gene models. Number of predicted genes, average gene length (bp), average number of exon per gene, rRNAs, and tRNAs were shown in gene models page

Organism	Number of predicted genes (≥99 bp)	Average gene length (bp)	Average number of exons per gene	rRNAs	tRNAs
Bipolaris papendorfii UM 226	11 015	1425	2.95	18	128
Cladosporium sphaerospermum UM 843	9652	1482	2.26	42	196
Daldinia eschscholtzii UM 1400	10 822	1483	2.87	29	168
Daldinia eschscholtzii UM 1020	11 120	1616	2.82	28	156
Herpotrichiellaceae sp. UM 238	9409	1544	1.93	9	67
Ochroconis mirabilis UM 578	13 435	1411	2.57	14	71
Pleosporales sp. UM 1110	14 074	1512	2.59	25	101
Pyrenochaeta unguis-hominis UM 256	12 545	1517	2.76	33	121

Table 3. Comparative genomic statistics. Assembly size (bp), number of contigs, contigs N50, contigs GC content (%), number of scaffolds, scaffolds N50, and scaffolds GC content (%) were shown in genomic statistics page

Organism	Assembly size (bp)	Number of contigs	Number of contigs (≥200 bp)	Contigs N50 (bp)	Contigs GC content (%)	Number of scaffolds (≥1000 bp)	Scaffolds N50 (bp)	Scaffolds GC content (%)
Bipolaris papendorfii UM 226	33 397 470	374	374	146 099	50.65	–	–	–
Cladosporium sphaerospermum UM 843	26 892 198	877	867	92 815	55.67	155	969 659	55.32
Daldinia eschscholtzii UM 1400	35 760 939	1944	1939	33 562	46.8	104	701 334	46.51
Daldinia eschscholtzii UM 1020	35 494 957	644	644	112 742	46.81	598	114 605	46.80
Herpotrichiellaceae sp. UM 238	28 370 377	232	217	270 646	49.83	128	455 601	49.83
Ochroconis mirabilis UM 578	34 611 065	603	544	220 443	52.1	163	1 170 353	51.84
Pleosporales sp. UM 1110	36 912 818	500	498	308 776	51.14	419	312 067	51.14
Pyrenochaeta unguis-hominis UM 256	35 484 281	286	286	457 153	50.4	254	481 751	50.35

Table 4. Functional annotation for all predicted gene models. All predicted gene models were functionally annotated on the basis of KEGG, EC, GO KOG, Pfam, Interpro, CAZyme, secondary metabolite, peptidase, and virulence factor classification. Individual functional annotation from this track is linked to corresponding page for additional information

Organism	KEGG	EC	GO	KOG	PFAM	InterPro	CAZyme	CAZyme unique	Secondary metabolite	Peptidase	Peptidase unique	Virulence factor
Bipolaris papendorfii UM 226	1206	2381	7154	6296	–	8282	729	669	32	153	150	442
Cladosporium sphaerospermum UM 843	999	1817	6070	5853	7092	6842	605	566	16	136	134	405
Daldinia eschscholtzii UM 1400	975	1962	6471	6168	7690	7959	664	619	47	181	178	602
Daldinia eschscholtzii UM 1020	998	1883	6224	6195	–	7996	660	618	45	187	184	606
Ochroconis mirabilis UM 578	1012	2029	6829	6909	8923	9397	590	559	14	186	179	401
Pyrenochaeta unguis-hominis UM 256	1337	2555	7616	6813	–	9223	808	725	21	192	187	454

