UNIVERSITY OF CALIFORNIA,
IRVINE

Acquisition, processing and analysis of electrocardiogram in awake zebrafish

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SIENCE

in Electrical Engineering

by

Tai Le

Thesis Committee:
Assistant Professor Hung Cao, Chair
Assistant Professor Zhou Li
Professor Peter Burke

2021

DEDICATION

To

my parents, friends and my wife

who always support and encourage me

Unless you try to do something beyond what you have already mastered, you will never grow.

Ralph Waldo Emerson

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my committee chair, Assistant Professor Hung Cao. I first met him in a seminar which he presented his research work in HUST. Since then, we have worked together for several years. I feel so fortunate to have an opportunity to be his student, receiving invaluable support in both academia and daily life. Furthermore, Dr. Cao's laboratory is a highly collaborative and friendly environment. We have lab dinners and social events together often. His laboratory has provided me opportunities to work with other talented students and has cultivated in me a passion to continue working and excelling in the field of electrical/biomedical engineering. Without his guidance and persistent help this MS thesis would not have been possible.

I would like to thank Professor Xiaolei Xu (Mayo Clinic) and Michael P. H Lau – president at Sensoriis, Inc for their enthusiastic support in both research experience and financial support.

I would like to thank my committee members, Professor Zhou Li and Professor Peter Burke who are very supportive and encourages me to finish the work.

I would also like to extend my gratitude to our wonderful team members in the Hero Lab. Their enthusiasm and friendliness significantly support me to finish the work and they are by my side in many memorable moments. Without them, I could not imagine I can finish this work.

Finally, I would like to acknowledge the financial support from the NSF CAREER Award #1917105 under Professor Hung Cao, the NIH SBIR grant #R44OD024874 under Professor Hung Cao and Michael Lau, and the NIH HL107304 and HL081753 under Professor Xiaolei Xu. I also thank Lauren Schmiess-Heine for managing the zebrafish facility.

ABSTRACT OF THE THESIS

Acquisition, processing and analysis of electrocardiogram in awake zebrafish

by

Tai Le

Master of Science in Electrical Engineering

University of California, Irvine, 2021

Assistant Professor Hung Cao, Chair

The zebrafish model has been demonstrated as an ideal vertebrate model system for a diverse range of biological studies. Along with conventional approaches, monitoring and analysis of zebrafish electrocardiogram (ECG) have been utilized for cardio-physiological screening and elucidation. ECG monitoring has been carried out with fish treated with anesthetic drugs, rendering the short period of time in recording the signals (<5 minutes). In this work, a prolonged sedation system for continuous ECG monitoring of multiple zebrafish was proposed and developed. The design allows the minimization of the effects of anesthetic drugs and temperature variation to zebrafish cardiac signals for experiments that need continuous data acquisition in real time. Moreover, a graphical user interface (GUI) was built in National Instruments LabView platform for real-time recording, processing and analysis. The program provided important features, such as signal de-noising, characteristic wave detection. Further, machine learning-based approach was first utilized for anomaly detection, yielding above 90% of accuracy to classify different types of heart disease with zebrafish ECG data. Last but not least, the device showed a robust ECG acquisition and analytics for various applications including arrhythmia in Na^+-induced sinus arrest,

temperature-induced heart rate variation, and drug-induced arrhythmia in *Tg(SCN5A-D1275N)* mutant and wildtype fish. The multiple channel acquisition also enables the implementation of randomized controlled trials on zebrafish models. Two fish per group may be utilized simultaneously in identical conditions of experiments. Finally, the developed ECG system holds promise and solves current drawbacks to greatly accelerate other cardiovascular studies and drug screening applications using zebrafish.

CHAPTER 1: INTRODUCTION

1.1. Motivation

Cardiovascular diseases (CVDs), are currently the leading cause of death globally, taking an estimated 17.9 million people dying annually which is 31% of all global deaths [1]. Combined effects of population growth, aging of populations, and epidemiological changes in CVDs have resulted in increasing global deaths from CVDs [2]. In the US, about 655,000 people die from heard disease each year which is 1 in every 4 deaths according to the Centers for Disease Control and Prevention (CDC) [3]. Especially, since the COVID-19 pandemic has started in late 2019, it significantly contributes to mortality rate in patients with CVDs of 10 % or more [4]. In attempt to reduce the number of dead people caused by CVDs, the National Institutes of Health (NIH) and other institutions such as the National Science Foundation and American Heart Association have provided multi billions of dollars (e.g., nearly 2.5 billion dollars in 2020 from NIH) in research toward heart disease therapies [5].

The zebrafish serves as an ideal model for cardiovascular studies because of its similar homology to humans in both morphology, physiology, and genetics. Moreover, with small size, low cost for maintenance, short generation time, and optical transparency in embryo stage, zebrafish offers advantages over traditional mammalian models. Despite having only two discernible chambers in the zebrafish heart compared to four in human hearts, the zebrafish heart possesses a similar contractile structure with an analogous conduction system [6, 7]. Additionally, 70% of all human genes have orthologues in the zebrafish genome, indicating its applicability in studying genetic pathways [8]. Therefore, the zebrafish model is appropriate in the study of SSS and the correlation of related genetic pathways to the electrophysical phenotype via ECG. Currently, several research groups have

developed systems to assess zebrafish ECG. Regarding sensor design, conventional needle electrodes are commonly used. Lin *et al.,* [9] designed and tested the needle electrode with different materials, including tungsten filament, stainless steel and silver wire to investigate the recorded signal quality. Along with a portable ECG kit, the authors aim to provide a standard platform for research and teaching laboratories. The needle system is also deployed in other studies [9-12] to conduct biological and/or drug-induced research. Although those demonstrated promising results, , the needles need to be gently inserted through the dermis of zebrafish in order to collect favorable signals. This could cause injury to the fish's heart, thus possibly changing signal morphology [11]. Moreover, it requires an intensive effort in precisely positioning the electrode on the tiny heart to achieve a clear ECG signal. Therefore, several alternative probe systems have been developed, including the micro-electrode array (MEA) and the 3D-printed sensors. Our team and others have demonstrated the use of MEA for acquisition and provided the signal with favorable signal-to-noise ratio (SNR) with high spatial and temporal resolution [13-15]. For instance, we presented a MEA array covering the fish's heart which enables site-specific signals [14, 15]. Cho *et al.,* [16] developed a MEA printed on a flexible printed circuit board (FPCB) based on a polyimide film for multiple electroencephalogram (EEG) recording for epilepsy studies. Although the MEA allows multiple signal recordings, only one fish can be assessed at a time due to the limiting number of channels on acquisition. Moreover, most studies have used bulky and expensive acquisition tools to collect data and then transfer them through a cable to a computer. These require a designated benchtop area to conduct experiments, and they can potentially encounter errors due to instability of cable and connectivity. In the market, commercially available systems, *e.g.* the one from iWORX (Dover, NH) with a compact

amplifier, can provide improved system mobility. However, several challenges have not been resolved, such as **i**) the current systems only record for a short period of time (3-5 mins), which result in inconsistent results among different fish; **ii**) the ECG acquisition requires anesthetized animals, rendering it stressful to the fish and inadequate to provide intrinsic cardiac electrophysiological signals; **iii**) manual one-by-one measurement limits the ability of conducting studies required to test with a large number of fish; and **iv**) signal processing has been carried out offline with exorbitant effort.

1.2. Objectives

In this work, a zebrafish ECG monitoring system with capable of obtaining long-term ECG recordings for multiple fish simultaneously is presented. Two protypes of the system are described in detail. An in-house electronic device is developed, leveraging the Internet of Thing (IoT) capability with wireless data transmission and data processing on a mobile application. This enhances the mobility and versatility of the system as well as support distanced collaborations to conduct research on zebrafish models. The system is validated through numerous experiments, showing its potential with 1) simultaneous measurement for up to 4 fish; 2) continuous ECG recording for up to 1 hour compared to several minutes of other systems; 3) reduction in arrhythmic side effects with the use of 50% lower Tricaine concentration.

Furthermore, we investigate a specific electrophysiological phenotype, namely sinus arrest, induced by sodium chloride on mutant fish *Tg(SCN5A-D1275N)* to demonstrate that our proposed system can be used as a screening tool to detect and elucidate zebrafish cardiac arrhythmic symptoms.

1.3. Thesis outline

Chapter 1 – This chapter presents the current problem with cardiovascular disease as well as the advantages of zebrafish as a model for cardiac studies. Moreover, literature review also is presented in this chapter.

Chapter 2 – The main content in this chapter is to cover ECG acquisition design, and ECG processing and analysis for zebrafish as well as the experimental design. Specifically, in the ECG acquisition system, two prototypes of the prolonged system and in-house electronics system are described in detail. For ECG processing and analysis, the work highlights difficulties of processing the zebrafish ECG, and introduces two platforms to solve the problems, including a graphical user interface (GUI) using LabView platform and machine learning-based approach.

Chapter 3 – This chapter provides results of system implementation with various experiments and characterizations. Some limitations of the current design and discussion on each result are also presented.

Chapter 4 – This chapter summarizes the content of this MS thesis and highlights potential impact of this work to cardiac studies. In addition, the future work also is mentioned, thus suggesting some points to improve the system and experiments.

CHAPTER 2: DESIGNS, METHODS, AND IMPLEMENTATION

2.1. ECG acquisition for awake zebrafish

As aforementioned, current approaches only provide a short ECG measurement with high dose in anesthesia and without proper environment control which leads to side effects in cardiac rhythm. This section provides in detail on two of the ECG monitoring systems for awake zebrafish, including electrode and apparatus development, mechanical system, and electronics design. These attempts are to tackle with such obstacles of current approaches.

2.1.1. Prolonged Zebrafish system – Prototype 1

Microelectrode array (MEA) membrane: To collect the ECG signal from zebrafish, 4-channel MEA membranes for ECG acquisition in adult zebrafish was first developed humans [13-15], yielding high spatial and temporal resolution with favorable SNR. The sensors were based on a polymer substrate, such as parylene C or polyimide so that they can be gently contact the fish's chest as shown in (**Fig. 1d**). The MEA membranes were fabricated by patterning sputtered or e-beam evaporated metals (200 nm Au on 20 nm Cr) by wet etching followed by an encapsulation process by another polymer, leaving only the electrodes and contacts exposed (**Fig. 1a**). The working electrodes (WE) were designed in

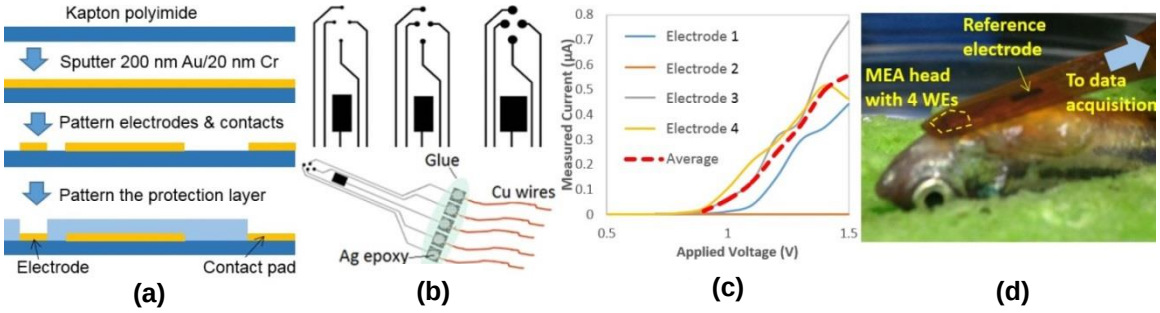


Fig. 2. 1. (a) Fabrication processes of the MEA membranes. **(b)** Different electrode sizes and the complete device. **(c)** Impedance curves of one 300-μm MEA membrane. **(d)** The MEA on the fish.

various sizes and shapes in order to compare the performance. The reference electrode (RE) was much larger to maintain proper electrode-tissue interface. Silver epoxy was used to form electrical connections using thin wires, and then all contacts were protected by glue using a glue gun (**Fig. 2.1.b**). Electrodes with different sizes were characterized using an impedance analyzer, and we chose round electrodes with a diameter of 300 µm for most of the cases owing to the performance and relatively small size. Initially, a Faraday cage on a vibration-free table was preferred [31]; however in order to reduce complications in experimental setup, we conducted experiments in regular lab benches then applied rigorous signal processing to enhance SNR using wavelet filtering and the thresholding technique humans [13-15]. Prior to experiments, to validate connectivity and characterize the performance of the fabricated electrodes, each MEA membrane was immersed in a saline solution while applying a range of voltages between 0 V and 1.5 V across the WEs and RE. The resulting currents were then measured. A result from a 300-µm MEA membrane is showed in **Fig. 2.1.c**. Our MEA membranes have provided a novel tool to study the zebrafish heart; however, measurement was done with fish under anesthesia (**Fig. 2.1.d**), indicating the obtained ECG signals were not intrinsic, which a new version of this MEA with the support of apparatus was developed.

