RESEARCH PAPER   OPEN ACCESS

Cardia Mutant: An Android Application for Revealing Information of the Genes Involved in Cardiovascular Disease

Faiza Fatima[1], Zirwa Anwar[1], Zainab Sajid[1], Abdul Jabbar[2], Syed Muhammad Zaigham Zia[3*], Rubina Kousar[1], Fazeela Zaka[4], Sara Siddique[1], Taiyyibah Basharat[5]

[1]*Department of Biochemistry, University of Agriculture, Faisalabad, Pakistan*

[2]*Institute of Microbiology, , University of Agriculture, Faisalabad, Pakistan*

[3]*Institute of Biochemistry and Biotechnology, PMAS Arid Agriculture University Rawalpindi, Pakistan*

[4]*Department of Animal Breeding and Genetics, University of Agriculture, Faisalabad, Pakistan*

[5]*Department of Plant Breeding and Genetics, University College of Agriculture, Sargodha, Pakistan*

Abstract

Mobile devices such as smartphones and tablets have become an integral part of everyday life, due to the rapid development of their hardware and software, and their increased functionalities. Even so, only a few mobile apps have been developed in the field of bioinformatics to date, for providing fast and potent access to sequential data. To facilitate the researchers working on cardiovascular diseases, an offline Android application has been developed and reported here. This app is aimed at providing a user-friendly platform to the researchers for retrieving the genes associated with various cardiovascular diseases and visualizing the reported mutations in respective genes. The application provides both DNA as well as protein sequences of the genes with mutation positions highlighted. Furthermore, online access to the cardiovascular diseases related literature through PUBMED and sequence similarity search though the Basic Local Alignment Search Tool (BLAST) is presenting the app worth using for the researchers in this field.

* **Corresponding Author:** Syed Muhammad Zaigham Zia ✉ zaighamzia4@gmail.com

Introduction

Cardiovascular diseases are the class of diseases that affect the heart, blood vessels, or both. The term 'heart diseases' is considered as 'cardiovascular diseases'. This disease is due to the blockage of blood vessels, chest pain that can lead to the death of patients. The other conditions such as heart muscles, valves, or rhythm are considered as heart diseases. Most of the cardiovascular traits in which multiple genes are involved are complex to describe. Mapping of such genes also requires more information and analysis. The portions of the genome that contain the disease-causing gene are determined by exploring the relationship between the diseases and their associated genetic markers (Amani and Sharifi, 2014).

Years ago, working on the Human Genome Project led to enormous contributions in understanding the genetics of cardiovascular diseases. Before this project, several genes involving Mendelian cardiovascular diseases were characterized (Kathiresan and Srivastava, 2012). These are very simple because only a single gene mutation is enough to cause diseases so, Mendelian diseases are also known as monogenic. Examples are myocardial infarction, QT-long syndrome and hypertrophic cardiomyopathy. However, most of the cardiovascular diseases are polygenic, which may be either inherited from parents to offspring or come from environmental changes (Lvovs *et al.*, 2012). Before the completion of the Human Genome Project, most of the attempts for determining the genetic causes of polygenic cardiovascular diseases were unsuccessful (Ferro *et al.*, 2011).

Cardiovascular diseases are life-threatening diseases throughout the world, with millions of deaths per year. Among cardiovascular diseases, Ischemic heart disease is especially fatal. Further, there are more heart-related diseases with a significant fraction of deaths, such as hypertensive heart disease and inflammatory heart disease. In light of such statistics, we must reduce these numbers, not only in health care environments such as hospitals or primary health care facilities but also in patients' homes and workplaces. To aid in this initiative, it is important to acknowledge the use of smartphones and tablets as devices that have become essential to users in recent years (Kumar, 2014; Shaima *et al.,* 2016).

The aim of the study is that in our application, the user will be able to get a complete genetic basis of his/her disease by just entering the disease name or gene name. This application highlights the mutated region in the DNA sequence of the gene as well as its protein sequence from which a user can easily identify the mutation. This app will have a user-friendly interface. It will be freely available offline. The application will be error-free and accurate data will be provided to the users as the data is taken from papers published in well-reputed journals. The user will also be able to search papers about diseases from PUBMED and can also go to BLAST to find similar sequences. The application will be efficient and faster than the existing systems. So, this app will be a new step in the medical sciences. It will be easily accessible and user-friendly and free to use. Moreover, there is no android application for the mutations of cardiovascular diseases, so it will be a great contribution.

Materials and methods

After a complete study of the previous system, the proposed system has been designed (Gates, 1920). The technical aspects of the new system will be converted into source code to develop the required application. The front-end which will provide a user interface will be developed in XML. The back-end of the project, which will keep records, will be designed in SQLite Server. The application system will be coded in Android Studio. System components typically consist of hardware, software, data and workers. A system's design consists of specifying components, their attributes and their relationships. The problem of system engineering is to design and implement a system that meets the needs of system stakeholders.

For the system development process, the proposed system will use the application development life cycle

or System Development Life Cycle (SDLC). It is a sequential design process, in which a development process proceeds through the phase of communication, planning, analysis, design, testing, deployment and maintenance. The process of rising software will, therefore, a series of activities that will build that software. SDLC will distribute in various activities and assigns, as these activities were solicited throughout the software development effort.

System Development Life Cycle (SDLC)

SDLC consists of a series of events that explains which steps are taken for the development of the application. These steps may involve identification of the problem, data collection and analysis and development of the systems, etc. The proposed application will be developed after the analysis of the previous systems, reviewing their properties and drawbacks. Following steps are involved in SDLC:

Identification of the problem

Undoubtedly, for the creation of anything remarkable, it is necessary to know its background. First of all, we searched for already existing applications and software related to cardiovascular disorders and viewed their capabilities and features (Miller, 1994). Most of the developed applications focus on physical conditions and treatments. We also studied different research papers and articles on cardiovascular diseases, their genetics and mutations from different sites and journals. The development of our application is largely based on these literature studies (Zhang and Shmulevich, 2006).

