DengueInfo: A web portal to dengue information resources


a Novartis Institute for Tropical Diseases, Chromos #05-01, 10 Biopolis Road, Singapore 138670, Singapore
b Genome Institute of Singapore, Genome #02-01, 60 Biopolis Street, Singapore 138672, Singapore
c Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA
d Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

Received 7 October 2006; received in revised form 3 February 2007; accepted 7 February 2007
Available online 11 February 2007

Abstract

DengueInfo (http://www.dengueinfo.org) is a web portal and database that brings clarity to dengue research by integrating the growing number of complete genome sequences of dengue virus with relevant literature and curated epidemiological information. Additionally, it represents a repository of ongoing prospective and retrospective studies of dengue disease severity. We intend the database to be a flagship resource for the dengue community, providing standardized and high quality information and facilitating research into key aspects of dengue biology and assisting in its control. To aid this process we also introduce a standard nomenclature for dengue isolates inspired by globally accepted systems used for influenza virus.

Keywords: DengueInfo; Dengue virus; Genomics; Epidemiology; Nomenclature

1. Introduction

Dengue is a mosquito-borne viral disease that affects millions of individuals per year and imposes a particularly heavy burden on developing countries in tropical and subtropical regions. The disease is now endemic in more than 100 countries, placing some 2500 million people at risk of infection each year (WHO, 2002; Kroeger et al., 2004). Typically, a dengue viral infection will lead to a self-limiting dengue fever (DF). However, the incidence of dengue epidemics is increasing and a larger percentage of cases are progressing to severe dengue disease, usually classified as dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS) (Gubler, 2002). The causes of DHF and DSS are poorly understood as are the reason(s) for the increasing incidence of severe disease, making dengue a first-rank threat to public health. Additionally, a dengue vaccine is complicated by the existence of four viral serotypes (DENV-1 to DENV-4). The compilation of dengue genomic and associated information into DengueInfo will, we hope, accelerate understanding of these phenomena.

As large scale sequencing of the complete genome of dengue virus (DENV) is now feasible, it is imperative to link these data to pertinent epidemiological and geo-medical information. To this end we have created the DengueInfo database to serve as a portal to dengue genomic and related information. Initially the database has been seeded with the complete genome sequences of DENV collected from GenBank, literature sources and personal communications. As part of ongoing prospective and retrospective studies we have added several additional genome sequences along with detailed associated data.

2. Access to DengueInfo and data retrieval tools

DengueInfo may be queried using a simple-search interface which presents a web form with a text field where search terms may be entered. Records matching any search term entered will be retrieved. More powerful searches are possible with the complex search interface. Multiple criteria can be searched including disease severity, year of isolation, serotype, isolate, country of isolation, etc. Exclusions can be added so that
sequences that match the exclusion can be filtered from the returned results.

Irrespective of the search type the results are returned in a unified interface: GenBank accession numbers and a short description of each sequence are displayed ordered by a score (either the number of search criteria matched or bit-scores in the case of BLAST queries). Sequences of interest can be selected for further analyses including: downloading the sequence (or subsequence) in FASTA or GenBank format, amino acid translation, and displaying clinical information for the selected records. For many isolates additional metadata is available including the serotype, disease severity (DF, DHF, DSS), country of origin and date of sampling. Additionally, isolates will be marked as wild type or engineered if they have been modified or attenuated for vaccine studies. Links to PubMed are provided if there are publications relevant to the virus. For isolates that are part of clinical studies we also present extensive clinical parameters collected at the time of the study.

DengueInfo implements two information aggregators that act as alert services for recent news and research articles about dengue. News articles are collected from Yahoo! News using an RSS feed. Abstracts of recent research articles are retrieved from Entrez PubMed using NCBI's eUtils interface. Summaries of these news and research articles are displayed with links to the original sources.

The website also permits molecular analysis of genome sequences. DENV sequences can be aligned and submitted to a mutation viewer or used to build a phylogenetic tree. Mutations in dengue proteins with crystal structures can also be submitted to the SwissModel homology modelling server.

3. A standardized dengue nomenclature

To improve clarity and help unify the dengue research field we have adopted a standardized isolate naming scheme inspired by the nomenclature used for isolates of influenza virus. An example standard identifier will appear as follows: D1/AU/TSV01/1999. The first part identifies the viral serotype (DENV-1 to DENV-4). The second two letters identify the country of isolation (using ISO country codes (ISO 3166-1:1997)). The third part is the common or laboratory name of the isolate, if available. Finally, the last 4 digits are the year of isolation. Where information is missing we propose that X’s be used (four in the case of a missing date); thus the most ambiguous identifier would be DX/XX/XX/XXXX. In rare cases where forward slashes appear in the original strain name (/) these should be replaced by a dash (–).

This simple and powerful scheme provides unique and informative names for isolates, and allows retrospective re-naming. We believe this nomenclature will greatly enhance dengue research and encourage others to adopt it.

4. Conclusion

The DengueInfo portal has been available for over 18 months and attracts between 500 and 700 unique users each month. A recent upgrade has seen the addition of new functionality and data; we have included over 100 newly sequenced complete DENV genomes along with clinical data from more than 300 individuals. We believe the portal has the potential to become a focal point for dengue related information and we encourage all interested parties to use and contribute to this free community effort. We hope it may also serve as a template for other flaviviral diseases such as Yellow Fever, Japanese Encephalitis and West Nile Fever.

DengueInfo is accessible for public and commercial use at http://www.dengueinfo.org.

Acknowledgements

We thank Hilmar Lapp from the Genomics Institute of the Novartis Foundation (GNF) for his detailed review of the database and interface design. We also thank Frans Verhoef from the Genome Institute of Singapore (GIS) for his conceptual design of the complex search interface. Much of the current design was implemented by KOOPrime Pte. Ltd. (www.kooprime.com).

References


WHO Fact Sheet No. 117, 2002.