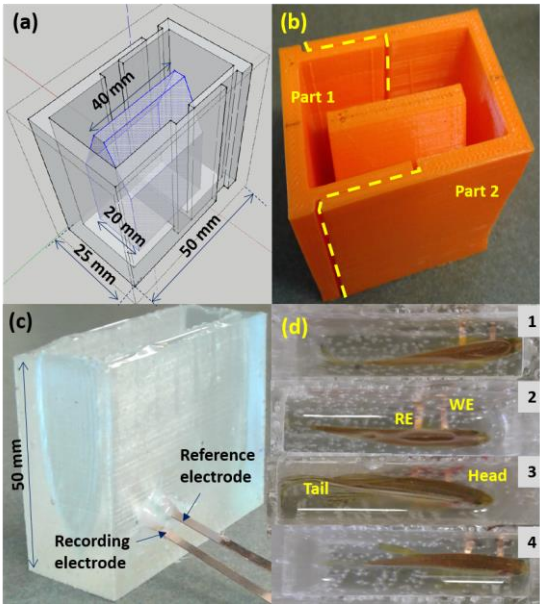


Fig. 2. 2. (a) 3D design in SketchUp software; (b) 3D-printed mold with parts 1&2 assembled; the dashed line shows the boundary of the two parts; (c) PDMS mold with integrated WE and RE electrodes; and (d) 4-chamber apparatus with fish.

Apparatus and 2-channel electrode: The apparatus was designed aiming to keep

zebrafish comfortable, in order to minimize the unwanted effects and prolong reliable recording [17]. Generally, the length of an adult fish body varies from 17 mm to 33 mm [18], thus the housing was designed with a length of 50 mm, a height of 50 mm and a width of 25 mm. For the ease of demolding after curing PDMS, the mold was split into two pieces as shown in **Fig. 2.2.b.** An additional part (in the middle, blue, **Fig. 2.2.a**) with a height of 43 mm and a maximum width of 20 mm was used to make a tapered shape of the housing to minimize the movement of the fish when loaded. First, the curing agent and PDMS monomer (*Sylgard 184, Dow Corning, Midland, MI*) were mixed together with a ratio 1:10 in weight, respectively. Subsequently, the mixed solution was degassed for approximately 1 hour before being poured into the rectangular mold. Then, the middle part was securely placed. After the PDMS was cured by putting on a hot plate under 100^{o}C for 5 hours, it was removed from the rectangular box. The middle part was gently pulled out to make sure the apparatus remains the desired shape (**Fig. 2.2)**.

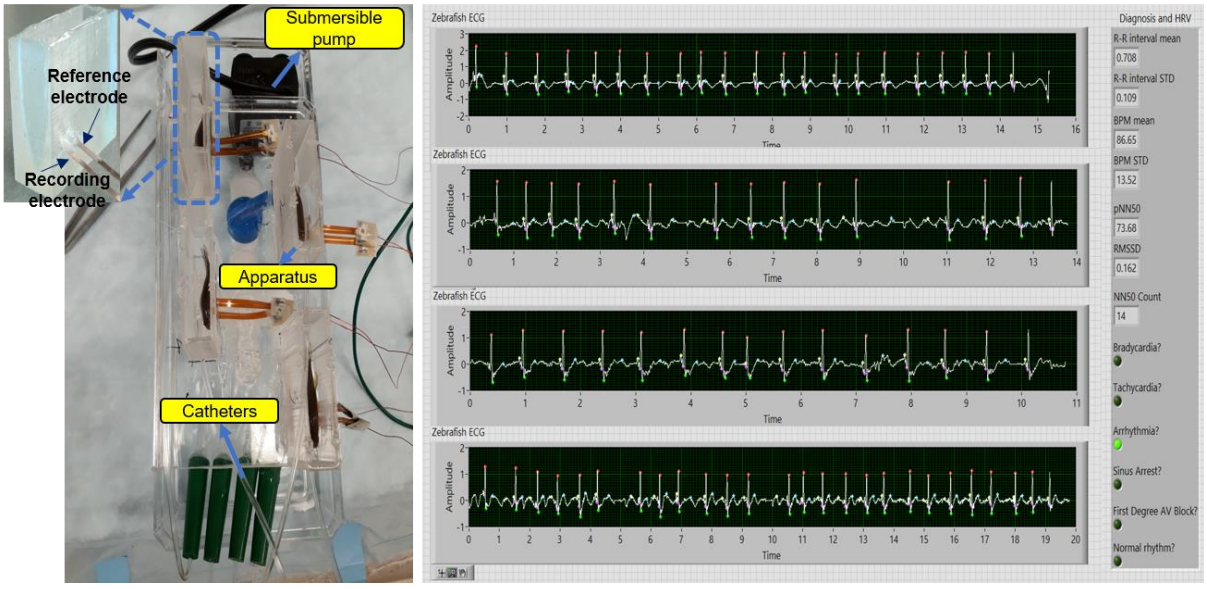


Fig. 2. 3. The prolonged anesthesia system for continuous zebrafish ECG monitoring (left) and the LabView GUI (right) for 4 fish simultaneously.

To integrate ECG electrodes, two strips of 125-μm thick polyimide (Kapton, DuPont, Wilmington, DE) with pre-sputtered Cu electrodes were inserted from the side of the apparatus through two thin-cut slits. The slits were then sealed by applying additional PDMS followed by a post-curing process. The strips were placed so that when the fish is loaded into the housing, the two electrodes would securely position at the chest and abdominal areas, acting as recording (WE) and reference (RE) electrodes, respectively (**Figs. 2.2.c and 2.d**).

Mechanical design: The entire system consists of three unique subsystems: the housing system, recording system, and anesthetic drug delivery system [19]. The housing system includes in-house apparatuses made of polydimethylsiloxane (PDMS) with 3D-printed molds. The apparatus was designed to fit a range of fish comfortably and secure them for recording (**Fig. 2.3 –** left panel). The detail of design was described in the previous section and could be found in our previous work [17]. The drug delivery system is illustrated in **Fig. 2.3**, including a submersible pump, a 4-way splitter, four catheters, a reservoir and tubing. The peristaltic pump placed inside the reservoir was to create the proper flow rate of system, which can be done by adjusting a valve connected to the pump. Here, a flow rate of 5.5-6 ml/min was chosen since it would provide zebrafish with adequate oxygen during the recording. The flow design of the system starts with the submerged pump inside the reservoir. After traveling through the pump, the solution is split 4 ways at the splitter. It goes through the catheters and into the fish directly though their mouths. Subsequently, the water exits through the gills and traveled back to the reservoir via the hole in the bottom of the apparatus. The ECG signals were collected with a high-gain differential amplifier (*A-M Systems, Sequim, WA*) and an in-house LabVIEW (*National Instruments, Austin, TX*) program (**Fig. 2.3** – right panel) was used to display, analyze and record the data which would be

discussed further in Section 2.2 . The gain was set at 10,000 with filters including a bandpass of 0.1-500 Hz and a notch at 60 Hz to eliminate the unwanted interferences. All signals were then fed into a data acquisition device (DAQ 6000, *National Instruments, Austin, TX*), being digitized at a sampling rate of 1,000 Hz.

2.1.2. Prolonged Zebrafish system – prototype 2

Needle-based electrodes: Although the advantage of MEA is inevitable, some drawbacks during the experiments should be taken into account. For instance, the contact between fish and the MEA is highly dependent on the fish's weight as the fish have to lay on the electrode to make contact, which leads to instability on the collected ECG signal.

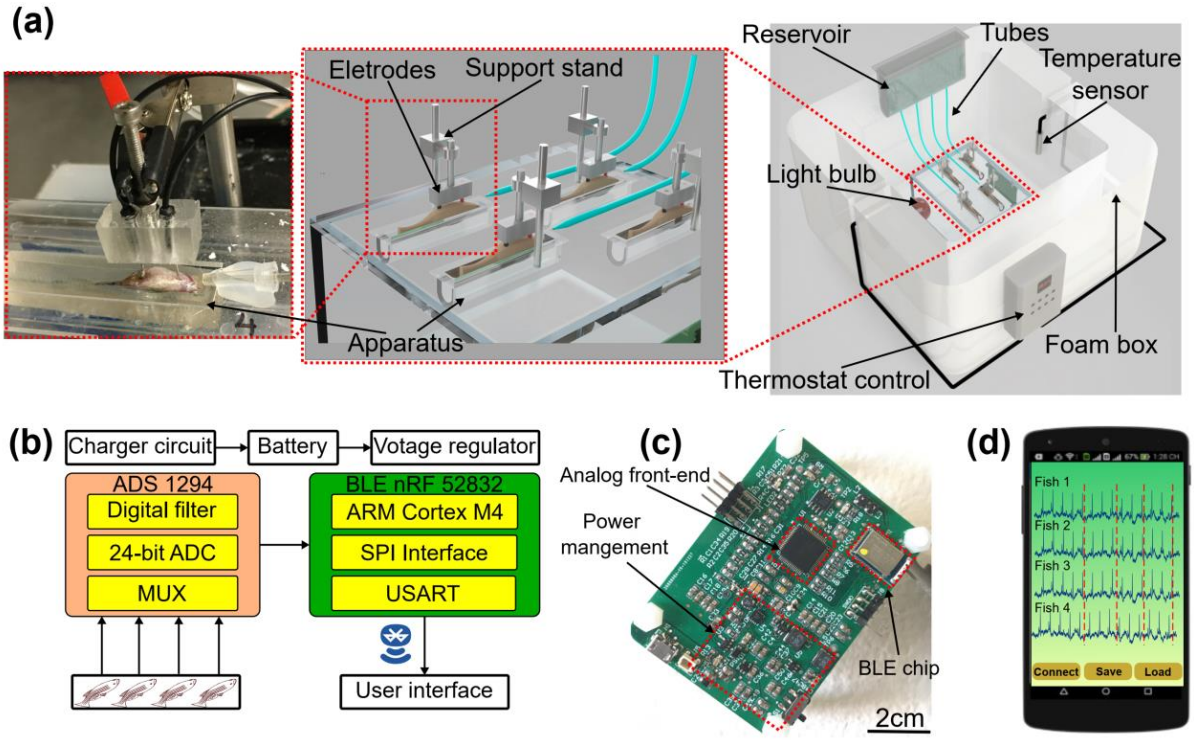


Fig. 2. 4. The prolonged ECG system for multiple adult zebrafish recording. (a) the prolonged ECG mechanical design: the reservoir containing solution, the tube system dropping the solution on the fish, the electrode and support stand recording the ECG signal. (b) System-level block diagram showing analog front-end chip, signal transduction, wireless transmission for the ECG signal to user interface. (c) In-house electronic board having system-on-chip for wireless transmission, power management connecting to the electrode for ECG acquisition. (d) User interface of mobile application receiving ECG signal from multiple fish.

Moreover, given the low Tricaine concentration continuously provided to the fish, unexpected movement was observed during the measurement with the MEA. This also reduced the SNR of the collected signal. Therefore, needle-based electrodes were used as the alternative. Specifically, 2 copper electrodes with the diameter of 0.8 mm were used for working electrode and reference electrode. To prevent these electrodes from causing injury to the fish, electrode' s tip was designed as ball shape. The general characteristics of the copper electrodes are relatively high conductivity and high tensile strength, enabling the needle to penetrate the fish dermis easily and provide strong electrical signals.

Mechanical design: The previous system used pump to create circular solution system; however, the pump's vibration significantly caused noise in the ECG signal. Therefore, in this version the perfusion system used four syringes instead along with four valves and tubing and the system took gravity effect to provide solution to the fish (**Fig. 2.4a**). Four syringes contained low dose Tricaine solution, continuously providing the solution to the fish through the tubing system to reduce their aggressiveness and activity while maintaining their consciousness. Four valves adjusted the solution's flow rate within a range of 5.5 - 6 ml/min [19]. Housing apparatuses and sensors were improved from the previous work [19]. Specifically, multiple side-fitted housings were made of polydimethylsiloxane

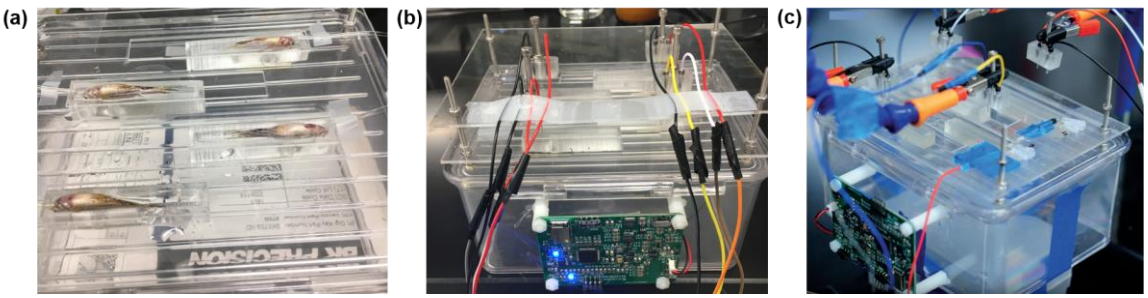


Fig. 2. 5. The prolonged system configuration: (a) fish placement in chambers; (b) electrodes placement with thumb screw; (c) electrodes placement with stand supports.