Data collection

Data collection is the second technique in an investigation. Data required for the application is collected from different literature databases (such as PUBMED, MEDLINE, OMIM) and other sources like NCBI, BLAST, etc. The data includes various cardiovascular diseases, their associated genes, their positions, sequences, functions, mutations and other details. The DNA and protein sequences are taken from NCBI. The positions of mutations are recorded from papers published in well-reputed journals. This

data is arranged in a database developed on the SQLite server which is then used in the application.

Designing

After data collection, the most creative and challenging phase of the system life cycle is system design. In this phase, the specific target mobile platform is decided. It is decided that whether the mobile application will be available online or offline and whether it will be available as a free version or trial version. During the designing process, all the requirements for the accomplishment of our idea are taken into account. Moreover, it is assumed that a user-friendly interface is designed. For this purpose, a prototype is first developed and tested.

Modeling Language (UML) Unified

The unified modeling language is a standard language to represent architecture modules and data flow representation of a system. UML provides the precise notation that we need when modeling the software system.

Use-case diagram

A use case diagram is a graphical depiction of a user's interaction with the different elements of the system. A use case diagram can define the boundary of the system, different types of users (actors), different types of use case relationships among actors and use cases.

Class diagram

Class diagram also display relationships such as containment, inheritance and associations. Another common relationship in the class diagram is a generalization. A generalization is used when two classes are similar but have some differences.

Development

Based on all the above levels, the application is developed using different languages, servers and tools. This process involves the complete coding for the application. The major platform for android development is Java and Android Studio. The minimum set and the most efficient set of interface components should be used to design the layout so

that the application can support as maximum mobile operating system platforms as possible (Lakhanpaul and Jehan, 2006).

The hardware and software requirements for developing the system are the following:

Hardware requirements

As we are developing our application in Android Studio, the minimum hardware requirements for Android Studio in the computer system are the following:

Processor: Microsoft® Windows® 8/7/Vista/2003 (32 or 64-bit).

RAM: 4GB Minimum, 8GB recommended.

Hard Disc: 250GB Minimum.

Screen Resolution: 1280×800 Minimum.

Software requirements

Android Studio: Android Studio is the official integrated development environment (IDE) built on IntelliJ IDEA software which is designed for android application development. The First Android device came to the market in 2008. It is openly available for Windows, macOS and Linux systems. It provides the fastest tools for building android applications on all types of android devices. It enables the programmer to update the code during the running of the application without the need of restarting the application or creating a new package file. Android studio has a code editor that helps us in writing the code more efficiently and faster by giving suggestions and make this productive. The virtual device of the android studio called Emulator helps the programmer to run the app faster than real devices and allows the testing of their app on various devices. Android studio has modules that allow dividing the code into different sections according to their functionality and testing them independently. Android studio is designed for teams, so every member will remain connected with the project and can make changes or updates (Gan, 2016).

Android becomes a fascinating platform for customers, developers and users. Android OS is philosophically different from Apple's iPhone operating system (iOS) in many ways. Android has a higher market share worldwide. Android mobile and tablet devices have a 62.94% market share and Apple's iOS mobile and tablet devices have 33.39% market share since March 2017.

Java: Java is a high-level, dynamic and interpreted programming language. It was first developed by Sun Microsystems in 1995. The first Java version was released under the name of Java 1.1 in February 1997. Java versions have the nickname Java Development Kit (JDK). Until now JDK 1.9 or Java 9 is released. Standard Development Kit (SDK) is an Android software developer's kit that is used by mobile application developers to develop Android mobile applications using Java.

It is a front-end tool used for making webpages interlinked and provides different online programs. Almost all web browsers have a built-in JavaScript engine. Java engines are incorporated in many types of web servers and non-web programs (such as PDF Software) and in the run-time environment to be used for writing mobile and desktop applications. Java is a prototype-based language with first-class functions making it a multi-paradigm language supporting object-oriented programming (Alhussain *et al.*, 2012). XML: It stands for Extensible Markup Language. It is a markup language that has its standards to encode documents in a format that is easily understandable by humans as well as machines. It was developed by World Wide Web Consortium. XML provides simplicity, generality and usage over the internet. It has a text-based data format that is mainly used for the data structure.

Numerous formats are based on XML such as RSS, SOAP and XHTML. XML formats are the basis for many office-productivity tools such as Microsoft Office and Apple's iWork. Apple is implementing an XML-based registry. XML is also a common source for exchanging data over the internet. In Android, XML is used for designing layouts. For app development in Android, Java is for handling and interpreting the data and XML is used to design the

front end of the app known as User Interface (UI) (Lambole *et al.*, 2017).

SQLite Database: SQLite is a transportable, in-process, open-source and lightweight relational database. It stores large and complex data into large tables. It can execute complex queries that combine the data from multiple tables to create views. SQLite is lightweight in terms of setup complexity and resource usage. It is serverless means no configuration requirements. It can directly access its storage files from the disk. SQLite library size is less than 500kb. It requires minimum stack space and heap storage, these features make it a popular database engine for short data storage devices such as mobile phones, Personal Digital Assistant (PDA) and MP3. Android includes SQLite database engine for mobile apps. It supports features of most of the query languages such as SQL (Jan, 2015).

Database development

Every organization has some information needs. Organizations are required to store the data. These data can be store on various media and in different formats like storage on the hard drive, on electronic spreadsheets, or in databases (Alexandar, 2015).

