RNA_Seq Viewer: A Mobile App for Displaying NGS Gene Expression Data

Chien-Yuan Chiu, Yu-Chieh Lin and Wen-chang Lin

Institute of Biomedical Sciences, Academia Sinica

Taipei, Taiwan, R. O. C.

e-mail: jychiu_tw@yahoo.com.tw, joycelin18@gmail.com, wenlin@ibms.sinica.edu.tw

Abstract— In recent years, Next Generation Sequencing (NGS) technology has dramatically advanced genome researches and biomedical sciences, which generates and accumulates a large amount of biological data rapidly. Nowadays, there is a rapid expansion of tablets and smart phone devices, replacing the traditional personal computers and notebooks commonly used for information retrieval and display. However, there are little bioinformatic related mobile applications developed specifically for NGS data visualization on mobile devices. To demonstrate the feasibility of displaying large NGS data using the mobile devices, we designed and implemented a new iOS App - the RNA_Seq Viewer, for the visualization of The Cancer Genome Atlas (TCGA) gene expression datasets. This iOS App could efficiently display gene expression information systematically over the human chromosome framework. We have processed over 2,500 human cancer patients in nine cancer types retrieved from the TCGA web site. This mobile App could be potentially utilized in future personalized medicine applications by serving as the essential visualization core component to easily access the personal genomic and medical information through the cloud based data warehouses.

Keywords-bioinformatics; NGS; App; TCGA; RNA-Seq.

I. Introduction

In recent years, genome sequence information has been accumulated rapidly thanks to the completion of human genome project and the development of Next Generation Sequencing (NGS) platform technologies. Therefore, many bioinformatic databases and tools are generated for the visualization and analysis of massive genomic sequence information [1]. It is not an easy task to efficiently display vast amount of biological annotation information in addition to the genomic sequences and chromosomal organizations using web browsers. There are several renowned web-based bioinformatic tools and databases to accomplish such challenges, including UCSC Genome Browser [2] and Ensembl [3]. However, over the last few years, it is worth noting that tablets and smart phone devices (such as iPad and iphone) are widely popularized and disseminated on the World, replacing the traditional personal computers. These devices utilize the multi-touch technology as the core user interface, and gradually transform the user experience and machine interface design.

The computer software development sector is also heavily affected by this trend, with the rapid expansion of mobile applications (Apps) for mobile devices. Many developers have been concentrated on the App developments in addition to the traditional desktop software development. Few tablet Apps were created to retrieve and display text based genome annotation information [4]. However, there

are no bioinformatic mobile applications developed specifically for the visualization of huge genome information as well as the NGS sequence data. Since the widely adaptation of handheld smart devices, there become the common platform for daily information retrieval and exchange. Therefore, we attempt to develop such novel mobile application software for mobile devices and demonstrate the feasibility of visualization of genome information on the mobile devices.

We first focused on the visualization of The Cancer Genome Atlas (TCGA) data. TCGA is a comprehensive and coordinated effort to accelerate the understanding of the molecular basis of cancer through the utilization of large scale sequencing experiments using NGS platforms [5]. We have retrieved NGS RNA-Seq cancer gene expression information from thousands of cancer patients in various cancer types collected in TCGA project. Our RNA_Seq Viewer mobile application is designed to display the RNA-Seq gene expression information systematically for easy interrogation and visualization.

Herein, in Section II, the development environment and data source were introduced. In Section III, the represented screen displays were demonstrated and basic functionalities were briefly described. Finally, we discussed the future improvements of this new App in Section IV.

II. Materials and Methods

The development and implement of RNA_Seq Viewer was done using the Apple development software, Xcode 5.02. The programming language used was Object-C and database used was SQLite. All TCGA cancer level-3 RNA-Seq data were retrieved and gene expression RPKM (reads per kilobase per million) values were processed for SQLite tables. We obtained a total of nine different cancer types with over 2,500 cancer samples. The RNA_Seq Viewer App can be obtained freely through Apple iTune App Store. Additional information can be accessed from [6].

III. Results

RNA_Seq Viewer is a mobile application built for iOS to provide biologists a new user experience in interrogating large-scale NGS data intuitively on the iPad or iPhone. Users can further interrogate differentially expressed genes by interrogating differentially expressed regions from the overall chromosome display view and quickly zooming into regions of particular interest.

The RNA_Seq viewer displays gene expression information (RPKM value) on the chromosome level. Initially, an illustration of the 22 autosomes and two sex chromosomes are displayed (Figure 1). The navigational and

functional buttons are arranged on the bottom row. Users should first select the cancer type of interest by clicking the list button on the left. Detailed information includes the TCGA sample ID and sex/age, as well as cancer stage information on the bottom line (Figure 1).

Once selected, the cancer gene expression information of an individual patient is loaded and displayed. Since the matching normal tissues were not always available in TCGA, we used the average expression values from all available normal tissues for that particular cancer type to represent the background expression levels. The gene expression value of cancer tissue is illustrated by a blue line, and the expression value of normal tissue is illustrated by the overlapping gray line in the background. User then can zoom in and out of the screen to interrogate the regions of their interest (Figure 2).