(PDMS), providing comfort to the fish and thus minimizing unwanted movements. Moreover, the top and bottom of the apparatus were designed in such a way that the fish can lay comfortably on electrodes within the curved bottom. The top had a redundant part sticking on the wall to keep the fish from escaping the apparatus. With the thermo box, a specific temperature was set by the thermostat control, and the light bulb was turned on so that the box's temperature can be maintained at the setup temperature and vice versa. **Fig. 2.5** illustrates the fish and electrode setup for the system. Specifically, 4 fish were put upside down with the tube projected toward the fish's head (**Fig. 2.5a**). For electrode placement setup, the first way was shown in **Fig. 2.5b** using a thumb screw. The electrodes were easily maneuvered up and down along the Z axes and the electrodes could be lowered until they touch the fish. However, this configuration was not effective for the measurement with multiple fish as it took 5-10 minutes to adjust all electrodes. As an attempt to cope with this issue, stand supports were utilized instead such that they will hold electrodes and we can adjust only swings connected to stand support (**Fig. 2.5c**).

Electronic design: The in-house electronic and a mobile application were developed as shown in **Figs. 2.4b-d**. Specifically, the system includes a system on chip (SoC) supporting Bluetooth Low Energy (BLE), an analog front-end for zebrafish ECG signal acquisition, and a power-supply module with charge management. The SoC adopted nRF52832 from Nordic (*Trondheim, Norway*), which was a 64-MHz Arm Cortex-M4 CPU with a built-in BLE module. The analog front-end of ADS1299 from TI was well-known for biopotential measurements with eight low-noise, programmable gain amplifiers (PGAs) and eight high 24-bit resolution Delta-Sigma ADCs. Its high bit resolution provided both precision and dynamic range, allowing it to capture signals as high as 4.5 V and as low as 0.5 μV. The data rate was

configurable from 250 samples per second (SPS) to 16 kSPS for all eight channels. Digital signals were sent from the ADS modules to the microcontroller for preprocessing via SPI interface. The data ready pin of the ADS1299 module was triggered to the signal microcontroller once a new data package is ready to transfer. Since multiple zebrafish were recorded simultaneously, differential mode configuration was utilized in the ADS1299. This would preclude signals in each fish from affecting each other. Moreover, a bias electrode with a signal generated by an ADS1299 chip was used, enabling a feedback loop built in the chip to get better common mode rejection. The overall fish ECG system electronic specifications are shown in **Table 2.1**.

Table 2. 1. Prolonged ECG monitoring system

Parameters	Numbers
Number of channels	4 (can be up to 8)
Bandwidth	1 Hz – 150 Hz
Input Range	±2.5V
Resolution	24 Bit
Sampling rate	250 Hz
Common mode rejection ratio (CMRR © 60 Hz)	100 dB
Input-Referred Noise	1 μV
Enclosure	Plastic
Dimensions	6.3’’W x 4.8’’L x 3.9’’H

Mobile application development: An Android application developed in Java that connects to the ECG system via BLE communication for data collection, displaying, and logging. We developed an Android smartphone application in Java that connects to the Bluetooth Low

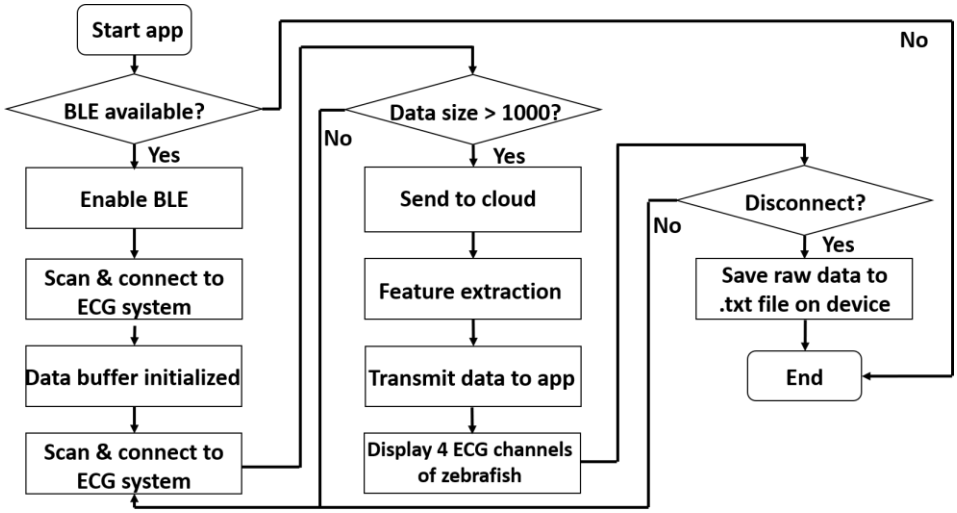


Fig. 2. 6. Diagram for wireless data transmission to mobile application.

Energy (BLE) device for data collection, displaying, and saving. Through BLE protocols, the application connects to the Zebra II and reads in multiple channels of data at a rate of 250 Hz. After accumulating 1000 data points, the input data is sent to the connecting cloud server for feature extraction of fish ECG. The results will start appearing on the application interface after at least 10 seconds of initially starting in the form of dynamic graphs – 4 ECG channels of zebrafish (**Fig. 2.6**). The user can disconnect the Zebra II at any time and save the raw data plus the time they are received in the application are recorded to a .txt file with a customizable name in the phone's external storage as well as on the cloud server as a backup. **Fig. 2.7** illustrates the user interface of mobile application, including main screen with 4 graphs showing ECG data collected from 4 zebrafish simultaneously. Once the device is turned on, the app will search device name in the device list and connect it automatically.

Fig. 2. 7. Diagram for wireless data transmission to mobile application.

2.2. ECG Processing and Analysis for awake zebrafish

2.2.1. LabView program

A LabVIEW GUI was developed to process and analyze zebrafish ECG data in real time [15]. The program was design to simultaneously collect and analyze ECG from zebrafish. As shown in **Fig. 2.3** – right panel, the graphic user interface describes 4 zebrafish ECG data displayed and labeled ECG wave components such as P, QRS, T waves and different types of heart disease classified for particular ECG data. An overview flowchart of the program is shown in **Fig. 2.8.**

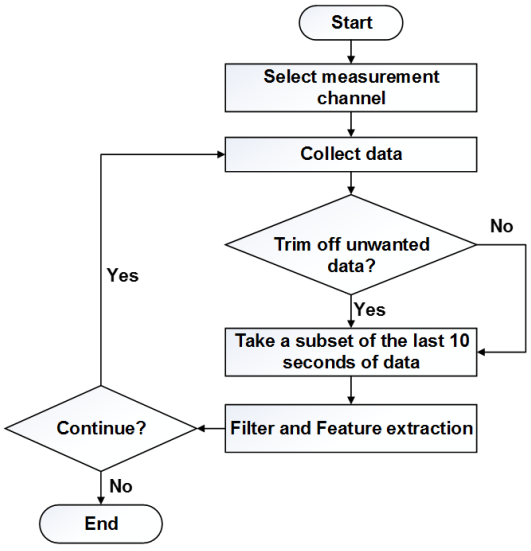


Fig. 2. 8. LabView program flowchart.

Pre-processing: For real-time recording and processing, first, ECG data are continuously collected at 1000 samples/second. They are checked to look for non-ECG periods, which

were due to uncertainties in measurement, such as unexpected strong movement of fish or electrode dislocation, then partitioned into 10-second long segments. Finally, each segment would be de-noised and analyzed along with abnormal detection. The process is iterated until the user stops recording the data. The method to remove non-ECG sections is illustrated in **Fig. 2.9**. In every five seconds, the coming data will be checked whether there are the glitches. Specifically, the maximum and minimum values of each interval will be detected (red and dashed orange lines in **Fig. 2.9**), then two thresholds (black lines) will be set up by taking the average of all maximum points and minimum points in each interval. Therefore, those meaningless points which are higher than the threshold could be found. Then, the distance between two such successive points would be found. If it is greater than 1000 samples, the data will be partitioned into different sub-segments (1st, 2nd and 3rd segments in **Fig. 2.9**).

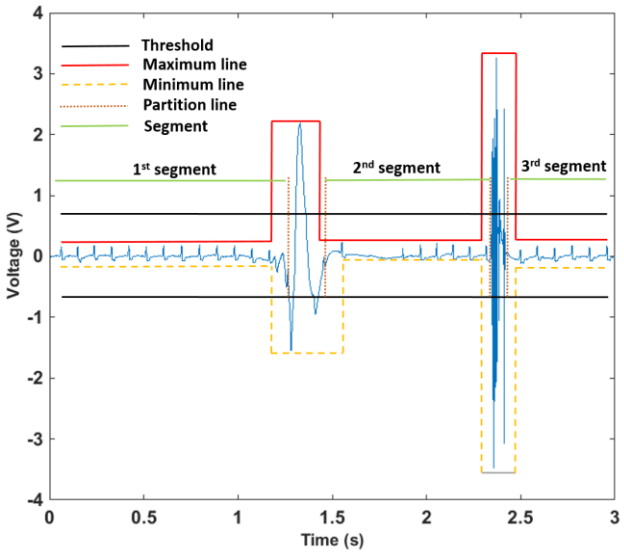


Fig. 2. 9. The illustration of our method to trim off unwanted data.

Once the non-ECG sections are removed, the data stream would be further processed to remove noise components such as baseline wander (breathing/gill motion interferences) caused by low-frequency and high-frequency noise. First, to remove high-frequency noise from the data, an initial Dolph-Chebyshev low-pass filter is applied using the Bio-signal Filtering VI as shown in Equation (1), which is part of the Biomedical Toolkit of LabView. The Dolph-Chebyshev window filter's Fourier transform is defined by equations reported

in [20, 21].

$$W(\omega_k)=\frac{\cos\left\{M\cos^{-1}\left[\beta\cos\left(\frac{\pi k}{M}\right)\right]\right\}}{\cosh\left[M\cosh^{-1}(\beta)\right]} \qquad (1)$$

where

M = Length of the window

α = - Sidelobe attenuation / 20

$$\beta=\cosh\left[\frac{1}{M}\cosh^{-1}(10^{\alpha})\right],\ k=0,1,2,\ldots,M-1$$

Lynch *et al.* [20] comprehensively compared this type of filter with others and the outperforming results were obtained as 1) the computational time was reduced from 6 hours to 3 hours with comparable performance; and 2) the Dolph window gave better attenuation in high-frequency ranges. Here, the Dolph-Chebyshev filter was used with a cutoff frequency of 40 Hz. As a result, the signals greater than 40 Hz were suppressed by a minimum of 40 dB. The baseline wander was eliminated using the Wavelet Denoise VI, which is a part of Digital Filter Design Toolkit of LabView, with the db06 wavelet. The selection of Daubechies mother wavelet showed the most effective result for de-noising ECG signals [22]. We examined the effectiveness of various mother wavelets by comparing different factors such as root mean square error (RMSE), root mean square bias (RMSB), and L1 norm, and revealed that the Deaubechies mother wavelet outperformed not only in terms of RMSE, but also the preservation of characteristic waves of ECG signals. Once the data were split into several segments with specific interval frequency, a soft threshold would be applied to suppress

coefficients which are smaller than the threshold. The new data were then reconstructed based on the new approximation and detailed coefficients.

Feature extraction: First, R peaks of the zebrafish ECG signal are detected using the Peak Detector VI within LabVIEW. Specifically, a threshold that equals 50% of the maximum amplitude in the signal is set up and points which are above the threshold value are the R peaks. Second, each R-R interval is used to find the P, Q, S, T wave. The Q-wave and S-wave peaks are searched for as the minimum points 50 ms before and after the R peak, respectively. The T-wave is defined as the maximum value between 15% and 55% of the R-R interval from the first R-wave in the interval. Finally, the P wave is the highest peak 65% to 95% of the R-R interval from the first R-wave in the interval. The R-R interval also is saved for the anomaly detection which is described in next section. Once all the ECG peaks are detected, typical features are calculated and reported in a document as shown in **Table I**.