The data is stored in databases in the form of tables. These tables are used to keep records of different queries to get appropriate reports needed by different users. We have developed our database in SQLite server which is then imported into the android application (Liu *et al.*, 2011).

Following is the detail of tables designed for our application:

Testing

After the development of the android application, it is tested to check its functioning and reliability. The testing is performed on various virtual and real devices such as smartphones, tablets, android wear and android TV devices. The Android Studio has its Emulator (virtual device) in which testing of the app can be performed.

Deployment

After going through various tests and feedback from the users, the application is ready to be launched. Before deployment, all the comments and other log files are removed from the application. Finally, the application is uploaded to the application store so that it can be used by the public.

Maintenance

Maintenance of the application is very important. The application is updated as any new information arrives related to the application. Improvements are also made upon the feedback of the users. Marketing of application, advertisement, and highlighting its unique features are also a part of the maintenance of the application.

Results and discussion

Animated splash screen

An animated splash screen is the first screen that is opened when the user opens the app (**Fig**. 4) It is front-page which disappears after some time and displays a new screen.

First activity

The main activity contains a brief introduction to the application. Users can move to the next activity by clicking on the "Get Started" button. It is shown in Fig. 5.

Second activity

The second activity allows the user to perform a search in two different ways i.e. search by a gene name, and search by a disease name (Fig. 6). Based on this choice, the user moves to the next activity.

List of diseases

If the user chooses to search by disease, then a list of diseases is displayed in the next activity. This list contains 20 cardiovascular diseases (Fig. 7).

The user can either select any disease by clicking on the disease name in the list (Fig.17 below describes of selecting disease name) or he can type any disease name in the search bar.

Table 1. Shows stores data about diseases and the corresponding genes.

Table 2. Shows stores data about DNA and protein sequences.

Disease detail

After selecting a disease name, next activity is opened which contains the names of one or more genes involved in that disease. By clicking on any gene name, the user can visualize the mutations in its DNA (Fig. 9) and protein (Fig. 10) sequences. A description of the protein positions and their types is also added in this activity.

Table 3. Shows the storage of data about positions of DNA and protein mutations.

	CODE	DNA_POSITION	ROTEIN_POSITIC	DESCRIPTION	M_CODE	
	Filter	Filter	Filter	Filter	Filter	
1	1	284	284	DNA ...	1	
2	2	609	67	DNA ...	2	
3	3	192	62	DNA ...	3	
4	4	124	42	DNA ...	4	
5	5	836	279	DNA ...	4	
6	6	922	308	DNA ...	4	
7	7	1691	506	DNA ...	5	
8	8	1241	414	DNA ...	6	
9	9	1441	481	DNA ...	6	
10	10	1444	482	DNA ...	6	
11	11	1727	576	DNA ...	7	

List of genes

If the user chooses to search by gene, then a list of genes is displayed in the next activity. The user can either select any gene by clicking on the gene name in the list (Fig. 11) or he can type any gene name in the search bar.
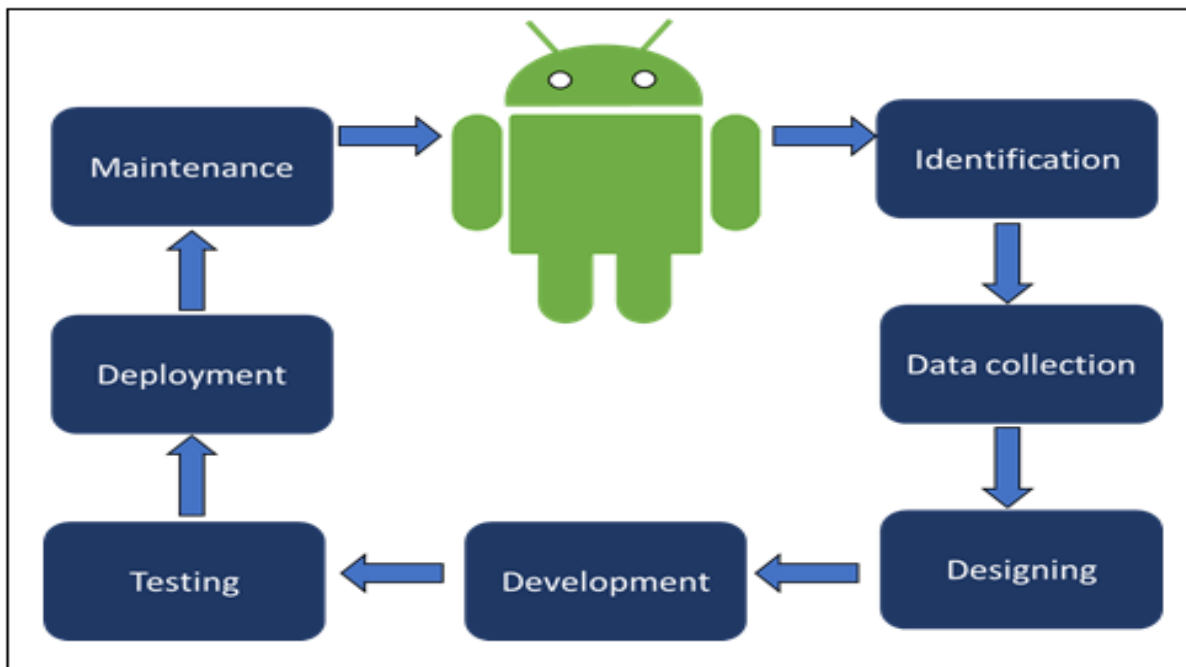
Fig. 1. App development cycle.

Gene detail

After selecting a gene name, the next activity contains the DNA and protein sequences of that gene. These details are shown in Fig. 12 and Fig. 13 respectively.