By adjusting the intensity bar on the right of the function row, the user can dim the blue line to a lighter color, which allows users to compare the expression difference between cancer and normal tissues, especially on the overexpressed genes in cancer. Once a region of interest is identified, users can click on the gene(s) of interest to activate the red-pin icon and show detailed gene information and expression values (Figure 2). The gene name/RefSeq ID and expression values are now displayed before the patient's sample information. The RPKM values of tumor (T) and normal (N) tissues are displayed (Figure 2). Users can then click on the NCBI (National Center for Biotechnology Information) button to retrieve the detailed gene information.
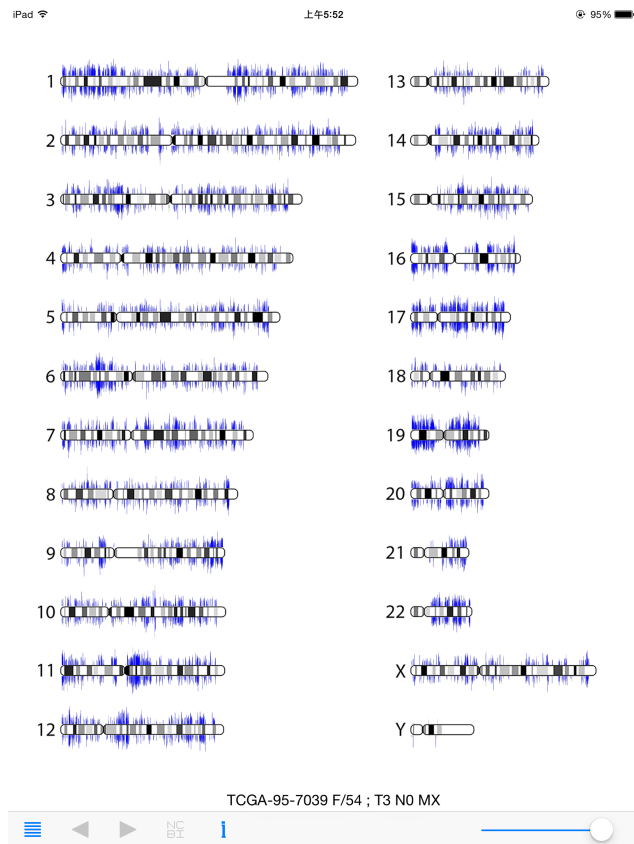

Figure 1.   User interface of RNA_Seq Viewer (global view).
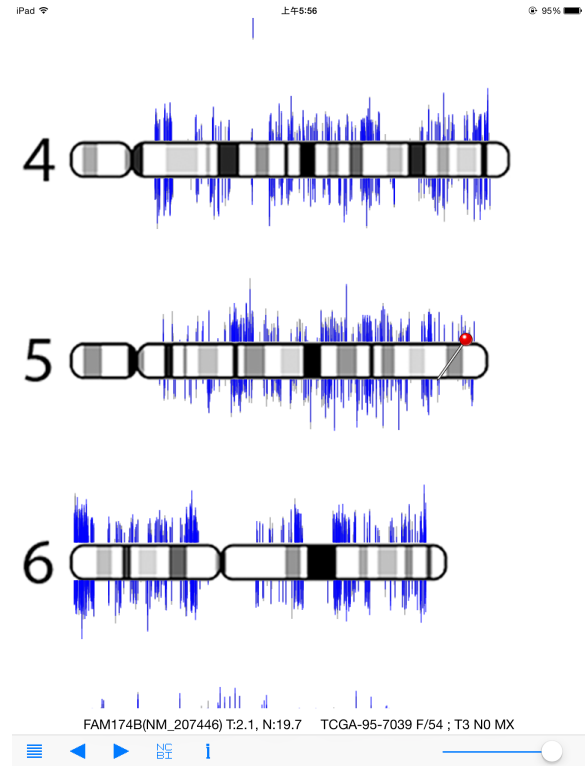

Figure 2.   User interface of RNA_Seq Viewer (detailed view).

Users can also submit their very own RNA-Seq data to display in our App here. Users simply obtain the user.agv db file from the web site and fill in the RPKM values of the normal and cancer tissues using any SQLite tools before synchronization to the mobile devices.

IV.   DISCUSSION

As proof of concept, we demonstrated that RNA-Seq data could be visualized and interrogated efficiently on the mobile devices. This mobile App is useful for displaying multiple genome data onto global chromosome context. In the future, it can be expanded to integrate additional NGS data and developed for Android platform. This App could be utilized as a fundamental visualization component for personalized genome information display and retrieval.

REFERENCES

[1]   S. C. Li, et al., "Identification of homologous microRNAs in 56 animal genomes," Genomics, vol. 96, Jan. 2010, pp. 1-9.

[2]   D. Karolchik, et al., "The UCSC Genome Browser database: 2014 update," Nucleic Acids Res., vol. 42, Jan. 2014, pp. D764-D770.

[3]   P. Flicek, et al., "Ensembl 2014,", Nucleic Acids Res., vol. 42, Jan. 2013, pp. 749-755.

[4]   C. Wu, I. Macleod, and A. I. Su, "BioGPS and MyGene.info: organizing online, gene-centric information," Nucleic Acids Res., vol. 41, Jan. 2013, pp. D561-D565.

[5]   Cancer Genome Altas Netwrok, "Comprehensive molecular portraits of human breast tumours," Nature, vol. 490, Oct. 2012, pp. 61-70.

[6]   http://tdl.ibms.sinica.edu.tw/NGS_viewer_1/NGS_Viewer/General_Introduction.html.