To assess the efficacy of the program in de-noising, signal to noise ratio (SNR) calculation was used as reported in [23].

$$\mathrm{SNR}=\frac{\sqrt{\frac{1}{b_{s2}-b_{s2}+1}\sum_{k=b_{s1}}^{b_{s2}}\frac{1}{t_{s2}-t_{s1}+1}\sum_{i=t_{s2}}^{t_{s2}}X_k^{\,2}(i)}}{\sqrt{\frac{1}{b_{n2}-b_{n2}+1}\sum_{k=b_{n1}}^{b_{n2}}\frac{1}{t_{n2}-t_{n1}+1}\sum_{i=t_{n1}}^{t_{n2}}X_k^{\,2}(i)}} \quad (2)$$

where b_{s2}, b_{s1}, t_{s2}, t_{s1} are end beat number, start beat number, end time of the segment and start time for the segment used for calculating the signal, respectively. $X_k(i)$ are the amplitudes. b_{n2}, b_{n1} and t_{n2} are end beat number, start beat number and end time of the segment used for calculating the noise, respectively.

Table 2. 2. Summary of ECG Data Features

Feature	Result	Standard deviation	Unit
Number of R peaks	41	N/A	N/A
Average heart rate	73.12	15.09	BPM
Average R amplitude	0.49	0.04	mV
Average normalized P amplitude (in % compared to R amplitude)	0.18	0.11	%
Average RT interval	0.272	0.08	s
RMSSD	0.36	N/A	ms
NN50	29	N/A	N/A
pNN50	74.36	N/A	%

RMSSD: root mean square successive difference of intervals, NN50: The number of pairs of successive R-R interval that differ by more than 50 ms; pNN50: the proportion of NN50 divided by the total number of R-R intervals.

The signal value in Equation (2) was calculated using data from the R-wave of each heartbeat while the noise value was calculated using points from the space between the T-wave and P-wave of each cycle. The time intervals were the same for both the raw and filtered data.

Anomaly detection: In addition to finding the ECG features, the program was designed to diagnose various ECG anomalies. Standards for fish cardiac electrocardiography were established, which was based on our experience and observations after years working with zebrafish. Sinus bradycardia which is induced if the heartrate (HR) is less than 90 beats per minute (BPM). The HR value above 150 BPM can be diagnosed as sinus tachycardia. Sinus arrhythmia is a symptom if a difference in R-R interval length between the shortest and longest in the last 10 heart beats is greater than 0.16 seconds. If P-P interval is greater than twice of R-R interval, the signal may be sinus arrest. Finally, the symptom of first-degree AV block is presented when P-R interval is

greater than 0.20 seconds long. All anomalies can be detected by calculating the time differences between various ECG features such as R-R, P-R and P-P interval. **Fig. 2.10** depicts the detection of sinus arrest in a zebrafish ECG signal when comparing the difference between P-P and R-R interval (*see the red arrow*).

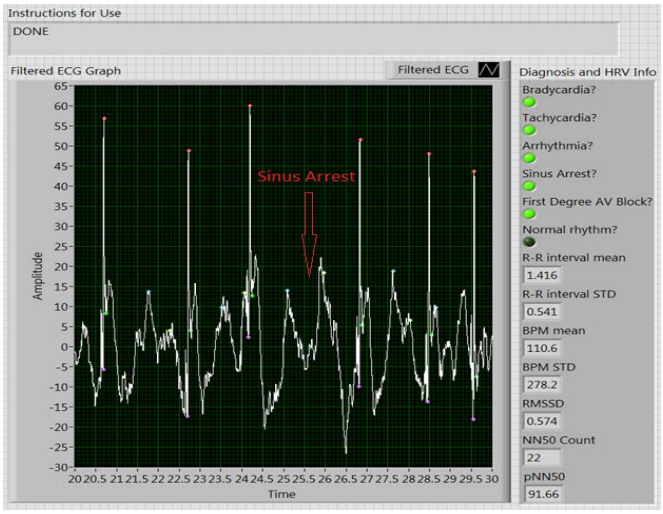


Fig. 2. 10. A zebrafish ECG signal with sinus arrest detected.

2.2.2. Machine learning (ML)-based for zebrafish analysis

Following classic approach based on zebrafish ECG morphology, machine learning-based approach was used to classify different types of heart disease with zebrafish ECG data. In this work, two ML algorithms (Convolution Neural Network and K-means) were investigated. We attempted to demonstrate the use of ML for training and recognition of three specific patterns: Atrioventricular Block (AVB), Sinus Arrest (SA) and ST Elevation (STE). Each signal in the dataset consists of approximately 4 to 5 beats which is roughly equivalent to 3000 samples at 1000-Hz sampling frequency. The data was pre-processed and filtered prior to training, the dataset size and training parameters were adjusted as necessary and the results were compared.

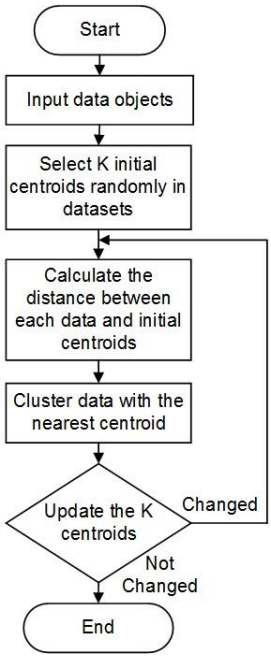


Fig. 2. 11. K-means clustering algorithm.

K-means Clustering: Clustering algorithms are often utilized for unlabeled feature detection and classification. K-means clustering

is a simple and straight-forward iterative approach and is one of the most well-known clustering algorithms in use today [24]. K-means has been successfully applied in a variety of fields, including human cardiology, obtaining high accuracy [25]. The high-level noise in the zebrafish ECG poses unique challenges. Moreover, each type of disease patterns has its own unique morphology, making clustering analysis is a good method for classification.

K-means works by partitioning data into k-numbers of clusters with one centroid defined for each cluster. The initial values of the centroids may be randomly selected. For each data point, the distance to each cluster is calculated. If the data point is closest to its initial cluster, the data point is left in that cluster, otherwise it is moved to the closest cluster [26]. This is repeated until a complete pass is made. The details of this approach are illustrated in **Fig. 2.11.**

Artificial Neural Networks: Researchers in the field of human cardiology trained a Deep Neural Network (DNN) in order to detect premature ventricular contractions, yielding 99.41% accuracy [27]. Convolutional Neural Networks (CNNs), a type of DNN, consistently deliver great performance in image classification tasks. For instance, Google's Inception v3 is capable of classifying images from 1000 classes with a Top-5 error rate of of 3.58% [28]. While neural networks can boast some impressive accuracy, one of their main trade-offs is that training one from scratch generally requires an elephantine amount of labeled data. Google's Inception v3 model, for example, was trained on 1.2 Million labeled images. Similarly, the neural network used in [27] was trained on over 80,000 samples. Fortunately, it is still possible to get favorable results from an ANN with a smaller dataset by utilizing a technique called "**transfer learning**" [29]. Transfer learning is the generalization of

information learned for one task, to another task. In the context of DNNs, a previously-trained neural network can be repurposed to work on a different problem. In order to accomplish this, the classification layer is replaced, and the final few layers are trained on a new dataset. During training, the weights of these final layers are adjusted to reflect the new data. This approach is possible because the output of this penultimate layer is essentially a feature vector which is then used by the classification layer. If the features are similar enough to the new task, the layer weights only need to be slightly adjusted to fit the new data. This approach is applied here to investigate whether generic visual features from a pre-trained CNN may be applied to successfully bootstrap recognition of cardiac anomalies in small-dataset environments, the last layer of a pre-trained image recognition CNN is retrained utilizing ECG data.

Here, we used Matlab to transform the 1-dimensional ECG signals into 2-dimensional images, resulting in square greyscale plots. The amplitude values of each segment were mapped to greyscale colors. Each segment was then mapped to a square greyscale matrix using a "z-layout", with the first sample occurring in the upper left hand corner of the matrix and the last sample occurring in the lower right hand corner. The signals were truncated to 2916 sample to accommodate the square layout. The CNN was constructed using Keras (an open source neural network library written in Python) with a Tensorflow backend [30, 31]. Google's Inception v3 model was imported without its final layer and the model was configured with inception's ImageNet pre-trained layer weights. The following layers were then added to the top of the inception architecture: a 2D global average pooling layer, two densely-connected layers with 1024 and 256 parameters both utilizing Rectifier Linear Unit (ReLU) as the activation function, then a final densly-connected layer using Softmax

activation. The layers on the original Inception model were frozen and the new layer was trained for 50 epochs utilizing only the output features from the original model, with rmsprop utilized as the activation function and Sparse Categorical Cross-Entropy used for the loss function. The layers after the final convolution layer (the final 12 layers) were then unfrozen and the model was retrained to fine-tune the weights on the 12 unfrozen layers. The model was trained for 400 epochs, the loss function employed was sparse categorical cross-entropy, the metric used was accuracy and the optimizer used was stochastic gradient decent with a learning rate of 0.0001 and a Nesterov momentum of 0.9, sparse categorical cross-entropy was used as the loss function.

Algorithm Evaluation: To ensure the validity of these approaches, 10-fold cross validation was employed, where the models were trained from 9 non-overlapping data subsets and tested from the remaining subset, and this process was repeated until all sets have been used for both testing and training and the results were then averaged. Precision, recall and F1 score were used as criteria for classifier evaluation [32]. These parameters are related to True Positives and False Positives (TP/FP) which refer to the number of predicted positives that were correct/incorrect, and similarly for True and False Negative (TN/FN) [32]. Here, TP_{AVB}, TP_{SA}, TP_{STE} are true positives; FP_{AVB}, FP_{SA}, FP_{STE} are false positives; FN_{AVB}, FN_{SA}, FN_{STE} are false negatives; and TN_{AVB}, TN_{SA}, TN_{STE} are true negatives for 3 classes AVB, SA and STE, respectively. The average accuracy, precision, recall and F1 score are calculated respectively by equations (1), (2), (3) and (4).

$$\text{Average Accuracy} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i+TN_i}{TP_i+TN_i+FP_i+FN_i} \quad (1)$$

$$\text{Precision} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i+FP_i} \quad (2)$$

$$\text{Recall} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FN_i} \quad (3)$$

$$\text{F1 score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

where N is number of classes and i (1, 2 and 3) represents AVB, SA and STE, respectively. Results were plotted and confusion matrices were used to indicate the efficacy of each training model. The number of correct and incorrect predictions of testing data will be illustrated in the confusion matrix along with the actual values. The diagonal of the matrix represents the outcomes where the predicted label was equal to the actual label.

2.3. Experimental design and preparation

All experiments were compliant with the Institutional Animal Care and Use Committee (IACUC) protocol (#AUP-18-115 at University of California, Irvine). Adult zebrafish targeted to the experiments were opened chest by doing surgery to form an incision and after two days they would be ready for the experiments.

2.3.1. Zebrafish husbandry

Mutant *Tg(SCN5A-D1275N)*, a transgenic zebrafish arrhythmia model bearing the pathogenic human cardiac sodium channel mutation *SCN5A-D1275N*, was used to characterize and validate device performance, study sinus node dysfunction, and perform drug high throughput screening assays. Correlation between clinical phenotype and the mutant line has been reported for bradycardia, conduction-system abnormalities, episodes of SA [33].

Adult wild/mutant-type zebrafish with the age of 13 to 20 months (body lengths approximately 3-3.5 cm) were used in this study. Zebrafish were kept in a circulating system

that is continuously filtered and aerated to maintain the water quality required for a healthy aquatic environment. The fish room was generally maintained between 26-28.5ºC, and the lighting conditions are controlled with 14:10 hours (*i.e.,* light: dark).

All animal protocols in this study were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) protocol (#AUP-18-115 at University of California, Irvine). All drug administration experiments (section 1-2, supplementary document), chemicals, reagents, and materials were performed in accordance with relevant guidelines and regulations.

2.3.2. Drug administration

To anesthetize fish, we used a buffered solution of 200 parts-per-million (ppm) Tricaine (Sigma, USA) [17]. Tricaine was dissolved in distilled water to a final concentration of 7,000 parts-per-million (ppm) as a stock, and the pH value was adjusted to 7.2 with sodium hydroxide (Sigma).

Amiodarone (Sigma) was dissolved in water at 65ºC for 2 hours and stocked as 900 μM at 4ºC. Before use, the solution was re-dissolved at 65ºC for 1 hour [34]. The fish were immersed in a tank with 100 μM amiodarone for 1 hour and then returned to fish water for 15 min before conducting experiment.

Sodium chloride (NaCl) was dissolved in water and stocked as 1.8 ‰ and the solution was re-dissolved into smaller doses (e.g., 0.1, 0.3, 0.6, 0.9 ‰) and keep the fish stay in that solution up to 30 mins before returning to fish water for the experiment of evaluation of Na+ sensitivity in the development of sinus arrest (SA) in *Tg(SCN5A-D1275N)*.