This activity also allows the user to perform BLAST on the DNA or protein sequences of any gene to find similar sequences.
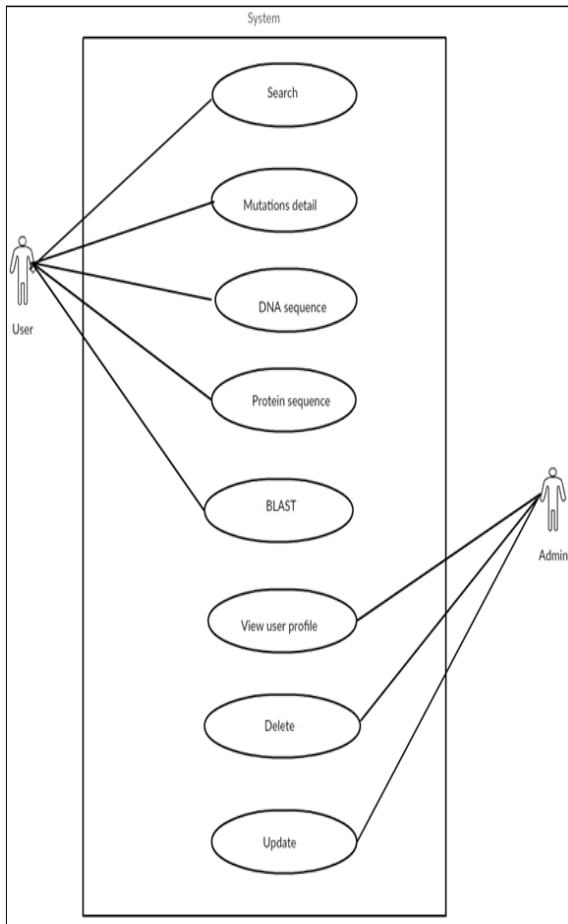
Fig. 2. Use case diagram for the system.

BLAST Tool

Online access to the BLAST tool is also provided in the application to search for similar sequences (Figs 14 & 15).

When the user clicks on the "BLAST" button, then a new window is opened in which the user performs BLAST on the sequence from our application.

Moreover, the user does not need to manually copy the sequence because it is copied by the app. The user just needs to click "paste" in the BLAST.

PUBMED Tool

The important feature of this app is that it provides online access to PUBMED, which is the most widely used platform to search for different research papers and review articles. So, if the user wants to search for an article while using the app, he will be able to do so with just one action. In our application, if the user clicks on the "PUBMED" button then the PUBMED homepage is opened in a new window where he can perform a search (Figs 16 & 17)).
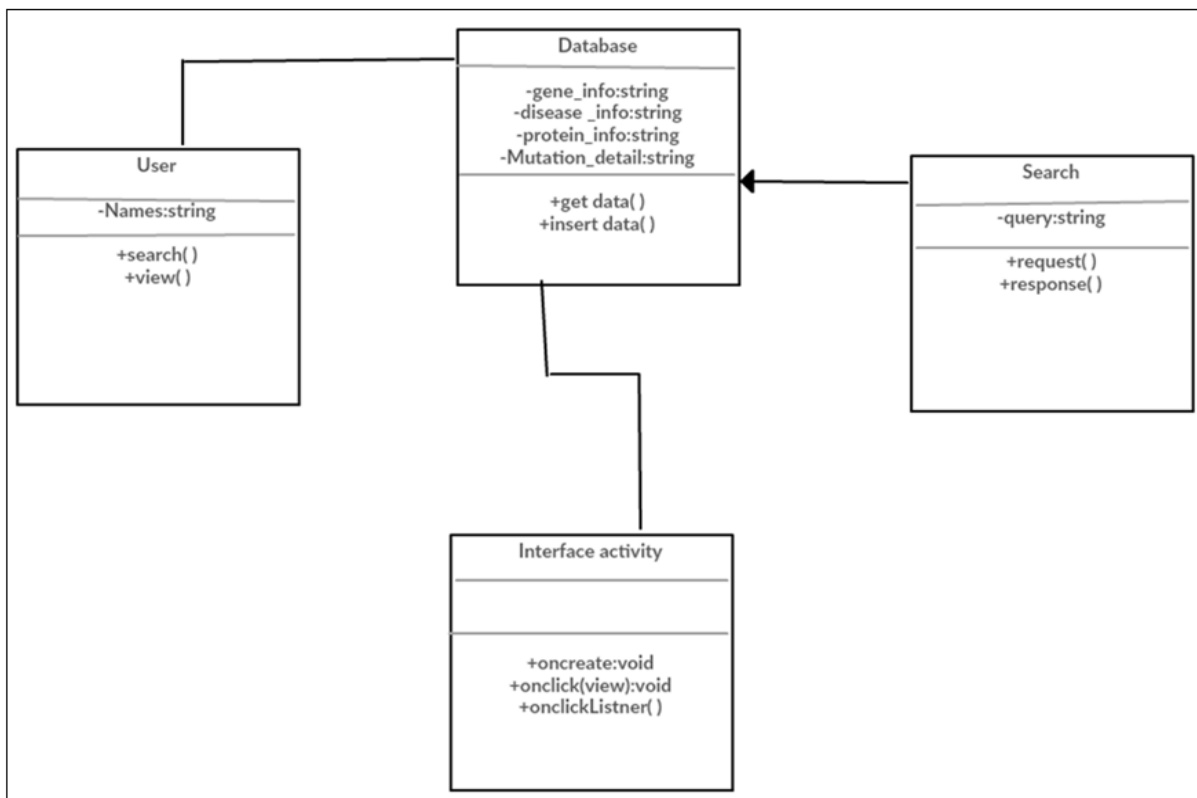
Fig. 3. Class diagram for system.