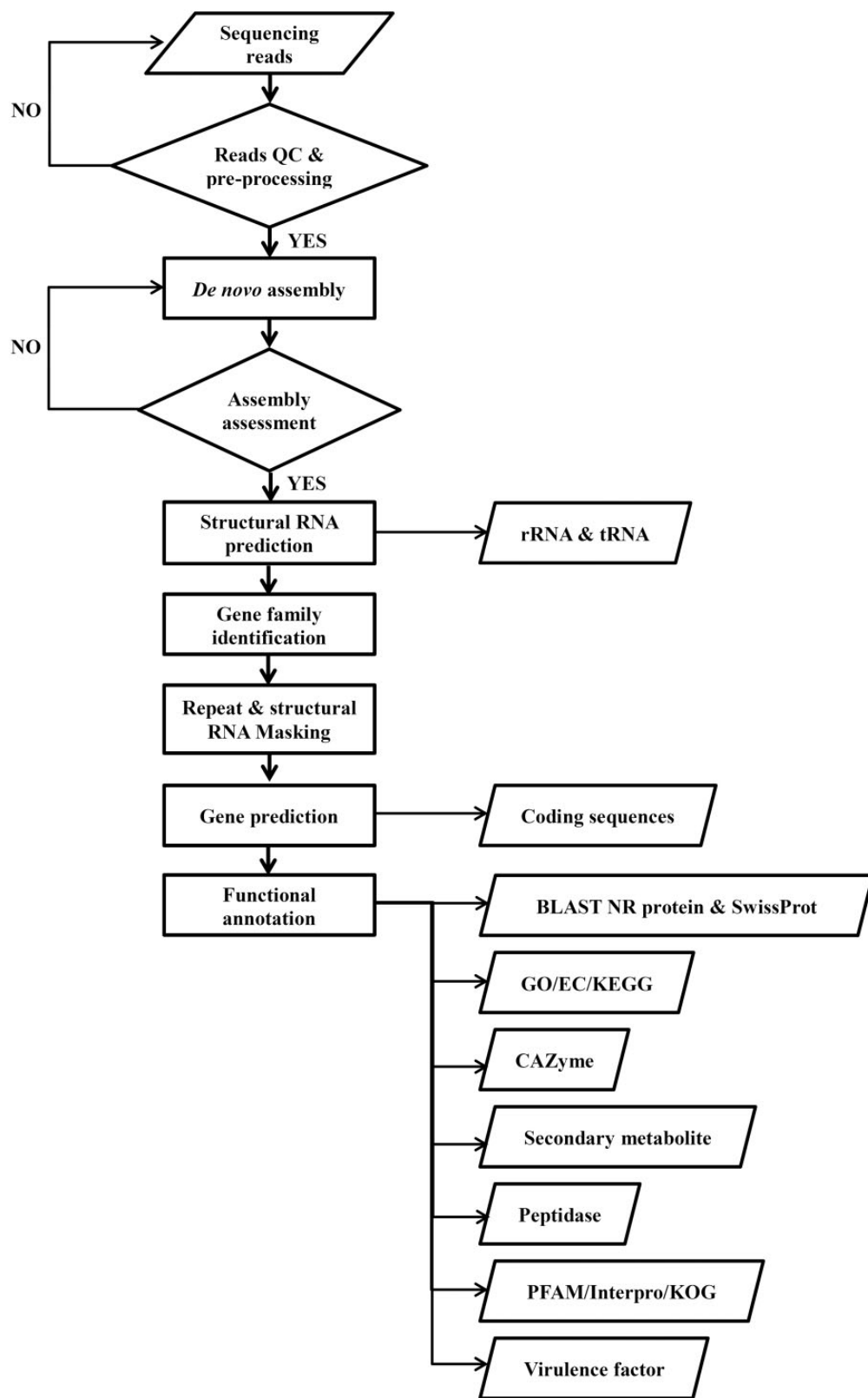| Gene ID	Family	Query start–end	Predicted active site residues	Predicted metal ligands	Hit name	Hit start–end	E-value	Merops annotation	Family type
UM256_ gene_133	M35	218–352	E305	H304, H308, D317	MER001399	216–351	$6.90e^{-36}$	Penicillolysin (*Penicillium citrinum*)	M35.001
UM256_ gene_284	S12	38–405	S67, K70, Y174		MER179257	31–389	$1.50e^{-44}$	Family S12 unassigned peptidases (*Aspergillus flavus*)	S12.UPW

Figure 3. Workflow for the compilation of pipeline for all genomic data in DemaDb.

Gene Details

| Gene Annotation | Gene Sequence | Peptidase | Interpro | EC | GO |

Gene Annotation

Source Genome	*Pyrenochaeta unguis-hominis*
Gene ID	UM256_gene_133
NR Annotation	hypothetical protein
SwissProt Annotation	Neutral protease 2 homolog SNOG_02177
Protein Size(aa)	352
Interpro	IPR001384 Peptidase M35, deuterolysin IPR024079 Metallopeptidase, catalytic domain

Gene Protein Sequence

```
>UM256_gene_133
MKVQVFSIAALASLASAFSDALSKRDSPLSVTLEQTGNTDVKAVITNTGAEDLKLFKTGTF
LDDSHVEKIEVFQAGEKVAFEGIRLRVSTANLDESAFKILAAGESIEASFDIAVAHDLGAG
GDFDLVTEGAFAYAELDSTVIAGAVPFISNTVTAAVNGEQAAKVRRDYQDLAKRTVVQSDC
TGTRRTATTTALSNCASLARTAASAAVNNAAKLNEYFKSTSSTTATTVQRVFNNVVTQCGS
STSGNSRYYCTDILGACSSGVLAYTYPSTSQMVNCPLFFTGLPALSRTCHAQDQATTVLHE
MTHLTQVKGTSDYGGYGYNFVRSLSAAQNLNHADTYTLFAQALYAGC
```