CHAPTER 3: RESULTS AND DISCUSSION

3.1. Demonstration of prolonged system

Fig. 3.1 presents the zebrafish ECG signal measured by the prolonged anesthesia system, and the recovery transition of the zebrafish from anesthesia to normal stages. Being treated with 50% of the original Tricaine concentration, fish were highly active, thereby interfering the signal (**Fig. 3.1a**) with signal-to-noise ratio (SNR) of 34.88 dB. Similarly, fish under the Tricaine concentration (150 mg/L) induced noise due to gill's motion, making the SNR of 50.4 dB (**Fig. 3.1b**); however, noise level was not as strong as that in fish with the

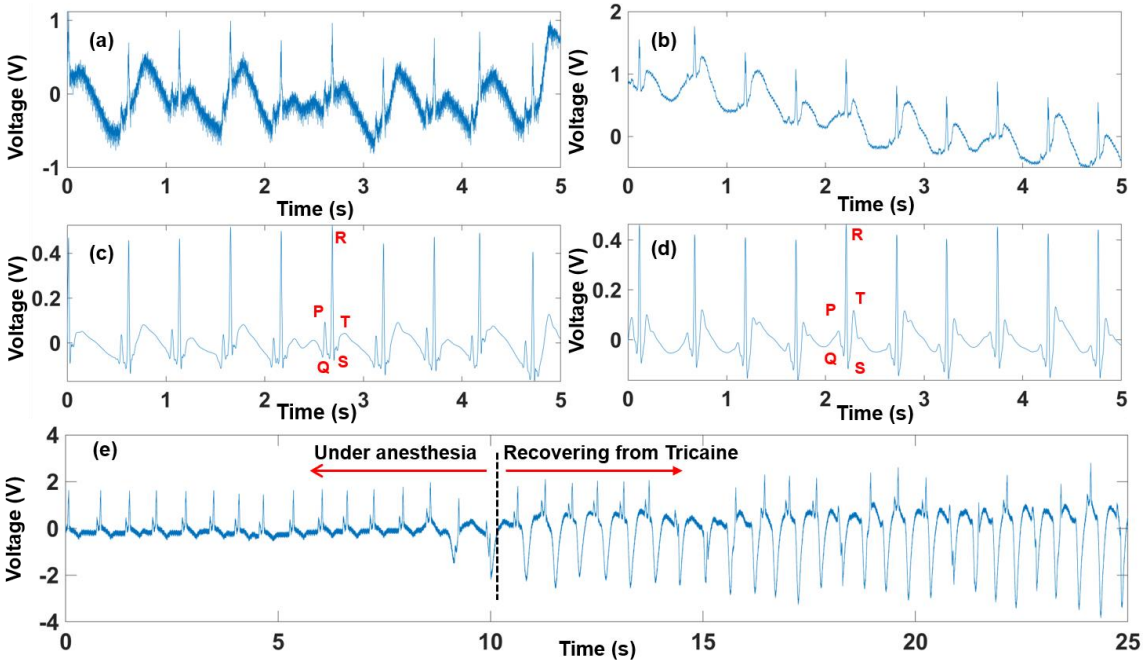


Fig. 3. 1. The zebrafish ECG collected with 100 mg/mL and 150 mg/mL Tricaine both raw data (a) & (b) and processed data (c) & (d). The transition stage when fish recover after being under anesthesia (e).

Tricaine concentration (100mg/L). **Fig. 3.1c** and **Fig. 3.1d** illustrate the processed ECG signals from the data collected by fish under 50% and 75% Tricaine, respectively. **Fig. 3.1e** depicts the moment when the fish started waking up. We can clearly notice the recovery transition with significant changes in the ECG signal. Specifically, the gills' motion started

dominating the signal as shown in the right side of the black dotted line in **Fig. 3.1e** compared with that noise level once the fish was under anesthesia. However, there were no abrupt and significant changes in the heartrate.

3.2. Investigation of side effects of Tricaine and variable temperature cardiac rhythm

Tricaine (MS-222) and temperature has been shown to affect cardiac functions of adult zebrafish and the HR of the treated subjects [9, 35]. As a result, they may skew the measurement of zebrafish physiological parameters. At low temperatures (*e.g.*, 18 - 20ºC), myocyte activity is reduced as a natural adaptive mechanism to aid survival during colder climates or reasons [36], which leads to a reduction in HR. At higher temperatures, increased HR facilitates greater cardiac output to support a higher metabolic activity/demand for

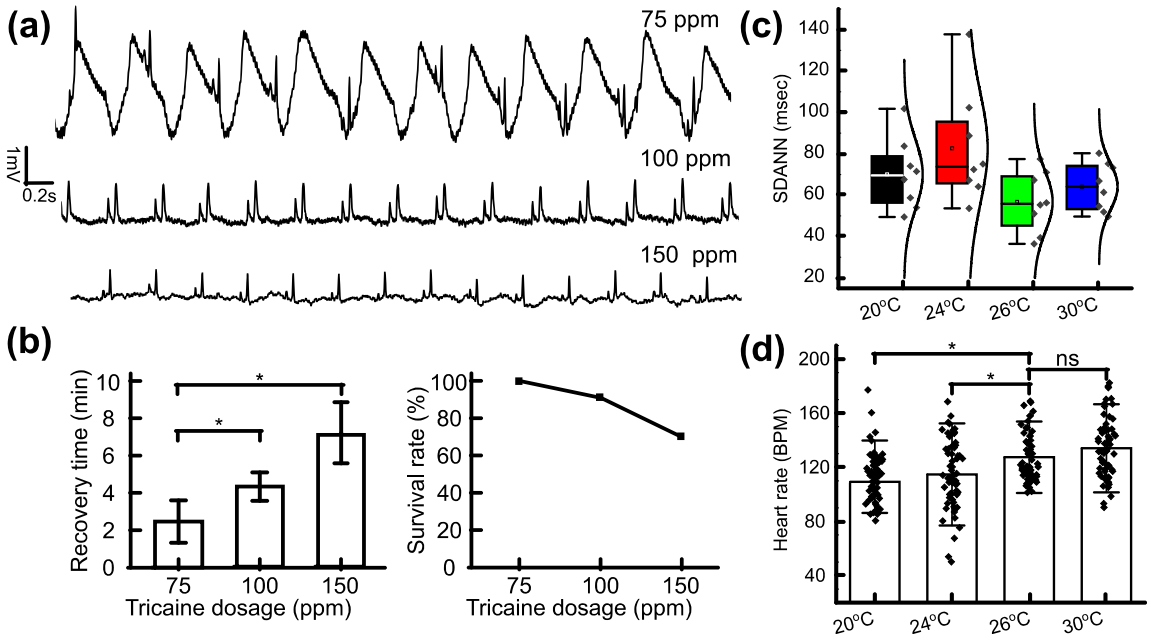


Fig. 3. 2. Investigation of tricaine and temperature to reduce cardiac rhythm side effect. (a) ECG morphology example recorded by different Tricaine concentrations. (b) Bar chart comparing recovery time needed after treatment for each Tricaine concentration. Line graph describing the survival rate of zebrafish treated by different Tricaine concentrations. (c) SDANN in WT fish with different temperatures. (d) HR in WT fish with different temperatures. *p <0.05; **p < 0.01 (one-way analysis of variance). *ns* indicates not significant.

oxygen consistent with normal biological rate function. Therefore, with an optimal environment temperature, our prolonged ECG system will help lower the tricaine concentration utilized, which can reduce the cardiac rhythm side effect as well as maintain the intrinsic ECG morphology.

As shown in **Fig. 3.2a**, the ECG morphology from eight fish per concentration group (75, 100, and 150 ppm) of Tricaine was observed. The ECG signal showed gill motion noise. The noise interfered with ECG waveforms such as P, T and QRS waves with 75 ppm Tricaine, while the signal appeared to be more stable under 100 ppm and 150 ppm treatment, providing clear ECG waves. After the 40 min measurement, the recovered time and survival rate of fish were collected (**Fig. 3.2b**). It was found that fish under higher Tricaine concentration needed longer recovery time. Specifically, it took an average of 7 min to recover the fish under 150 ppm compared to the 3 min and 4.2 min for fish treated under 75 and 100 ppm, respectively. Furthermore, with 150 ppm treatment, the survival rate was about 75%, while other concentrations yielded survival rates above 90%. It reflected the effect of a high dose used under a long period of time measurement, which was similar to the dose for fish euthanasia (*i.e.*, 168 ppm) [37]. Given the recovery time and survival rate along with ECG morphology, the dose of 100 ppm was the optimal one for the prolonged measurement. The heart rate variation (HRV) for zebrafish treated in 100 ppm showed no significant difference among the 30 min measurement, making the average standard deviation (STD) of 17 beats per minute (BPM). In contrast, the HRV for fish treated in 75 ppm was determined to have a 22 BPM difference based on STD. The recorded variable could be associated to the native response of individual fish to tricaine.

Given the optimal Tricaine concentration, we next investigated the effect of different temperatures on ECG acquisition. The zebrafish ECG system was placed in the middle of the home-made incubator. Temperature within the chamber ranged from 20ºC to 32ºC, as measured by a thermometer inside the chamber and controlled by a thermostat with accuracy of ±1˚C. Prior to recording the ECG signals, the impedance of the electrodes was measured on zebrafish skin, ensuring the signal stability during the long-term measurement. As shown in **Fig. 3.2c**, SDANN at 26ºC had the lowest value with the range of 36 millisecond (msec) to 75 msec, while the SDANN at 24ºC is highest with the range of 50 msec to 139 msec.

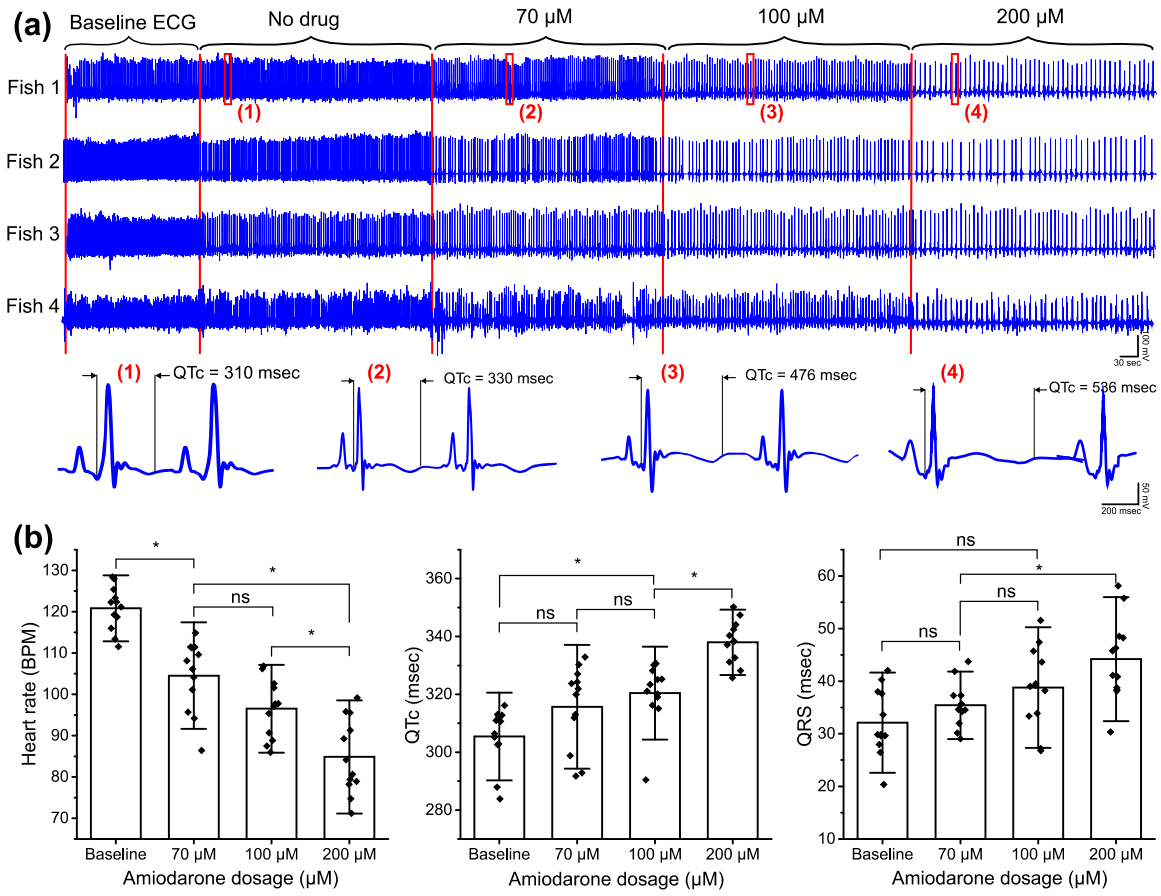


Fig. 3. 3. Demonstration of the prolonged ECG system showing the ECG morphology in response to different amiodarone concentrations. (a) The representative of ECG signal recorded by the proposed system and its change of signal morphology due to different amiodarone dosages in real time. (b) Bar chart describing the discrepancy of HR, QTc interval and QRS interval in ECG signal with different amiodarone dosages (n = 8 fish). *p <0.05 (one-way analysis of variance with Turkey test). *ns* indicates not significant.