After having a thorough study of cardiovascular diseases, we have come to know that there is a lot of literature available on these diseases.

Fig. 4. Frontpage.

However, there are few databases for these diseases. Moreover, these databases contain knowledge for some specific diseases. For example, the CAD gene is a database that contains the knowledge of coronary artery diseases only.
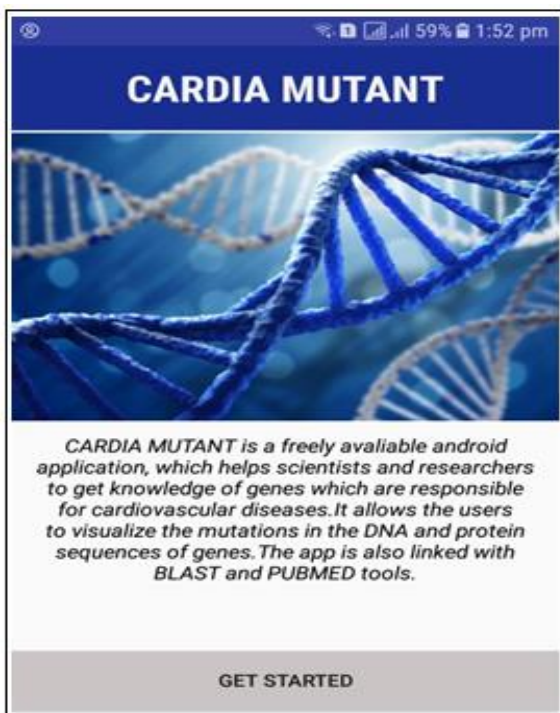

Fig. 5. First activity.

It contains information such as gene expression, gene annotation, gene pathways, protein-protein interactions, etc.
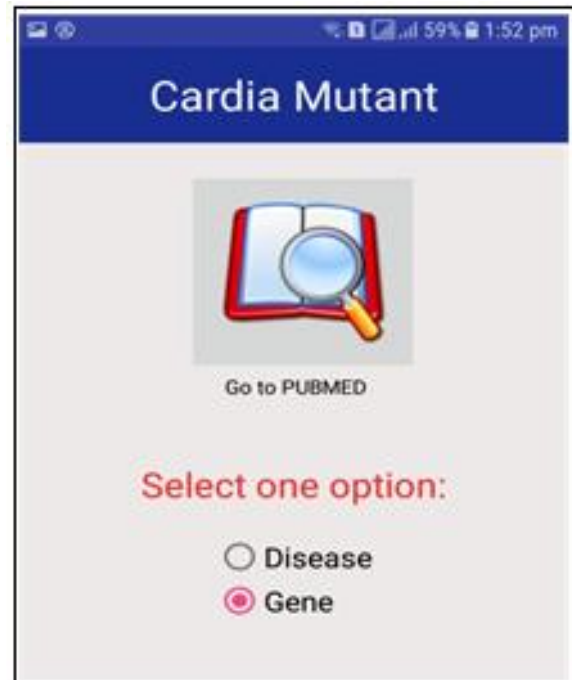
Fig. 6. Second activity.
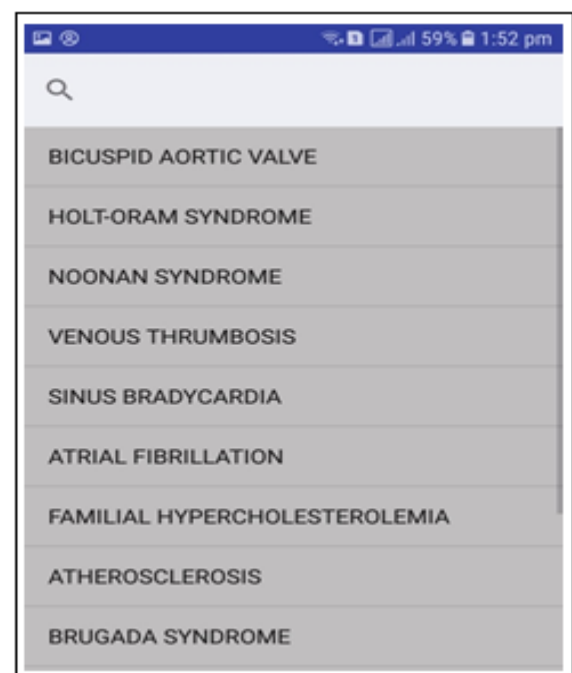
Fig. 7. List of diseases.

There is another database named CardioGenBase which covers major cardiovascular diseases. It contains information such as literature evidence, ontology, pathways, single nucleotide polymorphism, protein-protein interaction network, normal gene

expression, protein expressions in various body fluids and tissues.

It also contains some tools like gene-disease association finder and gene expression finder.
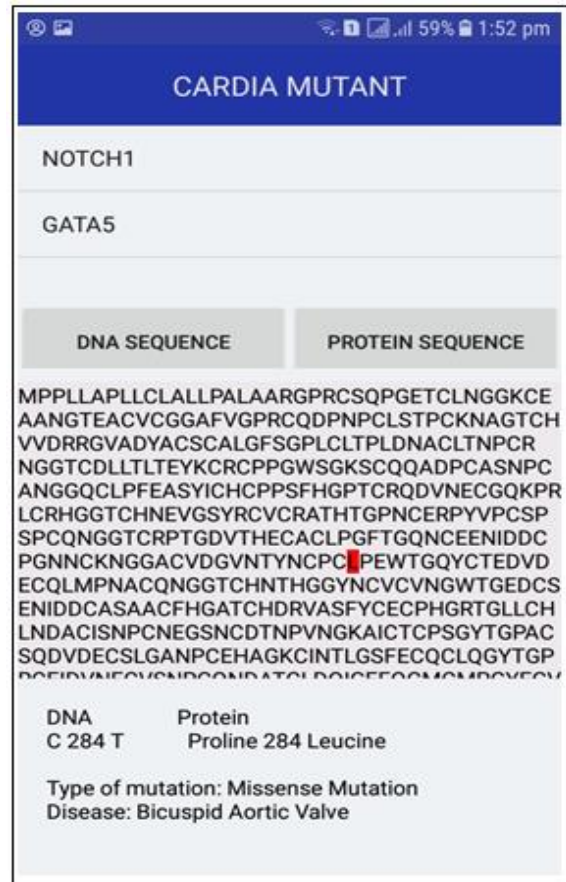