[Download] [View in GBrowse]

Peptidase

Family	M35	**Hit Start**	216
Query Start	218	**Hit End**	351
Query End	352	**E-value**	$6.90e^{-36}$
Predicted Active Site Residues	E305	**Merops Annotation**	penicillolysin (Penicillium citrinum)
Predicted Metal Ligands	H304, H308, D317	**Family Type**	M35.001
Hit Name	MER001399		

Interpro

Interpro Library	Interpro Signature	Signature Description	Match Start-End	E-value	Interpro ID	Interpro Description
PRINTS	PR00768	Deuterolysin metalloprotease (M35) family signature	228 - 253	$1.80e^{-1.80}$	IPR001384	Peptidase M35, deuterolysin
PRINTS	PR00768	Deuterolysin metalloprotease (M35) family signature	254 - 283	$1.80e^{-1.80}$	IPR001384	Peptidase M35, deuterolysin
PRINTS	PR00768	Deuterolysin metalloprotease (M35) family signature	195 - 224	$1.80e^{-1.80}$	IPR001384	Peptidase M35, deuterolysin
PRINTS	PR00768	Deuterolysin metalloprotease (M35) family signature	289 - 317	$1.80e^{-1.80}$	IPR001384	Peptidase M35, deuterolysin
Pfam	PF02102	Deuterolysin metalloprotease (M35) family	1 - 346	$7.40e^{-7.40}$	IPR001384	Peptidase M35, deuterolysin
SUPERFAMILY	SSF55486	-	179 - 352	$8.79e^{-8.79}$	-	-
Gene3D	G3DSA:3.40.390.10	-	179 - 352	$5.30e^{-5.30}$	IPR024079	Metallopeptidase, catalytic domain

EC

EC:3.4.24.0

GO

Aspect	GO ID	GO Description
C	GO:0005576	extracellular region
F	GO:0004222	metalloendopeptidase activity
F	GO:0046872	metal ion binding
P	GO:0006508	proteolysis

Figure 4. The layout of gene details page. The information, including gene ID, NCBI nr annotation, SwissProt annotation, GO annotation, protein size, domain sites, EC number and detailed functional annotation reports is available for every predicted gene.

detailed annotation report for every predicted protein in the individual genome. For example, users can explore peptidase families, the range of peptidase, active site residues, ligands for catalytic metal ions and *E*-value for the match in all predicted peptidases (Tables 4). In-depth multidimensional analysis can be performed using these data to obtain clues about their fungal lifestyle, adaptability, mating development, mechanisms underlying pathogenicity and drugs resistance. The details for every predicted gene, including gene ID, NCBI nr annotation, SwissProt annotation, GO annotation, amino acid sequence, protein size, domain sites, EC number, as well as the detailed functional annotation reports are available in the DemaDb (Figure 4). Functional annotations for *Herpotrichiellaceae* sp. UM 238 and *Pleosporales* sp. UM 1110 are in progress and will be included in an updated version of DemaDb.

Genome Browser

Generic Genome Browser (GBrowse) developed by GMOD (27) is integrated into DemaDb to allow graphical web visualization of the genomic data. The DemaDb genome browser is accessible through the links at each genome project portal page and genome's gene annotation & classification page, allowing the users to explore the genes of interest in a single-genome context. It also provides the navigation of the genomic regions for all eight fungal genomes in DemaDb, which can be freely switched through the drop-down menu of the data source box. Landmark or Region textbox acts as a universal search box, in which user could search by entering several types of inputs, including, but not limited to, range, gene name, chromosome name and description. It displays the predicted features of gene models along with their functional description (Figure 5). Single click on each gene track features will bring up a page showing additional annotation data retrieved from DemaDb database. By clicking on the gene ID, users will be brought to the detail description page of the annotation and amino acid sequence of the gene. The mRNA/CDS tracks are linked to additional gene details describing the gene structures and sequences (Figure 5). Additional tracks such as the DNA/GC Content track, 6-frame translation track and frame usage track can be displayed and configured using the toolbar.

