Bonn-Aachen International Center for Information Technology

(B-IT)

University of Bonn

Master Program in Life Science Informatics

Master Thesis

**Landscaping of COVID-19 Clinical Trials for the Discovery of Insightful Patterns on Ethnicities**

Submitted by : Ram Kumar Ruppa Surulinathan

First Reviewer : Prof. Dr. Martin Hofmann-Apitius

Second Reviewer : Prof. Dr. Juliane Fluck

Internal Supervisor : Dr. Marc Jacobs

*A thesis submitted in fulfillment of the requirements for the degree of M.Sc. in Life Science Informatics in the*

Bonn-Aachen International Center for Information Technology (B-IT)

*in collaboration with*

Fraunhofer Institute for Algorithms and Scientific Computing (SCAI)

August 2021

**ACKNOWLEDGEMENTS**

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| 1 | Introduction | 1 |
|  |  |  |
|  |  |  |
|  |  |  |
| 2 | Theoretical Background |  |
|  |  |  |
|  |  |  |
|  |  |  |
| 3 | Materials and Methods |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| 4 | Results |  |
|  |  |  |
|  |  |  |
|  |  |  |
| 5 | Conclusion |  |
|  |  |  |
|  |  |  |
| 6 | Scientific References |  |

**LIST OF FIGURES**

**LIST OF TABLES**

**LIST OF ABBREVIATIONS**

**ABSTRACT**

**INTRODUCTION**

**THEORETICAL BACKGROUND**

2.1 SARS-CoV-2 a novel coronavirus

The Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) a novel strain and fatal coronavirus was first identified in Wuhan city of China in December 2019. The common symptoms include increased body temperature, dry cough, nausea and body pains. The rapid spread of the virus posed a threat of life to the global environment. According to the records maintained by Worldometer, the top ten most affected nations include USA, India, Brazil, France, Russia, UK, Turkey, Italy, Spain and Germany.

The coronaviruses generally is classified under the family Coronoviridae and subfamily Coronavirinae which is subdivided into four genera namely Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus (Mittal et al., 2020). The SARS-CoV-2, a member of Betacoronavirus genera whose sequence is 96% homologous to the bat coronavirus. Its primary reservoir is considered to be bats and transmitted to human beings through an intermediate host called Pangolin (Zhao et al., 2020).

The SARS-CoV-2 also called by COVID-19 is spherical in structure with positively stranded RNA genome packed inside the nucleocapsid protein (N) and enveloped by the membrane glycoprotein protein (M), envelope protein (E), and the spike protein (S). The typical virus lengths between 26.4 and 31.7 kb with the GC content ranging between 32% and 43% thus indicating to be the largest RNA virus (Mousavizadeh & Ghasemi, 2020).

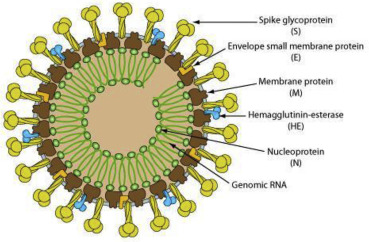


Figure 1 Structure of novel coronavirus , Adapted from (Mousavizadeh & Ghasemi, 2020)