Fig. 8. Genes.
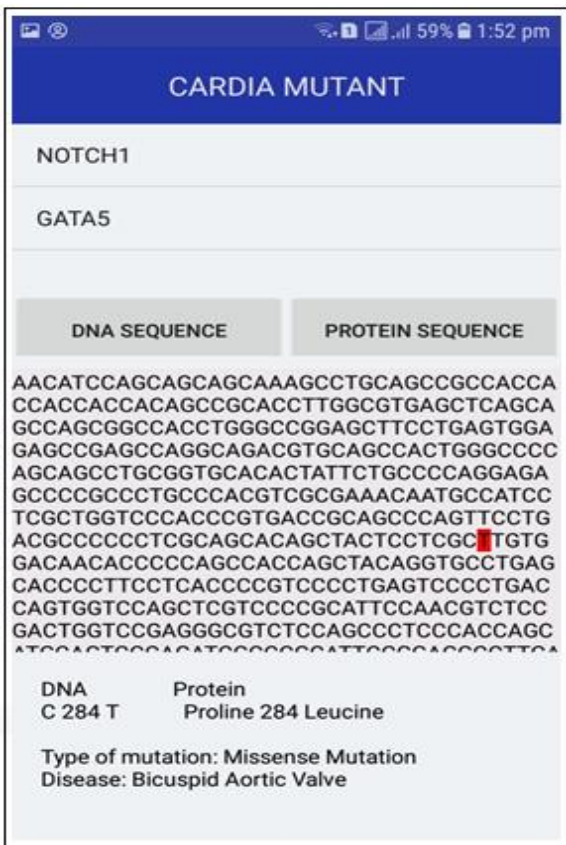

Fig. 10. Protein mutations.
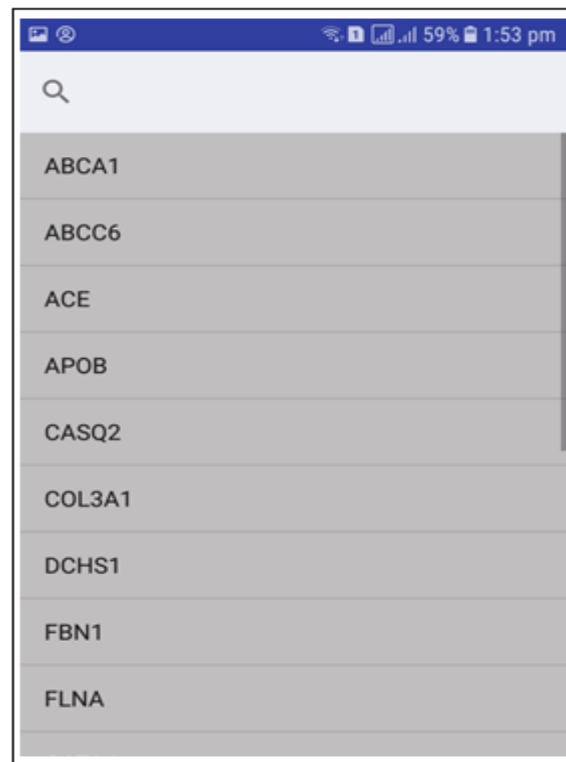

Fig. 9. DNA mutations.

Fig. 11. List of genes.

However, our application is different from these databases in the sense that it contains information about the mutations of genes involved in most common cardiovascular diseases with their DNA and protein sequences.
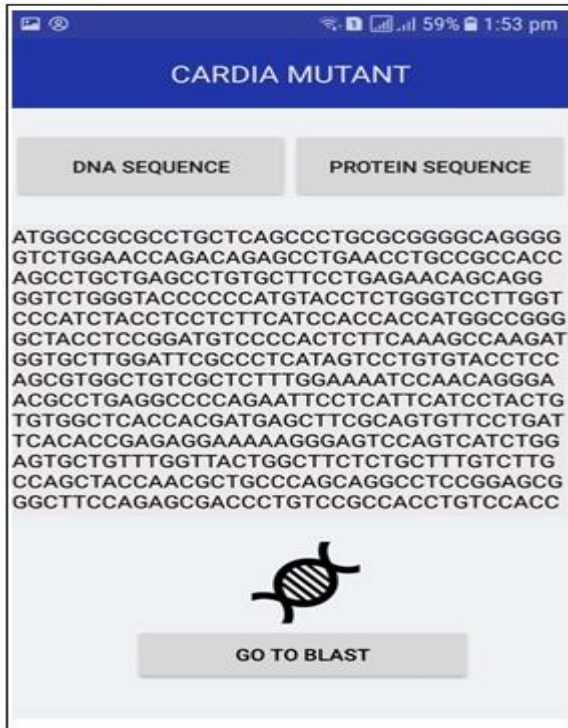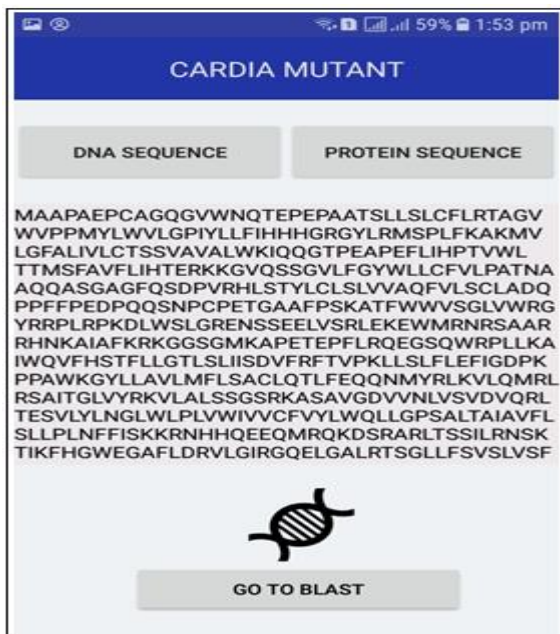