Moreover, the data distribution from HR collected every 5 min at 26ºC was the most condensed (**Fig. 3.2d**). Thus, under 26ºC, the HR was most stable.

3.3. Response analysis to drug treatment in real time with the Zebra II system

One of the key novelties of the Zebra II system is the capacity to test drugs with different dosages on individual fish with a continuously prolonged assay. First, we explored the effect of Amiodarone with different dosages to zebrafish ECG morphology and ECG wave intervals. 3 doses of Amiodarone were consecutively filled in the reservoir to feed the fish while collecting the ECG data, and each dose lasted around 5 min. As shown in **Fig. 3.3a**, the changes in response to different dosages in all four fish were obvious. Zooming in on the data collected from fish 1 in different amiodarone dosages as denoted from (1) to (4), the QTc interval displayed considerable changes. For instance, the QTc interval was 310 msec without drug treatment, and it tended to increase after the fish was treated. Specifically, it was 330 msec with 70 μM of Amiodarone, while it was 476 msec and 536 msec with 100 μM and 200 μM of Amiodarone, respectively. **Fig. 3.3b** described the overall changes in terms of QT prolongation, QRS interval, and HR in response to different Amiodarone concentrations. With the increase of amiodarone dosage, QT prolongation and QRS interval showed an increase while the average HR was decreased.

In drug response studies, it is most ideal to conduct comparatively real time measurements for the same fish due to biological variability between individuals. Our perfusion system is equipped with multiple chambers for multiple drug doses in order to provide seamless transitions of multiple dosage treatments, allowing for a more native measurement. Our system was able to demonstrate that a longer acquisition time enables

the treatment of multiple drug doses, as indicated by our successful documentations of dose response Amiodarone-associated ECG changes as well as Na^+-associated arrhythmic symptoms described later. The utility of our perfusion system can also be expanded to include the testing of multiple drugs to assess the arrhythmic effects due to drug-drug interactions, a developing field of study [38]. The chambers in the perfusion system can also contain multiple drugs for the precise modulation of zebrafish drug intake to accurately determine the effects of each tested drug as well as the onset of potential drug-drug interactions. While our ECG analysis of *Tg(SCN5A-D1275N)* indicates that Methamphetamine (Meth) did not improve SA frequency and HR, it demonstrated the assessment of effects of multiple drugs, as seen from the prolongation of QT after treatment of 50 μM of Meth. In both groups, the QTc was longer (by 350 msec for wild type and 385 msec for *Tg(SCN5A-D1275N)*) after Meth treatment. Thus, the robust performance of the system allowed incorporation of multiple drugs with different effects (*e.g.,* antagonistic effects) in a single continuously prolonged assay to study drug-drug interactions on a specific arrhythmic phenotype, which was previously heavily performed on the short time course of current systems. Extending ECG acquisition in merely-sedated fish allowed the measurement of interactive effects of different drugs on a specific phenotype by a prolonged screening course. Different ECG phenotypes were recorded in our experiments using the prolonged real-time courses (over 40 min) to provide intuitive insights into drug interaction effects, demonstrating the potential of our system to evaluate drug efficacy. As shown in **Fig. 3.4c** for the Na+ sensitivity experiment, the average SDNN was 125 msec for wild type fish, but was 255 msec for *Tg*(*SCN5A-D1275N*), consistent with reduced conduction velocities due to Na^+ ion channel disfunction [39].

3.4. Evaluation of Na^+ sensitivity in the development of sinus arrest (SA) in *Tg(SCN5A-D1275N)*

To the best of our knowledge, there are no research studies established to investigate the Na^+ sensitivity in the development of SA. Therefore, an experiment was devised to determine the Na^+ sensitivity of the variant *SCN5A-D1275N* by observing changes in the frequency of SA episodes, the HR, and the QTc in the transgenic fish after treatment of 0.1, 0.3, 0.6, 0.9, and 1.8 ‰ of 5 M NaCl. WT fish (n = 12, aged 1.5 years) and *Tg(SCN5A-D1275N)* fish (n = 8, aged 10 months) were used in this experiment. Our developed system showed the susceptible effects of excess sodium ions to abnormal cardiac rhythm of the *Tg(SCN5A-D1275N)* mutant as a demonstration of characterizing different arrhythmic phenotypes. **Fig. 3.4a** illustrates the ECG morphology of *Tg(SCN5A-D1275N)* with different NaCl dosages. We noticed that with a small NaCl dosage (0.1 ‰), the zebrafish started showing a reduction in HR, followed by a more significant decrease in higher dosages. According to the SA criteria (*i.e.*, RR interval is greater than 1.5 sec) determined in our previous work [33], sinus arrest appears more frequently after treatment with 0.6 ‰ NaCl and above. The result confirms a strong association between high sodium intake and cardiovascular diseases, which has been reported in hypertensive populations [40]. A high sodium diet is associated with alterations in various proteins responsible for transmembrane ion homeostasis and myocardial contractility. Recent studies provided important evidence that excess sodium promotes structural and functional impairment of the heart, especially in populations bearing mutant phenotypes of the major cardiac sodium channels, including $Na_v1.5$ and its corresponding gene *SCN5A*. However, there is a current lack of a functional prolonged device to characterize $Na_v1.5$ sodium channel variants phenotypically, including the functional response of $Na_v1.5$

to initiate action potentials based on variation of sodium concentration. Sodium-overload sinus arrest observed in this study may be associated with a rise of the intracellular Na^+ in heart muscle due to the gain-of-function of *Tg(SCN5A-D1275N)* for sodium ions traveling into the cell. Detection of Na^+-induced SA by the developed system implied that the mutation *Tg(SCN5A-D1275N)* is susceptible to excess Na^+ ions due to hastening epicardial repolarization and causing idiopathic ventricular conduction, which induce ECG changes and ventricular arrhythmias. Moreover, the mutation *D1275N* evokes the long QT syndrome

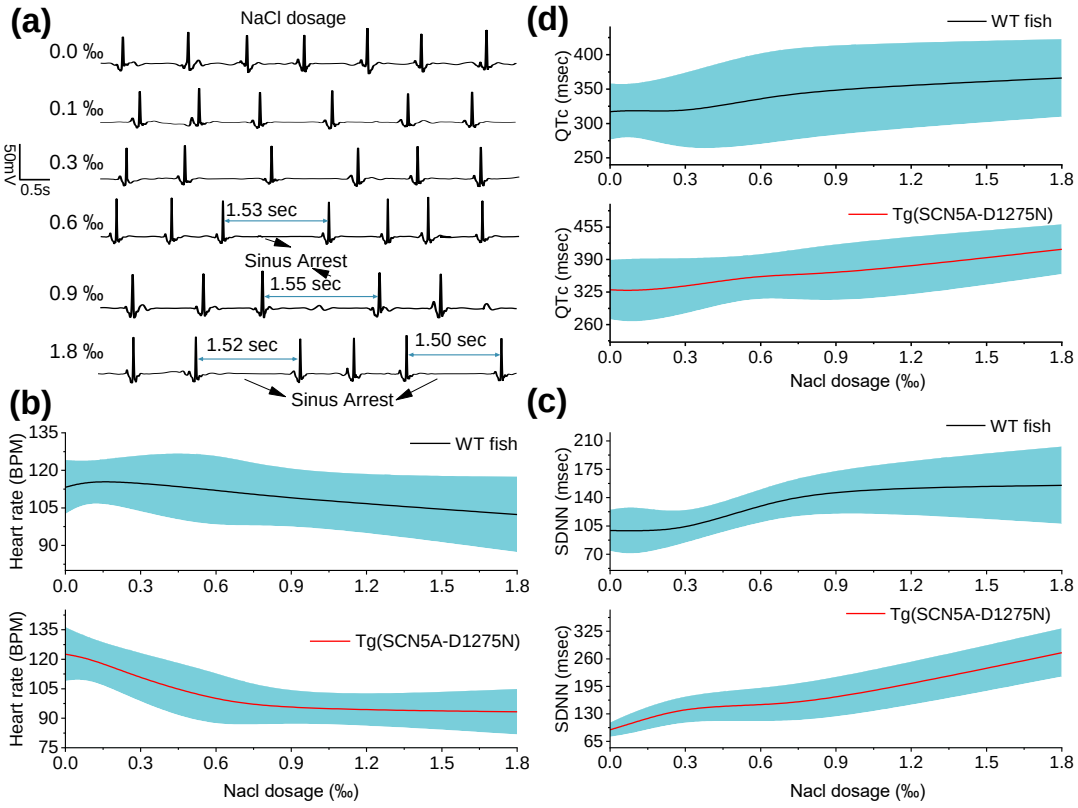


Fig. 3. 4. Evaluation of Na+ sensitivity in the development of sinus arrest in *Tg*(*SCN5A-D1275N*). (a) the representative ECG waveforms before and after NaCl treatment with different dosages. The SA appears more frequently in response to the increase of the NaCl dosage. (b) the average HR of wild-type fish (n = 12) and mutant fish (n = 8) with each dosage of NaCl. (c) SDNN of wild-type fish and mutant fish. (d) QTc values of two types of fish after treatment with different NaCl concentration.

(LQTS), which is caused by excessive I_{Na} detected by the developed system with continuously prolonged measurement. Recorded data by the developed system was consistent with clinical reports indicating that Brugada syndrome in human and animals reported Na^+-induced abnormalities in ventricular conduction [41, 42]. Thus, an overload of

Na^+ ions can cause destabilized closed-state inactivation gating of *D1275N* that may attenuate the ventricular conduction delay, shown in arrhythmic parameters.

As shown in **Fig. 3.4b**, the HR started significantly dropping in the 0.6 ‰ NaCl treatment in the mutant. In contrast, NaCl did not show a profound effect to the control group (WT fish) as evidenced by the slight decrease in the HR responding to different NaCl levels. It was worth noting that these WT fish were at 1.5 years old, which could attribute to an increase of SA [33], causing the slight reduction of HR in the experiment. In terms of HRV, *Tg(SCN5A-D1275N)* fish showed a remarkable increase with the high NaCl dosages (0.9 and 1.8 ‰) compared with other dosages. This provided evidence that the *Tg*(*SNC5A-D1275N*) triggered more SA under NaCl treatment. Moreover, the QTc interval in response to NaCl dosages were also measured, and it exhibited a similar pattern shown in SDNN measurements. (**Fig. 3.4d**).

Our notable finding is that the excessive sodium ions caused SA in *Tg(SCN5A-D1275N)* at 0.6 ‰, 0.9 ‰, and 1.8 ‰, corresponding to 1.53 sec, 1.55 sec and 1.52 sec, respectively. Additionally, slower HR and prolonged QTc were observed only in mutant fish. These results provide a significant association between the increased frequency of SA, slower HR, and prolonged QTc with increased sodium intake in SSS mutants. According to previous reports [43, 44], the $Na_v1.5$ sodium-channel protein can disrupt the heart's electrical activity and

lead to a dramatic decrease of HR. The slow-conducting *Tg(SCN5A-D1275N)* phenotype has been proved by voltage-clamp measurement [45, 46], in which data were consistent with our finding. In the eight mutant fish, the average QTc intervals were 385 msec at the upper physiological limit, indicating that the QTc intervals in *Tg(SCN5A-D1275N)* fish were generally more prolonged than wild type animals. Moreover, excess Na^+ ions caused not only slow HR and prolonged QTc but also increased SA frequency in *Tg(SCN5A-D1275N)* (**Fig. 3.4**).

