Infection, Genetics and Evolution 13 (2013) 261–266

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

MicrobPad MD: Microbial pathogen diagnostic methods database

B.C. Han [a,b,c,1], X.N. Wei [a,b,c,1], J.X. Zhang [c], N.Q.T. Truong [d], C.L. Westgate [e], R.Y. Zhao [f], Y.Z. Chen [a,b,c,*]

[a] State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, West China Medical School, Sichuan University, Sichuan 610041, P.R. China
[b] Computation and Systems Biology, Singapore-MIT Alliance, National University of Singapore, E4-04-10, 4 Engineering Drive 3, Singapore 117576, Singapore
[c] Bioinformatics and Drug Design Group, Department of Pharmacy, Centre for Computational Science and Engineering, National University of Singapore, Blk S16, Level 8, 3 Science Drive 2, Singapore 117543, Singapore
[d] Department of Biological Science, Faculty of Science, National University of Singapore, 10 Kent Ridge Road, Singapore 117546, Singapore
[e] Institute of Biomedical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom
[f] Central Research Institute of China Chemical Science and Technology, No. 18 Building, 5 Area, Anwai, Anhuali, Chaoyang District, Beijing 100011, P.R. China

ARTICLE INFO

Article history:
Received 12 June 2012
Received in revised form 21 October 2012
Accepted 26 October 2012
Available online 21 November 2012

Keywords:
Microbial pathogen
Diagnostic methods
Database

ABSTRACT

Medical pathogens induce infections, illnesses and sometimes serious medical conditions in the infected hosts. Diagnosis of these pathogens is important for proper treatment and investigation of pathogenesis processes. Molecular techniques have been developed for facilitating accurate, sensitive and low-cost diagnosis of these pathogens. Based on these techniques, diagnostic devices have been developed for a number of pathogens. More devices are needed for comprehensive coverage of medical pathogens. To facilitate the development of these devices, a database with integrated information about diagnostic methods, targets, and primers/probes for the known bacterial, fungal and viral pathogens is needed. We developed the microbial pathogen diagnostic methods database MicrobPad MD (http://bidd.nus.edu.sg/group/MicrobPad/MicrobPad.asp or http://pha-bidd.nus.edu.sg/group/MicrobPad/MicrobPad.asp) to provide comprehensive information about the molecular diagnostic techniques, targets, primers/probes, detection procedures and conditions, and tested diagnostic accuracies and limit of diagnosis for 314 bacterial, fungal and viral species from 61 genera. While available, additional information such as pathogen strains and hosts, tissue distribution or habitats, cultivation methods, biochemical characteristics, virulence factors, morphology, diseases, symptoms, treatment and prevention methods are provided. Our Database covers 242 gene targets, 700 primers/probes, 340 virulence factors, and 261 diseases. Cross-links to the NCBI genome and SwissProt/UniProt databases are provided.

1. Introduction

Medical pathogens of bacterial, fungal, and viral species induce infections, illnesses and sometimes serious medical conditions in the infected hosts (Ginocchio, 2011; Joseph and Read, 2010; Liljemark and Bloomquist, 1996; Preuner and Lion, 2009). Diagnosis of these pathogens is important for proper treatment and investigation of pathogenesis processes, and extensive efforts have been made for developing molecular techniques that enable fast, accurate, sensitive and low-cost diagnosis of these pathogens (Ginocchio, 2011; Joseph and Read, 2010; Russek-Cohen et al., 2011; Tenover, 2011). Based on these molecular techniques, advanced diagnostic devices have been developed for a number of medical pathogens (Endimiani et al., 2011; Ginocchio, 2011). More devices are needed for comprehensive coverage and faster diagnosis of medical pathogens, and for direct detection of multiple species (Ecker et al., 2010; Endimiani et al., 2011; Tenover, 2011).

Several databases have been developed and explored for providing the information and tools about the molecular diagnostic methods of specific classes of pathogenic species. For instance, the RIDOM website provides medical micro-organism differentiation services based on the analysis of small subunit ribosomal 16S rDNA sequences (Harmsen et al., 2002). An expanded MicroSeq 500 16S rDNA sequence library database (Cloud et al., 2004) and an integrated database network system (Conville et al., 2010) are useful for the identification of nocardia species. The fourth international spoligotyping database (Brudey et al., 2006) has been explored for the identification of *mycobacterium* species. A three-locus DNA sequence database is useful for the identification of the 69 *Fusarium* species associated with human or animal mycoses (O'Donnell et al., 2010). The 16SpathDB database supports automated identification of medically important bacteria by 16S rRNA gene sequencing (Woo et al., 2011). Another database provides

---

* Corresponding author at: Bioinformatics and Drug Design Group, Department of Pharmacy, Centre for Computational Science and Engineering, National University of Singapore, Blk S16, Level 8, 3 Science Drive 2, Singapore 117543, Singapore. Tel.: +65 6874 6877; fax: +65 6774 6756.

E-mail address: phacyz@nus.edu.sg (Y.Z. Chen).

[1] These authors contributed equally to this work.

1567-1348/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2012.10.017

pulsed-field gel electrophoresis patterns of epidemic-type oxacillin-resistant *Staphylococcus aureus* strains (McDougal et al., 2003). GenoBASE-pylori is useful for genotype searching of the human gastric pathogen *Helicobacter pylori* (Ahmed et al., 2007). TrED is a relational database that provides integrated access to various expression data of *Trichophyton rubrum* for developing effective diagnostic and treatment strategies (Yang et al., 2007).