Fig. 12. DNA sequences.

Fig. 13. Protein sequences.

Moreover, computer availability is not possible everywhere while everybody has android phones in their pockets nowadays. So, the android application

will be an easier source of information than databases. Besides, our application is offline, so it can be used everywhere whether there is internet accessibility or not.

Fig. 14. BLAST tool, web view.

With the advancement in technology, the number of android applications is increasing day by day. However, very little work is done in the field of medical sciences. Most of the apps contain information about the symptoms, diagnosis and precautions of diseases. Moreover, there is no specific android application that covers multiple cardiovascular diseases.
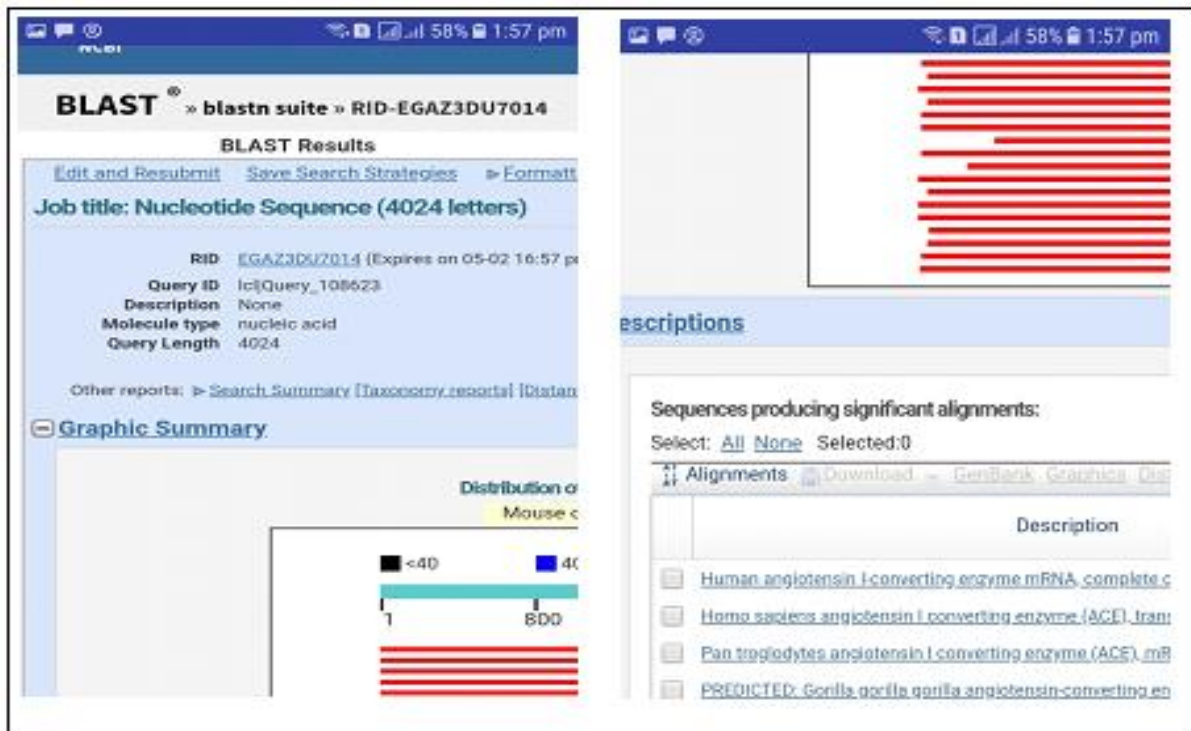
Fig. 15. BLAST Results.

