

Respiratory Disease Information Hub Project Proposal

Introduction

Respiratory diseases represent a significant public health concern worldwide. While various online resources like NCBI GenBank provide extensive information on respiratory diseases, we do not have a centralized database where users can readily access comprehensive information, including common names, gene annotations, reference sequences in FASTA format, major subtypes, and symptoms, all conveniently available in one location. Therefore, this project aims to bridge this gap by developing a Respiratory Disease Information Hub. This web-based application will serve as a centralized platform where users can easily access detailed information about respiratory diseases, aiding researchers, healthcare professionals, and the general public in understanding and combating these conditions effectively.

Functionality

Users can search for respiratory diseases using common or scientific names, with autocomplete support for efficient querying. Upon submitting a search query, an AJAX call is made to the backend server, which retrieves relevant information from a MySQL database based on the user's input. The retrieved information, including scientific name, common names, major subtypes, genome types, and symptoms associated with the respiratory disease, is dynamically updated on the user interface. For each respiratory disease displayed in the search results, users have the option to view its gene annotation by clicking on a clickable text labeled "View Gene Annotation." This directs users to a page displaying gene annotation details for the selected disease. Gene annotations are presented in a tabular format, including columns for gene name, product, 5' and 3' coordinates on the reference sequence, and the actual corresponding DNA sequence. The DNA sequence is initially truncated to show only the first few characters to ensure readability, but users can click on "View All" within the DNA sequence box to view the full sequence. At the top of the gene annotation page, there will be a sentence indicating the NCBI reference sequence used as the template for the gene annotation.

Wireframe Example

This wireframe example uses "hMPV" as the search query and is for demonstration only. The final appearance and layout may not exactly resemble the wireframe example shown.

1. As the user begins typing "hMPV" into the search bar, the system provides autocomplete suggestions based on the entered characters.

Respiratory Disease Information Hub

hM	Submit
hMPV	

2. After the user clicks the submit button, the page is updated to display information about hMPV.

Respiratory Disease Information Hub

<input type="text" value="hMPV"/>	<input type="button" value="Submit"/>
-----------------------------------	---------------------------------------

Scientific Name	Other Name(s)	Major Subtype(s)	Genome	Symptoms	
Human Metapneumovirus	HMPV, hMPV	A1, A2, B1, B2	Negative-sense single-stranded RNA virus	Cough, nasal congestion, sore throat, fever, and shortness of breath	View Gene Annotation

3. After the user clicks on the "View Gene Annotation" text, the system reloads the page to display gene annotation details for hMPV. The user can view the complete sequence of each gene by clicking on the "View All" text.

Respiratory Disease Information Hub

Human Metapneumovirus Gene Annotation
NCBI Reference Sequence: NC_039199.1

Gene	Product	5' Coordinate	3' Coordinate	Sequence
N	nucleoprotein	40	1224	ATGTCTCTCAAGGGATT ... (View All)
P	phosphoprotein	1248	2132	ATGTCATCCCTGAAGGA ... (View All)
...

Technical Requirements

UNIX OS and Filesystem: The project will be developed within a UNIX environment.

Placement/Organization of Files within a Web Application: Project files will be organized within the directory /var/www/html/ywang833/final/ following the principles of content separation.

Python CGI Programming: Python CGI scripts will handle user requests and generate content for display. It will execute SQL queries to fetch disease information from the MySQL database using mysql.connector.

CGI and HTML Templates: HTML templates will be used to define the layout of web pages, with placeholders for content generated by CGI scripts.

HTML5 Document Markup: The HTML document for this project will adhere to the syntax and structure defined by the HTML5 standard

Relational Database Schema Design: A relational database schema will be designed to store information about respiratory diseases, including disease characteristics, gene annotations, and reference sequences, with appropriate entities and relationships.

MySQL: MySQL will be used as the relational database management system.

Limited Page Styling with CSS: CSS will be employed for basic page styling.

JavaScript and JQuery: JavaScript and JQuery will facilitate client-side interactions, such as autocomplete functionality and dynamic content loading.