3.5. Anomaly detection with machine learning

Fig. 3.5 illustrates the ECG waveforms of control, AVB, SA and STE fish as well as their conversion images prepared for the CNN approach. **Fig. 3.6** depicts the results from the K-means clustering algorithm which had an overall accuracy of 76% while the CNN approach used with 488 samples, obtaining an overall accuracy of 93.26 %, are showed in **Fig. 3.6b**. Both the K-means model and the ANN performed best when distinguishing STE with an accuracy of 80% and 96%, respectively; while it was more challenging to differentiate AVB and SA for both approaches. **Table 2** illustrates evaluation metrics for the two different

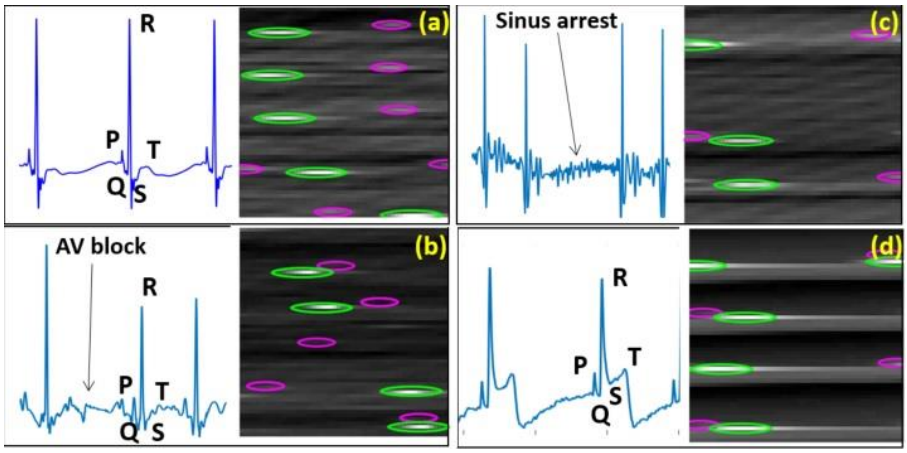


Fig. 3. 5. ECG patterns and their corresponding conversion images for CNN training. (a) Control fish; (b-d) Mutant lines with phenotypes of AVB, SA and STE, respectively. Green: R peaks; Pink: P peaks.

models, including precision, recall and F1 scores for each cluster. As shown in the table, F1 score obtained from CNN algorithm is higher than that from K-Means.

Table 3. 1. Classifier results

Abnormality	Precision		Recall		F1	
	K-Means	CNN	K-Means	CNN	K-Means	CNN
AV Block	0.78	0.95	0.70	0.89	0.74	0.92
ST Elevation	0.8	0.98	0.80	0.96	0.80	0.97
SA	0.73	0.86	0.80	0.95	0.76	0.90
Average	0.77	0.94	0.77	0.93	0.77	0.93

Unsupervised algorithms such as clustering compensate for a lack of labeled data by their ability to analyze unlabeled datasets and detect patterns that haven't been discovered before. In contrast, the CNN compensated by utilizing a supervised image-based approach which made labeling the training set much faster. These methods are thus a great fit to detect zebrafish ECG anomalies due to their capability to deal with less well-defined inputs and a lack of well-defined diagnostic criteria. CNN and K-means clustering approaches share the ability to extract latent features from complex data which makes both a good choice for this type of determination. Both of the methods employed allowed us to use quick visual evaluation of the signal by a human, allowing for more data to be labeled quickly for training. There are a few likely reasons for the lower accuracy of K-means relative to the CNN (Table 2). First, K-means clustering is known to suffer from problems due to high-dimensional data, especially in cases where data distribution causes the distance between the points to be uniform. This could potentially be addressed by reducing the dimensonality of datasets in the form of feature selection using methods such as principal component analysis (PCA)

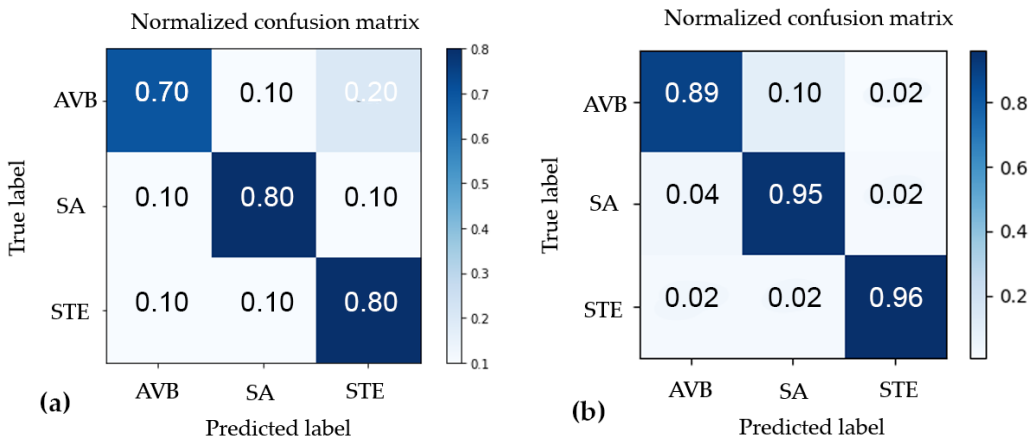


Fig. 3. 6. (a) The confusion matrix for k-means clustering based classification (b) The confusion matrix for CNN-based classification.

where irrelevant features or those that show high correlation with other features are removed. Second, it may be due to the randomized selection of initial centroids. The output of representative clusters from K-means is heavily dependent on the selection of the initial centroids. In the traditional K-means approach used here, these were randomly selected. To improve the accuracy, a different clustering algorithm could be used, features could be manually engineered to reduce their dimensionality, or K-means could be modified to allow for controlled selection of the centroids. In contrast to K-means clustering, the utilized CNN method performed beyond expectations given the small amount of labeled data. One aspect of the training which may be the key to the performance of this model, was curtailing the training to 400 epochs, as models trained beyond this point tended toward over-fitting. If more data were used to train the CNN, the training could likely be extended further than 400 epochs with a positive outcome. The utilization of an image-based method and a pre-trained CNN allowed us to circumvent many of the issues encountered with ECG data from zebrafish. While classification with K-means did not perform as well as the CNN did in this case, unsupervised clustering is an excellent tool for the discovery of common features within datasets. Further, the strengths of the two approaches create an opportunity to increase the

performance by combining both methods into a semi-supervised approach, possibly enabling more-efficient evaluation of zebrafish ECGs.

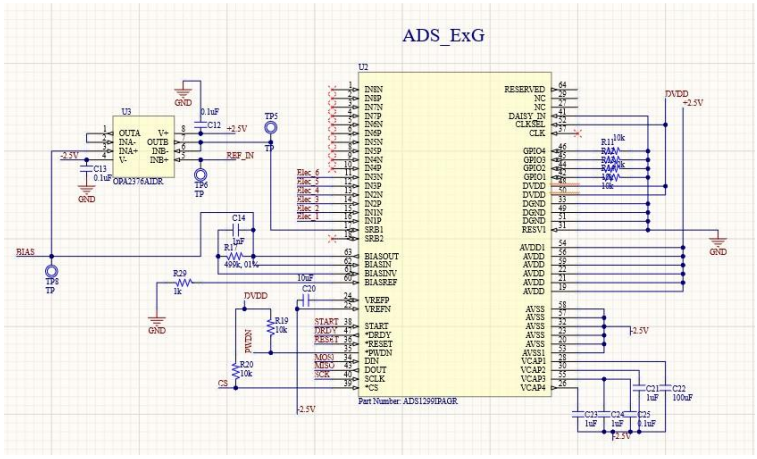
CHAPTER 4: CONCLUSION AND FUTURE WORK

The prolonged continuous ECG performance of our system increased diagnostic and monitoring yield in the detection of asymptomatic cardiac events and reduced ECG artifact to improve arrhythmia detection. The major novelties of the novel Zebra II ECG system lie in the prolonged measurement for multi-step experiments (up to 1 hour), high throughput screening with multiple fish, controlled environment with minimal side effects, and automated cloud-based analytics, among other controlled features. We have demonstrated and further deployed the system for phenotyping cardiac mutants in response to various drugs and environmental stimulus. Specifically, we have utilized our system to study Na^+ sensitivity of the variant *SCN5A-D1275N* by observing changes in the frequency of SA episodes, HR, and QTc. From our understanding, such an experiment has not been reported previously. Finally, the Zebra II can be used for a host of cardiac disease studies including phenotypic cardiac screening for genetic engineering studies and new drug screening applications.

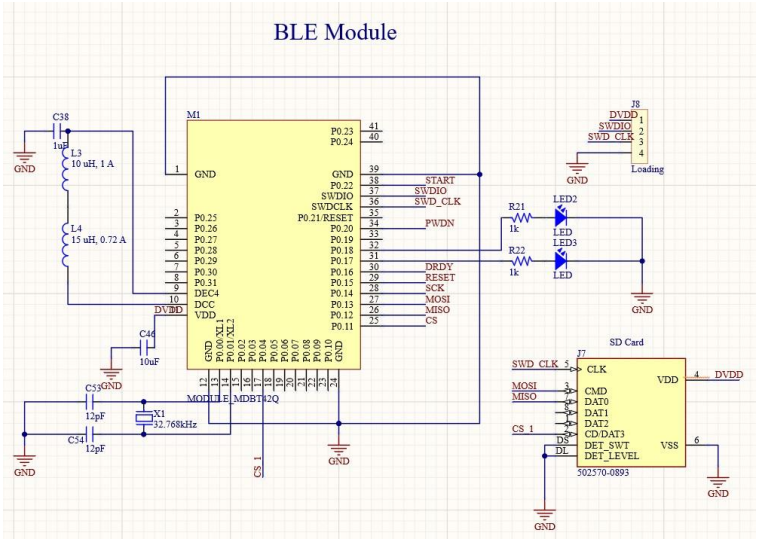
In the future, the prolonged ECG monitoring system can be improved so that more fish can be recorded simultaneously with favorable SNR. With the use of the low Tricaine concentration to enable longer ECG recording, the fish tended to exhibit unexpected strong movement, thus leading to the electrode dislocation. Thus, I may utilize a mechanism to automatically detect the impedance of electrode-tissue interface, which can be used to indicate the contact between electrode and fish's chest. More experiments for drug screening purpose with both wide type and mutant zebrafish will be conducted to validate the system.

APPENDIX

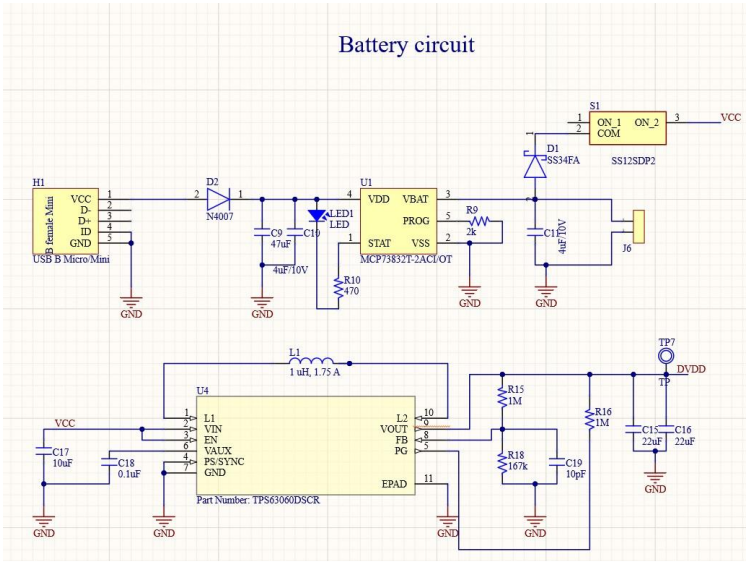
(a)



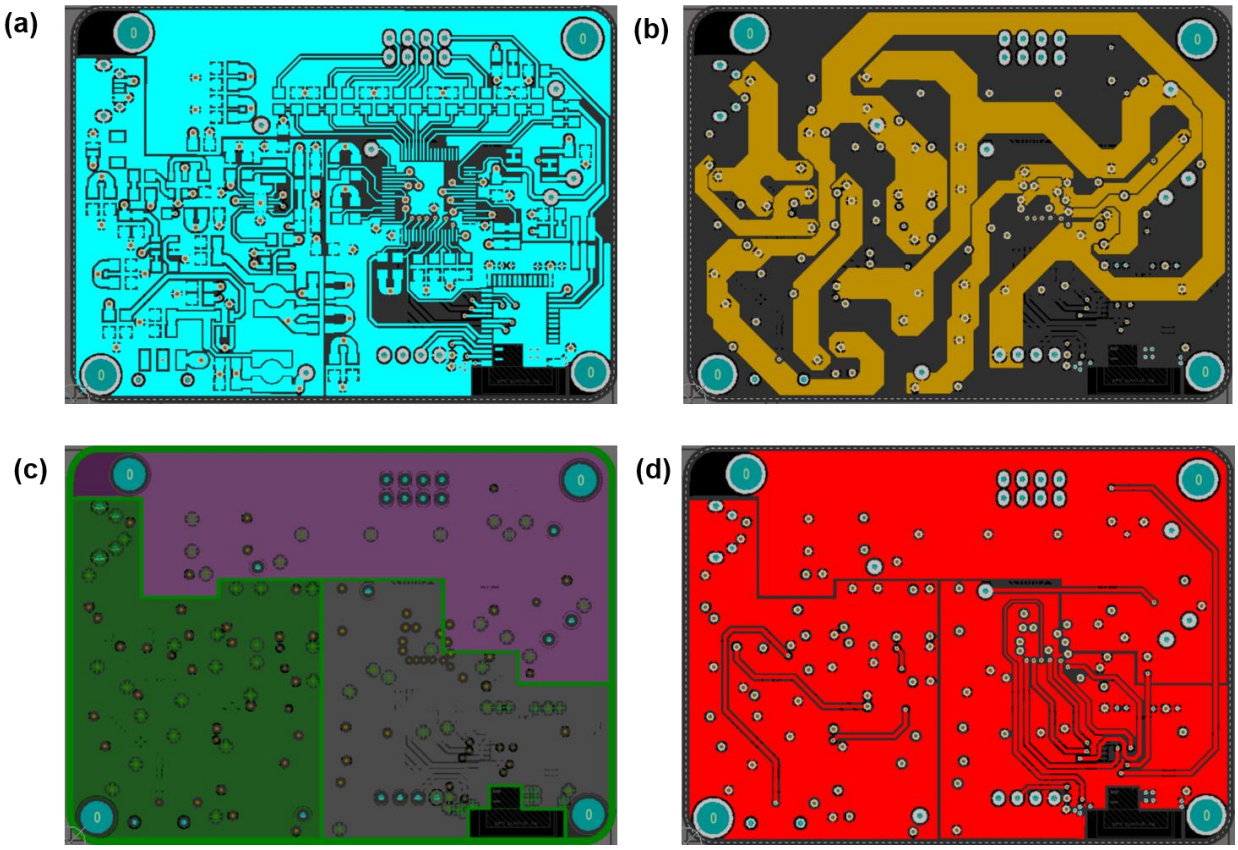
(b)



