TABLE OF CONTENTS

Chapter-I 1 – 18

Introduction
1.1 Bacteria
1.2 Bacterial classification based on metabolisms
   1.2.1 Types of bacteria based on metabolism
1.3 Reproduction and growth
1.4 Modern bacterial classification and identification
1.5 Association with other organisms
1.6 Pathogens
   1.6.1 Mechanisms of bacterial pathogenicity
   1.6.1.1 Bacteria cause pathogenesis by two broad mechanisms
   1.6.1.2 Pathogenicity islands
   1.6.1.3 DNA repeats (simple sequence repeats (SSRs))
   1.6.2 Treatment of bacterial infection and quick resistance development
1.7 Non-pathogenic beneficial bacteria
1.8 Importance of bacteria
1.9 Bacterial evolution
   1.9.1 Evolution is based on three principles
   1.9.2 Evolution of bacteria due to sharing of genetic material across species
   1.9.3 Co-evolution of bacteria with their hosts
1.10 Genomics & genome sequencing
1.11 Drug discovery
   1.11.1 Drug targets
1.12 Drug design
   1.12.1 Two major types of drug design
   1.12.1.1 Ligand-based drug design
   1.12.1.2 Structure-based drug design
1.13 Study

Chapter-II 19 – 43

UniDrug-Target: A Computational Tool to Identify Unique Drug Targets in Pathogenic Bacteria
Abstract
2.1 Introduction
2.2 Materials and methods
   2.2.1 Data collection and generation
   2.2.1.1 Proteomes information
   2.2.1.2 Enzymes, reactions, pathways and metabolites data
   2.2.1.3 Essential genes
   2.2.1.4 Choke points
   2.2.1.5 Domain information
   2.2.1.6 Active site information
   2.2.1.7 Clustering of protein clusters in bacteria to KO groups
   2.2.2 Algorithms
   2.2.2.1 Identification of chokepoints
   2.2.2.2 Identification of interruption sites/points (enzymes) in metabolic pathways and corresponding end metabolites
2.2.2.3 Identification of metabolic pathways disturbed by inhibition of dihydroxypirimidine dehydrogenase (NADP+) of \textit{M. tuberculosis}

2.2.2.4 Identification of variation at functional sites to determine the site specific uniqueness in pathogen protein sequences

2.2.3 Server architecture and implementation

2.2.3.1 Client interface

2.2.3.2 Server

2.2.4 Data representation

2.2.5 Implementation

2.3 Results

2.3.1 Identification of checkpoint reaction efficiency of pathogen-specific proteins

2.3.2 Identification of uniqueness in pathogen-specific proteins at domain level

2.3.3 Identification of unique cavity sites in pathogen-specific proteins

2.4 Discussion

Chapter-III 45 – 175

Simple Sequence Repeats Drive Structural and Functional Plasticity in Bacterial Proteomes towards Pathogen Evolution

Abstract

3.1 Introduction

3.2 Materials and methods

3.2.1 Genomic information

3.2.2 Proteomic information

3.2.3 Pathogenic and non-pathogenic classification information

3.2.4 Bacterial taxonomy data

3.2.5 Computational SSRs identification in bacterial genomes

3.2.6 Average SSRs density (ASD) calculation

3.2.7 Mapping of SSRs to their respective proteins and pathways

3.2.8 Construction of bacterial evolutionary trees

3.2.9 Construction of genome circles

3.2.10 Combining evolutionary tree and genome circles

3.2.11 Concentric genome circles for the polymorphic SSRs identification

3.3 Results

3.3.1 Distribution and density of SSRs in pathogenic vs. non-pathogenic bacterial genomes

3.3.1.1 Variation in SSRs distribution

3.3.1.2 Variation in SSRs density

3.3.2 Mapping SSRs to proteins and pathways

3.3.2.1 Clustering of proteins using KEGG orthologus grouping

3.3.2.2 SSRs content variation in proteins/protein clusters between pathogens vs. non-pathogens

3.3.2.3 SSRs content analysis in protein clusters contributing to the mode of pathogenesis

3.3.2.4 Proteins/protein clusters with lengthier SSRs

3.3.2.5 Properties of amino acids encoded by the SSRs

3.3.2.6 Mapping amino acids encoded by the SSRs to respective cavities and domains

3.3.3 SSRs content analysis vis-a-vis bacterial evolution

3.3.3.1 High SSRs containing bacteria are highly evolved bacteria

3.3.3.2 Increase in occurrence of ‘lengthier SSRs’ and ‘proteins with high SSRs content’ in bacterial genomes positioned down the evolutionary tree

3.3.4 Length polymorphisms in SSR loci among bacterial species and strains within a species

3.3.4.1 Properties of lengthier SSRs

3.3.4.2 Repeat unit occurrence frequency in pathogenic and non-pathogenic bacteria

3.3.4.3 Length polymorphisms in SSR loci among bacterial species and strains

3.4 Discussion

3.4.1 Selection of SSRs by pathogenic bacteria
3.4.2 SSR content analysis in proteins/pathways
3.4.3 Preferential proteins by the bacteria
3.4.4 SSRs content analysis of proteins involved in different modes of infection
3.4.5 Hydrophobic residues, domain and cavity regions encoded by SSRs
3.4.6 Occurrence of lengthier SSRs and SSR types in pathogenic and non-pathogenic bacteria
3.4.7 Bacterial evolution w.r.t SSRs

Conclusion 176

Bibliography 177 - 192

Annexure i - lvi