These databases and web-tools are highly useful for the development of diagnostic devices of specific classes of medical pathogens. To facilitate the development of diagnostic devices for more diverse groups of medical pathogens, a database with integrated information about diagnostic methods, targets, and primers/probes for the known bacterial, fungal and viral pathogens is needed. Therefore, we developed the microbial pathogen diagnostic methods database, MicrobPad MD, to provide comprehensive information about the molecular diagnostic techniques, targets, primers/probes, detection procedures and conditions, and tested diagnostic accuracies and limit of diagnosis for 314 bacterial, fungal and viral species from 61 genera.

2. Database construction

MicrobPad MD is intended as a comprehensive resource for facilitating the research, development, and evaluation of molecular diagnostic methods for faster detection of pathogens that conventional diagnostic methods are inadequate to meet the treatment demand. For instance, more than half of the Tuberculous meningitis (TB) cases cannot be confirmed microbiologically and the conventional diagnostic method CSF takes over 2 weeks time for the test outcome, resulting in many patients being treated on the basis of clinical suspicion before the diagnosis is confirmed (Thwaites et al., 2000). Partly for dealing with this problem, two PCR-based molecular diagnostic devices, TB Amplicor and E-MTD, have been developed and approved by the FDA for diagnosis of TB from clinical specimens (Yang and Rothman, 2004).

In addition to the development of MicrobPad MD as a resource for microbial pathogen diagnostic methods, we also aim to provide additional information useful for understanding the characteristics and mechanisms of the microbial pathogens. These include pathogen strains and hosts, tissue distribution or habitats, cultivation methods, biochemical characteristics, virulence factors, morphology, diseases, symptoms, treatment and prevention methods are provided for facilitating the study of the molecular mechanisms of medical pathogens. Cross-links to the NCBI genome and Swiss-Prot/UniProt databases are provided.

MicrobPad MD is a freely accessible public online database and the full version of the database in text format can also be downed from the download page. Users are recommended to use the web-version because of its userfriendly format. Our database continues to be regularly updated and supported. Queries and suggestions are welcome and can be sent via email link provided in the MicrobPad MD webpage. Users are also welcome to send their new data via email or the new data upload page.

3. Data collection and access

The relevant data were collected from the literature searched from the Pubmed database (Sayers et al., 2011) by using keyword combinations of "diagnosis", "diagnostic", "detection", "detect", "bacterial", "fungal" and "microbial", "viral", "pathogen", and "pathogenetic", and from the information described in such review journals as Expert Rev Mol Diagn, Nature Rev Microbiol, and Trends in Biotechnology. A total of 382 papers were collected, which report or describe the molecular diagnostic methods, gene targets, primers/probes, detection procedures and conditions, and the tested diagnostic accuracies and limit of diagnosis for 205 bacterial species from 25 bacteria genera, 17 fungal species from six fungal genera, and 92 viral species from 30 virus genera. We further extracted the information from these papers and additional literatures searched from Pubmed (Sayers et al., 2011) for finding the pathogen strains and hosts, tissue distribution or habitats, cultivation methods, biochemical characteristics, virulence factors, morphology, diseases, symptoms, treatment and prevention methods for each species.

The MicrobPad MD data can be accessed by keyword or customized search. The keyword search is case insensitive and wildcards are supported. In a query, a user can specify full name or any part of the name in a text field. Wild characters of '*' and '?' are allowed in text field. Here, '?' represents any one character and '*' represents a string of characters of any length. For example, input of 'toxin' in the query field finds entries containing 'toxin' in their names, such as alpha toxin, beta toxin, epsilon toxin, RTX toxin and enterotoxins. On the other hand, input of 'Clostridium*' finds all the species start their genus names with 'Clostridium'. In this case, '*' represents 'perfringens A', 'perfringens B', 'septicum', 'difficile', etc.

Customized search (Fig. 1) fields include genus name, species name, target name, disease indication and virulence factor. The result of a search is illustrated in Fig. 2, in which all entries that satisfy the search criteria are listed. This list includes the MicrobPad entry ID, genus name, species name, virulence factor, target gene, disease indications, and the number of diagnostic methods. The related species and diagnosis method page (Fig. 3) can be obtained by clicking the "MicrobPad ID" link of a selected MicrobPad entry. The page of species and diagnostic method contains two sections. The first and second section provides detailed description about the medical species and the diagnostic methods respectively. Further information about the genome of the species, target genes and virulence factors can be accessed via crosslink to NCBI genome databases (Sayers et al., 2011) and SwissProt/UniProt database (Consortium, 2011). The whole MicrobPad methods data can be downloaded via the download link.

4. Database usage and validation

Users of MicrobPad MD are expected to have basic knowledge about the popular molecular diagnostic techniques such as PCR, Multiplex PCR, real-time PCR, the diagnostic markers of microbial pathogens including the targets, primers, and probes, and the commonly used detection procedures. To facilitate the users for studying the relevant techniques, all the techniques used in the diagnostic methods described in the MicrobPad MD are provided in the help page. For searching MicrobPad MD, users are also expected to have the knowledge of at least one of the following items: pathogen genus name, pathogen species name, virulence factor, detection target name, and disease indication. Users can use both keyword search and browsing facilities (with pull-down manuals of pathogen, Target Name, Disease Indication and Virulence Factor lists) for selecting the relevant diagnostic method. Keyword search function supports incomplete word search such that all items that partially match the input keywords are displayed for user to select appropriate entries. Our database is built based on IIS HTTP server, ASP (Microsoft's server-side script engine for dynamically generated web pages) and Access (Microsoft's database manage system). There is no special requirement for client users. It can be easily accessed on various operation systems by common Internet Browsers such as Internet Explorer, Chrome, Firefox and Safari.

As an illustrative example, in order to find the diagnostic method for detecting the disease "Brucellosis" from clinical samples, the keyword "Brucellosis" can be entered into the MicrobPad MD