Then we realized that there should be an android application that will help understand the genetic causes of cardiovascular diseases. So, we developed an android application that will be very useful for the researchers working on cardiovascular diseases.
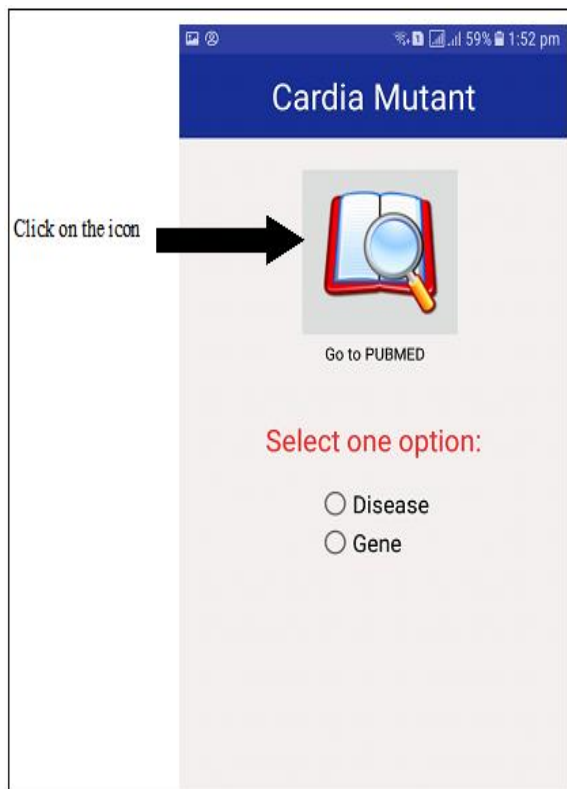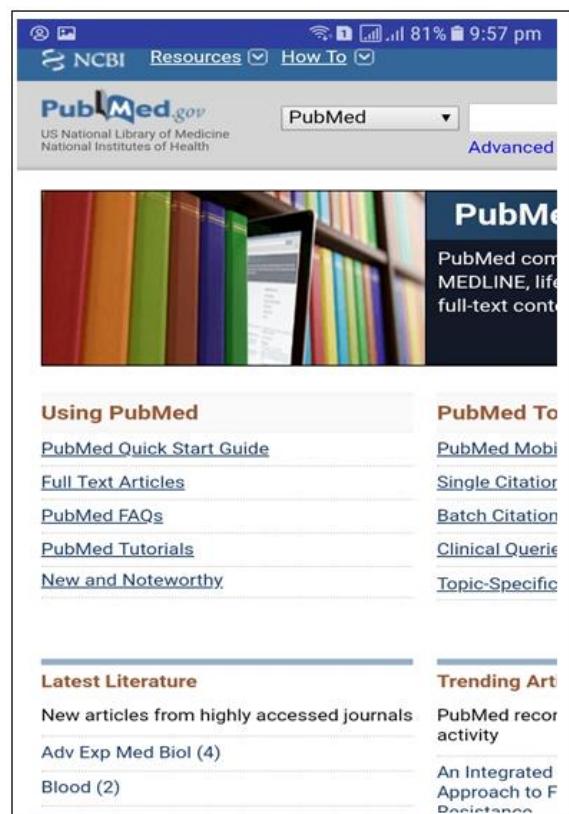

Fig. 16. PUBMED tool.

Fig. 17. PUBMED tool, web view.

Conclusion

Nowadays technology is becoming advanced day by day, and android users are everywhere. This is the

reason we have developed an android application that would help researchers and the medical community to find data about cardiovascular diseases caused by mutations in genes. The mobile app provides users amazing features like the ability to search a gene, search literature through PubMed, BLAST any protein or DNA sequence, and find SNP's location and much more. It is an offline app which means there would be no waiting for an internet connection to get data from our app. We have used different databases like NCBI, OMIM, SNP, and PUBMED to collect literature data about cardiovascular diseases for our App. These databases have been used to make sure that the information we are going to provide to our users is reliable and accurate.

References

Amani R, Sharifi N. 2014. Cardiovascular Disease risk factor. The Cardiovascular System-Physiology, Diagnostics and Clinical Implications **14,** 280-310.

Kumar A. 2014. Cardiovascular diseases: Are we overlooking some cardiovascular diseases risk factors/markers? Journal of Biomedical Sciences **3,** 1-3.

Shaima C, Moorthi PV, Shaheen NK. 2016. Cardiovascular diseases: Traditional and non—traditional risk factors. Journal of Medical and Allied Science **6,** 46-51.

Feero WG, Guttmacher AE, O 'donnell CJ, Nabel EG. 2011. Genomic Medicine Genomics of Cardiovascular Disease. The New England Journal of Medicine **365,** 2098–109.

Sekar Kathiresan, Deepak Srivastava. 2012. Genetics of Human Cardiovascular Disease. Cell **148(6),** 1242–57.

Gates RR. 1920. Mutations and Evolution. New Phytol **19(3–4),** 64–88.

Zhang W, Shmulevich I. 2006. Computational and statistical approaches to genomics. Comput. Stat. Approaches to Genomics 1–416.

Lakhanpaul Jehan. 2006. Single nucleotide polymorphism (SNP)– Methods and applications in plant genetics : A review.; (October 2006).

Liu H, Liu W, Liao Y. 2011. CADgene: A comprehensive database for coronary artery disease genes. Nucleic Acids Research **39(1),** 991–6.

Alexandar V, Nayar PG, Murugesan R. 2015. CardioGen base: A literature based multi-omics database for major cardiovascular diseases. PLoS One **10(12),** 1–10.

Miller JG, Chapman CM, Bruenger JA. 1994. CARDIAX .: Computer-Aided Review of Diagnosis Involving. 4210.

Gan SKE, Poon JK. 2016. The world of biomedical apps: their uses, limitations, and potential. Sci Phone Apps Mob Devices [Internet] **2(1),** 6. Available from: http://scientificphoneapps.springeropen.com/articles/10.1186/s41070-016-0009-2

Alhussian H, Zakaria N, Hussin FA, Bahbouh HT. 2012. A review of the current status of the Java programming on embedded real-time systems. 2012. Int Conf Comput Inf Sci [Internet] **2,** 836–42. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6297142

Lambole PR, Chatur PN. 2017. A review on XML keyword query processing. IEEE Int Conf Innov Mech Ind Appl ICIMIA 2017 - Proc (Icimia) 238–41.

Jan I, Shafi J. 2015. Sqlite- A Better Choice as an Embedded Database for Smart Phones. International Journal of Allied Practice, Research and Review **2(6),** 9–15.

Lvovs D, Favorova OO, Favorov AV. 2012. A polygenic approach to the study of polygenic diseases. Acta Naturae **4,** 59-71.