(c)



Schematic design



PCB design

REFERENCE

[1] (April 13). *Cardiovascular Diseases*. Available: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1

[2] G. A. Roth, M. H. Forouzanfar, A. E. Moran, R. Barber, G. Nguyen, V. L. Feigin*, et al.*, "Demographic and Epidemiologic Drivers of Global Cardiovascular Mortality," *New England Journal of Medicine,* vol. 372, pp. 1333-1341, 2015.

[3] (April 13). *Heart Disease in the United States*. Available: https://www.cdc.gov/heartdisease/facts.htm

[4] "Late-Breaking Science Abstracts and Featured Science Abstracts From the American Heart Association's Scientific Sessions 2020 and Late-Breaking Abstracts in Resuscitation Science From the Resuscitation Science Symposium 2020," *Circulation,* vol. 142, pp. e470-e500, 2020.

[5] (April 13). *Estimates of Funding for Various Research, Condition, and Disease Categories (RCDC)*. Available: https://report.nih.gov/funding/categorical-spending#/

[6] C. E. Genge, E. Lin, L. Lee, X. Sheng, K. Rayani, M. Gunawan*, et al.*, "The Zebrafish Heart as a Model of Mammalian Cardiac Function," in *Reviews of Physiology, Biochemistry and Pharmacology, Vol. 171*, B. Nilius, P. de Tombe, T. Gudermann, R. Jahn, R. Lill, and O. H. Petersen, Eds., ed Cham: Springer International Publishing, 2016, pp. 99-136.

[7] P. Nemtsas, E. Wettwer, T. Christ, G. Weidinger, and U. Ravens, "Adult zebrafish heart as a model for human heart? An electrophysiological study," *Journal of Molecular and Cellular Cardiology,* vol. 48, pp. 161-171, 2010/01/01/ 2010.

[8] K. Howe, M. D. Clark, C. F. Torroja, J. Torrance, C. Berthelot, M. Muffato*, et al.*, "The zebrafish reference genome sequence and its relationship to the human genome," *Nature,* vol. 496, pp. 498-503, 2013/04/01 2013.

[9] M.-H. Lin, H.-C. Chou, Y.-F. Chen, W. Liu, C.-C. Lee, L. Y.-M. Liu*, et al.*, "Development of a rapid and economic in vivo electrocardiogram platform for cardiovascular drug assay and electrophysiology research in adult zebrafish," *Scientific Reports,* vol. 8, p. 15986, 2018/10/30 2018.

[10] G. H. Chaudhari, K. S. Chennubhotla, K. Chatti, and P. Kulkarni, "Optimization of the adult zebrafish ECG method for assessment of drug-induced QTc prolongation," *Journal of Pharmacological and Toxicological Methods,* vol. 67, pp. 115-120, 2013/03/01/ 2013.

[11] C. C. Liu, L. Li, Y. W. Lam, C. W. Siu, and S. H. Cheng, "Improvement of surface ECG recording in adult zebrafish reveals that the value of this model exceeds our expectation," *Scientific Reports,* vol. 6, p. 25073, 04/29/online 2016.

[12] M. R. Stoyek, E. A. Rog-Zielinska, and T. A. Quinn, "Age-associated changes in electrical function of the zebrafish heart," *Progress in biophysics and molecular biology,* vol. 138, pp. 91-104, 2018/10// 2018.

[13] F. Yu, Y. Zhao, J. Gu, K. L. Quigley, N. C. Chi, Y.-C. Tai*, et al.*, "Flexible microelectrode arrays to interface epicardial electrical signals with intracardial calcium transients in zebrafish hearts," *Biomedical microdevices,* vol. 14, pp. 357-366, 2012.

[14] H. Cao, F. Yu, Y. Zhao, X. Zhang, J. Tai, J. Lee*, et al.*, "Wearable Multi-Channel Microelectrode Membranes for Elucidating Electrophysiological Phenotypes of

Injured Myocardium," *Integrative biology : quantitative biosciences from nano to macro,* vol. 6, pp. 789-795, 2014.
[15] M. Lenning, J. Fortunato, T. Le, I. Clark, A. Sherpa, S. Yi*, et al.*, "Real-Time Monitoring and Analysis of Zebrafish Electrocardiogram with Anomaly Detection," *Sensors,* vol. 18, p. 61, 2018.
[16] S.-J. Cho, D. Byun, T.-S. Nam, S.-Y. Choi, B.-G. Lee, M.-K. Kim*, et al.*, "Zebrafish as an animal model in epilepsy studies with multichannel EEG recordings," *Scientific Reports,* vol. 7, p. 3099, 2017/06/08 2017.
[17] T. Le, M. Lenning, I. Clark, I. Bhimani, J. Fortunato, P. Marsh*, et al.*, "Acquisition, Processing and Analysis of Electrocardiogram in Awake Zebrafish," *IEEE Sensors Journal,* vol. 19, pp. 4283-4289, 2019.
[18] C. Singleman and N. G. Holtzman, "Growth and Maturation in the Zebrafish, Danio Rerio: A Staging Tool for Teaching and Research," *Zebrafish,* vol. 11, pp. 396-406, 2014.
[19] T. Le, J. Zhang, X. Xia, X. Xu, I. Clark, L. Schmiess-Heine*, et al.*, "Continuous Electrocardiogram Monitoring in Zebrafish with Prolonged Mild Anesthesia," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020, pp. 2610-2613.
[20] P. Lynch, "The Dolph-Chebyshev Window: A Simple Optimal Filter," *Monthly Weather Review,* vol. 125, pp. 655 - 660, 1998.
[21] MathWorks. (2018). *Construct Dolph-Chebyshev window object - MATLAB*. Available: https://www.mathworks.com/help/signal/ref/sigwin.chebwin-class.html
[22] B. N. Singh and A. K. Tiwari, "Optimal selection of wavelet basis function applied to ECG signal denoising," *Digital Signal Processing,* vol. 16, pp. 275-287, 2006/05/01/ 2006.
[23] J. S. Lian, Sriram; Tsia, Ho-Chie; Wu, Dongsheng; Avitall, Boaz; He, Bin, "Estimation of Noise Level and Signal to Noise Ratio of Laplacian Electrocardiogram During Ventricular Depolarization and Repolarization," *Journal of Pacing and Clinical Electrophysiology,* vol. 25, 2002.
[24] S. S. Vijayarani S, " An Efficient Clustering Algorithm for Predicting Diseases from Hemogram Blood Test Samples," *Indian Journal of Science and Technology,* 2015 Aug; 8(17).
[25] P. P. Kumar Pandey, K.L. Jaiswal, Atul, "Classification Model for the Heart Disease Diagnosis," *Global Journal of Medical Research,* 2014-04-29 2014.
[26] M. P. S. K.A. Abdul Nazeer, "Improving the Accuracy and Efficiency of the k-means Clustering Algorithm," *Proceeding of the World Congress on Engineering,* vol. 1, 2009.
[27] V. V. C. Szegedy, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2015.
[28] J. C. J. Yosinski, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," *Advances in Neural Information Processing Systems, Montreal, Canada,* vol. 27, pp. 3320-3328, 2014.
[29] J. A. A. M. Badshah, N. Rahim, and S. W. Baik, "Speech Emotion Recognition from Spectrograms with Deep Convolutional Neural Network," *2017 International Conference on Platform Technology and Service (PlatCon),* pp. 1-5, 2017.
[30] F. Chollet, "Keras," 2.0 ed. https://github.com/fchollet/keras: GitHub, 2015.

[31] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro*, et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467,* 2016.
[32] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.,* vol. 45, pp. 427-437, 2009.
[33] J. Yan, H. Li, H. Bu, K. Jiao, A. X. Zhang, T. Le*, et al.*, "Aging-associated sinus arrest and sick sinus syndrome in adult zebrafish," *PLOS ONE,* vol. 15, p. e0232457, 2020.
[34] Y. H. Chen, H. C. Lee, R. J. Hsu, T. Y. Chen, Y. K. Huang, H. C. Lo*, et al.*, "The toxic effect of Amiodarone on valve formation in the developing heart of zebrafish embryos," *Reprod Toxicol,* vol. 33, pp. 233-44, Apr 2012.
[35] P. Sun, Y. Zhang, F. Yu, E. Parks, A. Lyman, Q. Wu*, et al.*, "Micro-Electrocardiograms to Study Post-Ventricular Amputation of Zebrafish Heart," *Annals of Biomedical Engineering,* vol. 37, pp. 890-901, 2009/05/01 2009.
[36] M. Maricondi-Massari, A. L. Kalinin, M. L. Glass, and F. T. Rantin, "The effects of temperature on oxygen uptake, gill ventilation and ecg waveforms in the nile tilapia, Oreochromis niloticus," *Journal of Thermal Biology,* vol. 23, pp. 283-290, 1998/10/01/ 1998.
[37] M. Matthews and Z. M. Varga, "Anesthesia and euthanasia in zebrafish," *Ilar j,* vol. 53, pp. 192-204, 2012.
[38] S. Zukkoor and V. Thohan, "Drug-drug interactions of common cardiac medications and chemotherapeutic agents," *American College of Cardiology,* pp. 653-655, 2018.
[39] A. L. George, Jr., "Inherited disorders of voltage-gated sodium channels," *The Journal of clinical investigation,* vol. 115, pp. 1990-1999, 2005.
[40] A. Mente, M. O'Donnell, S. Rangarajan, G. Dagenais, S. Lear, M. McQueen*, et al.*, "Associations of urinary sodium excretion with cardiovascular events in individuals with and without hypertension: a pooled analysis of data from four studies," *The Lancet,* vol. 388, pp. 465-475, 2016.
[41] C. Antzelevitch, "Brugada syndrome," *Pacing and clinical electrophysiology : PACE,* vol. 29, pp. 1130-1159, 2006.
[42] Y. Mizusawa and A. A. M. Wilde, "Brugada Syndrome," *Circulation: Arrhythmia and Electrophysiology,* vol. 5, pp. 606-616, 2012.
[43] M. Liu, K.-C. Yang, and S. C. Dudley, "Cardiac sodium channel mutations: why so many phenotypes?," *Nature Reviews Cardiology,* vol. 11, pp. 607-615, 2014/10/01 2014.
[44] D. Darbar, P. J. Kannankeril, B. S. Donahue, G. Kucera, T. Stubblefield, J. L. Haines*, et al.*, "Cardiac Sodium Channel (SCN5A) Variants Associated with Atrial Fibrillation," *Circulation,* vol. 117, pp. 1927-1935, 2008.
[45] W. P. McNair, G. Sinagra, M. R. G. Taylor, A. Di Lenarda, D. A. Ferguson, E. E. Salcedo*, et al.*, "SCN5A Mutations Associate With Arrhythmic Dilated Cardiomyopathy and Commonly Localize to the Voltage-Sensing Mechanism," *Journal of the American College of Cardiology,* vol. 57, pp. 2160-2168, 2011/05/24/ 2011.
[46] T. P. Nguyen, D. W. Wang, T. H. Rhodes, and A. L. George, "Divergent Biophysical Defects Caused by Mutant Sodium Channels in Dilated Cardiomyopathy With Arrhythmia," *Circulation Research,* vol. 102, pp. 364-371, 2008.