Pathway Browser

The DemaDb Pathway Browser based on the KEGG pathway database enables users to analyse molecular interaction and reaction networks for all KEGG annotated genes. Pathway Browser provides a summary of metabolic pathways of each fungus and number of genes in each pathway. Users can select the pathways that they wish to analyse without the need for laborious search for all genes in every KEGG pathway. The top 15 hit pathways are shown in the Pathway Browser (Figure 6[TQ2]). Individual elements such as EC number and genes that are involved in specific pathway are linked to the corresponding pages for additional details. Users can also freely search for pathway map for all KEGG annotated genes (Figure 6).
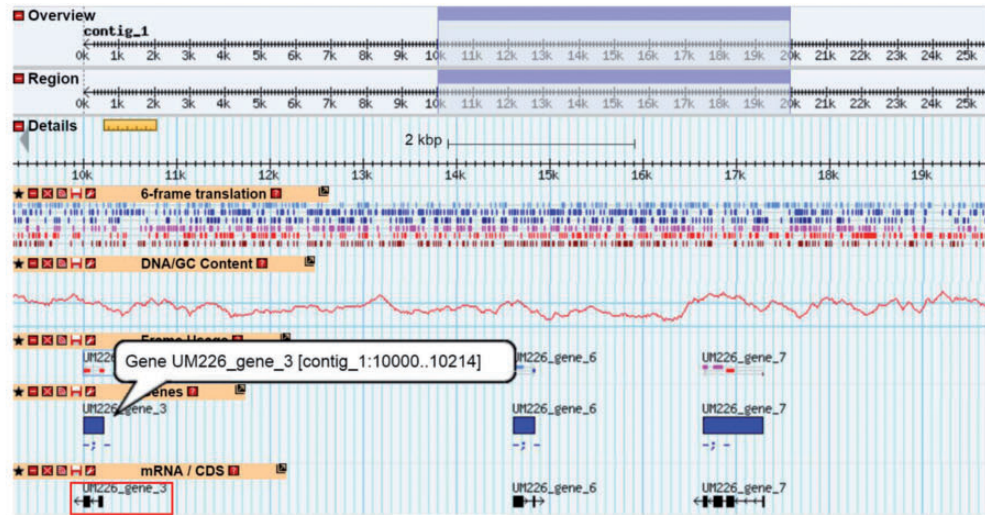
Downloadable Files

All assembly, nucleotides and proteins data can be downloaded in FASTA format to accommodate users who wish to pre-process the data based on additional parameters or conduct large-scale comparative genomic analysis.

Future Perspective

DemaDb will be updated whenever a new genome and its detailed analysis are completed. Details for each release will be displayed in the homepage. Users are also encouraged to deposit and integrate their fungal genomes into DemaDb, which will be coordinated by the principal investigator of DemaDb. The users can send their assembled sequences (fasta format), predicted genes (fna format) or functional annotated proteins (faa format) directly to the principal investigator where the contact details are available in the 'Contact' section. The quality of the raw data will be accessed and reviewed, and, if approved, will be annotated and deposited into the DemaDb. Such collective works are important for comparative genomic analysis to gain a better understanding of dematiaceous fungi functional diversity and evolution.

Up to date, we have produced a total of 398 ITS sequences from various dematiaceous fungi recovered from different anatomical sites. This large set of data will be archived into an updated version of DemaDb. Apart from the ITS sequences, isolation source as well as macroscopic and microscopic characteristics of these isolates will be incorporated to provide comprehensive information of each isolate. Also, a taxonomic engine will be constructed to allow users to identify their fungal isolate by conducting a similarity search against our curated database, and an ITS-based phylogenetic tree will be generated for taxonomic classification. This extensive collection of data would serve as a referral point for fungi isolated in tropical countries, such as Malaysia. As some fungal species appeared to be geographically restricted (28), this would provide a glimpse of idea on the commonly encountered dematiaceous fungal pathogen in Malaysia.
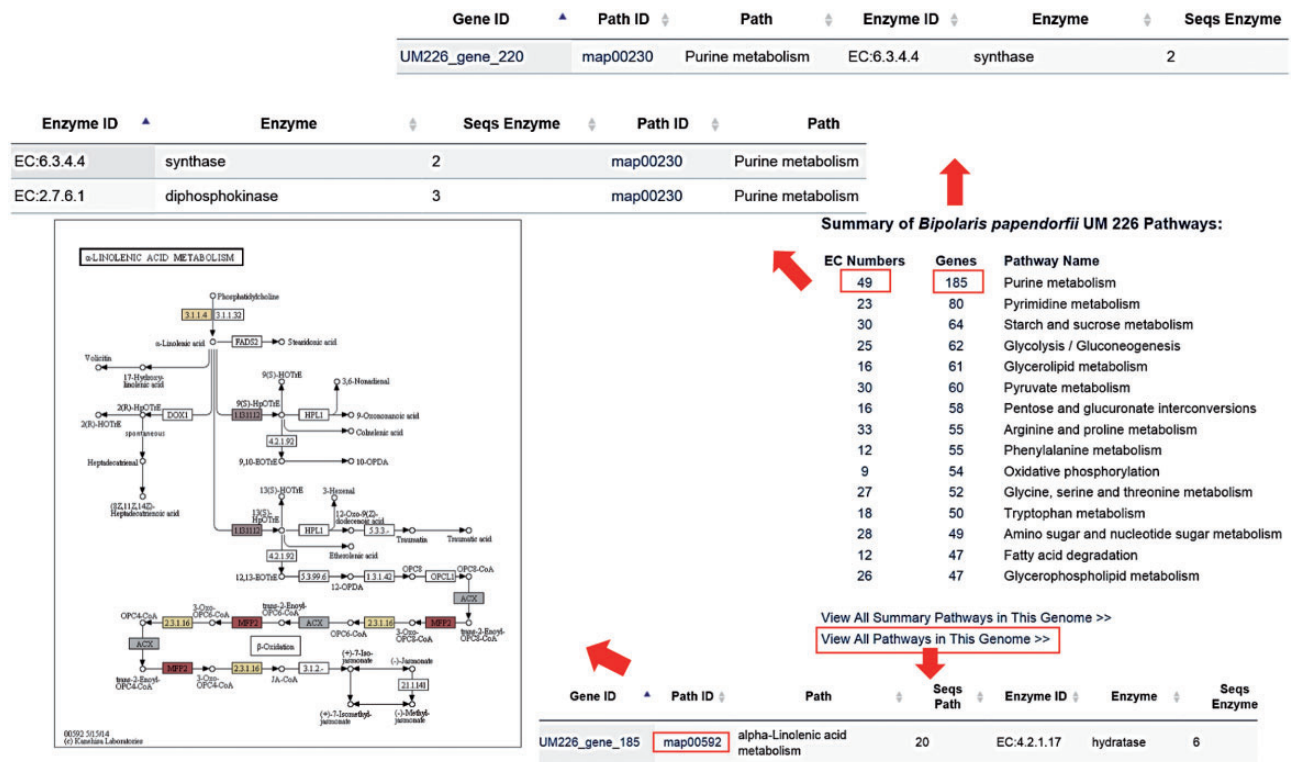
Figure 5. The layout of Genome Browser for *B. papendorfii* UM 226 genome. Horizontal tracks show genomic features of Contig 1. A highlighted blue rectangle indicates the genome region, that is, displayed in the details panel. A 6-frame translation track for DNA sequence is displayed in six different colour. The DNA/GC Content track represents the GC content in a given contig. In the genes track, blue bars indicate *B. papendorfii* UM 226 predicted genes. In the mRNA/CDs track, black arrows indicate *B. papendorfii* UM 226 mRNAs and their orientation. The information of a particular gene, including intron–exon organization, sequence and length is linked to each mRNA (black arrow).

Figure 6. The layout of Pathway Browser for *B. papendorfii* UM 226 genome. The total number of genes and EC numbers for each pathway is displayed in the Pathway Browser page. Single click on each EC Number or gene will bring up a page showing additional information of genes that are involved in the specific pathway. Each KEGG annotated gene is also linked to a pathway map. The coloured EC numbers indicate that the genes are mapped to that EC number in the pathway.

To the best of our knowledge, there is no integrated view of dematiaceous fungal gene expression profiles obtained from RNA-sequencing (RNA-seq). Currently, we are performing RNA-seq of these clinical isolates. In the future, the DemaDb database will collect these RNA-seq data and use a standardized method to identify the gene expression levels. Integration of RNA-seq data into the genome browser will be one of the new features in the upcoming DemaDb database.

Funding

Conflict of interest. None declared.

References

1. Elewski,B.E. and Hazen,P.G. (1989) The superficial mycoses and the dermatophytes. *J. Am. Acad. Dermatol.*, 21, 655–673.

2. Brandt,M.E. and Warnock,D.W. (2003) Epidemiology, clinical manifestations, and therapy of infections caused by dematiaceous fungi. *J. Chemother.*, 15, 36–47.

3. Zakaria,L., Yaakop,A.S., Salleh,B. *et al*. (2010) Endophytic fungi from paddy. *Trop. Life Sci. Res.*, 21, 101–107.

4. Velez,H. and Diaz,F. (1985) Onychomycosis due to saprophytic fungi. Report of 25 cases. *Mycopathologia*, 91, 87–92.

5. Guedes,F.L., Attili-Angelis,D. and Pagnocca,F.C. (2012) Selective isolation of dematiaceous fungi from the workers of *Atta laevigata* (Formicidae: Attini). *Folia Microbiol. (Praha)*, 57, 21–26.

6. Kumar,K.K. and Hallikeri,K. (2008) Phaeohyphomycosis. *Indian J. Pathol. Microbiol.*, 51, 556–558.

7. Yew,S.M., Chan,C.L., Lee,K.W. *et al*. (2014) A five-year survey of dematiaceous fungi in a tropical hospital reveals potential opportunistic species. *PLoS One.*, 9, e104352.

8. Kuan,C.S., Yew,S.M., Toh,Y.F. *et al*. (2015) Dissecting the fungal biology of *Bipolaris papendorfii*: from phylogenetic to comparative genomic analysis. *DNA Res.*, 22, 219–232.

9. Ng,K.P., Ngeow,Y.F., Yew,S.M. *et al*. (2012) Draft genome sequence of *Daldinia eschscholzii* isolated from blood culture. *Eukaryot. Cell*, 11, 703–704.

10. Yew,S.M., Chan,C.L., Soo-Hoo,T.S. *et al*. (2013) Draft genome sequence of dematiaceous coelomycete *Pyrenochaeta* sp. strain UM 256, isolated from skin scraping. *Genome Announc.*, 1, e00158–13.

11. Chan,C.L., Yew,S.M., Na,S.L. *et al*. (2014) Draft genome sequence of *Ochroconis constricta* UM 578, isolated from human skin scraping. *Genome Announc.*, 2, e00074–14.

12. Ng,K.P., Yew,S.M., Chan,C.L. *et al*. (2012) Sequencing of *Cladosporium sphaerospermum*, a Dematiaceous fungus isolated from blood culture. *Eukaryot. Cell*, 11, 705–706.

13. Haas,B.J., Zeng,Q., Pearson,M.D. *et al.* (2011) Approaches to fungal genome annotation. *Mycology*, 2, 118–141.

14. Schadt,E.E., Linderman,M.D., Sorenson,J. *et al.* (2010) Computational solutions to large-scale data management and analysis. *Nat. Rev. Genet.*, 11, 647–657.

15. Cachuela-Palacio,M. (2006) Towards an index of all known species: the Catalogue of Life, its rationale, design and use. *Integr. Zool.*, 1, 18–21.

16. Lomsadze,A., Ter-Hovhannisyan,V., Chernoff,Y.O. *et al.* (2005) Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.*, 33, 6494–6506.

17. Lagesen,K., Hallin,P., Rodland,E.A. *et al.* (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.*, 35, 3100–3108.

18. Lowe,T.M. and Eddy,S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, 25, 955–964.

19. Bateman,A., Birney,E., Cerruti,L. *et al.* (2002) The Pfam protein families database. *Nucleic Acids Res.*, 30, 276–280.

20. Quevillon,E., Silventoinen,V., Pillai,S. *et al.* (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.*, 33, W116–W120.

21. Conesa,A., Gotz,S., Garcia-Gomez,J.M. *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21, 3674–3676.

22. Tatusov,R.L., Fedorova,N.D., Jackson,J.D. *et al.* (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, 4, 41.

23. Yin,Y., Mao,X., Yang,J. *et al.* (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.*, 40, W445–W451.

24. Rawlings,N.D., Barrett,A.J. and Bateman,A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.*, 40, D343–D350.

25. Khaldi,N., Seifuddin,F.T., Turner,G. *et al.* (2010) SMURF: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.*, 47, 736–741.

26. Winnenburg,R., Baldwin,T.K., Urban,M. *et al.* (2006) PHI-base: a new database for pathogen host interactions. *Nucleic Acids Res.*, 34, D459–D464.

27. Stein,L.D., Mungall,C., Shu,S. *et al.* (2002) The generic genome browser: a building block for a model organism system database. *Genome Res.*, 12, 1599–1610.

28. Revankar,S.G. and Sutton,D.A. (2010) Melanized fungi in human disease. *Clin. Microbiol. Rev.*, 23, 884–928.