Developing an Ultra-Portable, Low-Cost Sanger Sequencing System

Mischa Johal (majohal67@gmail.com) - 2020

Abstract:

This work focused on the development of a low-cost, ultra-portable Sanger Sequencing system for DNA sequencing. Sample isolation by centrifugation, amplification by PCR, and sequencing by capillary electrophoresis utilizing orthogonal LIF detection would be miniaturized into a 130mm x 350mm x 10mm space, with the added goal of an integrated user control analysis and system through a Raspberry Pi. Separate centrifuge, PCR, and capillary electrophoresis unit design and partial prototyping were accomplished, with further work required on the Raspberry Pi integrated control system and spectrum analysis instrument.

Introduction:

Genetics has recently become increasingly popular in mainstream society, beginning with the publication of the human genome, expanding to life-altering processes such as CRISPR, and becoming universalized through companies such as Ancestry or 23AndMe. (Carver et al. 2017)

Behind all these innovations is “DNA sequencing”, the concept of determining the specific order of the A, G, T, and C bases that make up human genes. However, the vast majority of systems that sequence and analyze DNA are quite large, consist of multiple parts, and are extremely expensive for anyone outside a university or governmental lab. (Carver et al. 2017)

The goal of this study was to create a low-cost, ultra-portable (130mm x 350mm x 10mm), “laboratory” that would use Sanger Sequencing as it’s ultimate method of DNA analysis, with the intention of such as device to be deployed in resource-poor third-world countries where cheap, efficient disease genome sequencing is essential to vaccine creation. The components involved in this setup would be a centrifuge, a thermo-cycler suited for PCR applications, and an orthogonal LIF CE system with integrated DNA analysis, with the added goal of having each of these systems running on an integrated, user-controlled Raspberry Pi.

Sanger Sequencing Overview:

The process of DNA synthesis and replication in a cell involves DNA helicase, DNA polymerase, DNA template, and deoxynucleotides. DNA replication starts when DNA helicase unravels the double-helix structure to expose single-stranded DNA and form a replication fork. RNA primase introduces a primer that binds to the single-stranded DNA. DNA polymerase then binds to the replication fork and starts DNA synthesis by sequentially adding nucleotides to the 3′-hydroxyl end of the RNA primer bound to the DNA template. The result is the creation of an "extension product ." The extension product grows in the 5′ to 3′ direction by forming a phosphodiester bridge between the 3′-hydroxyl group at the growing end of the primer and the 5′-phosphate group of the incoming deoxynucleotide. The DNA sequence is copied with high fidelity because at each base on the DNA template, DNA polymerase incorporates the nucleotide that is complementary to that base . Thymine (T) is complementary to adenine (A) and guanine (G) is complementary to cytosine (C) because they can form hydrogen bonds with each other. (Thermofisher et al. 2016)

Sanger Sequencing (or dideoxy sequencing) takes advantage of the ability of DNA polymerase to incorporate 2′,3′-dideoxynucleotides - nucleotide base analogs that lack the 3′-hydroxyl group essential in phosphodiester bond formation (an oxygen has been removed from a normal dNTP). Sanger sequencing requires a DNA template, a sequencing primer, DNA polymerase, deoxynucleotides (dNTPs), dideoxynucleotides (ddNTPs), and reaction buffer. DNA polymerase adds a deoxynucleotide or the corresponding 2′,3′-dideoxynucleotide at each step of chain extension . Whether a deoxynucleotide or a dideoxynucleotide is added depends on the relative concentration of both molecules. When a deoxynucleotide (A, C, G, or T) is added to the 3′ end, chain extension can continue. However, when a dideoxynucleotide (ddA, ddC, ddG, or ddT) is added to the 3′ end, chain extension is terminated with dideoxynucleotides at the 3′ end. After this, the DNA Sequencing itself is completed by Capillary Electrophoresis (CE) with the addition of fluorescently labeled ddNTPs. These fluorescent dyes label the extension products and the components are combined in a reaction that is subjected to cycles of annealing, extension, and denaturation in a thermal cycler. Thermal cycling the sequencing reactions creates

and amplifies extension products that are terminated by one of the four dideoxynucleotides. The ratio of deoxynucleotides to dideoxynucleotides is optimized to produce a balanced population of long and short extension products. As the DNA mix is undergoing CE, a light source - whose wavelength is perfectly attuned to excite the fluorescent dye - is shined through the sample, which becomes excited and emits fluorescence. This light is captured by a spectrophotometer or similar instruments, and is fed into a computer which converts these intensity measures into an easily-readable graph. The peaks of fluorescence correspond to the individual base they were paired with, and it is in this way that shorter sequences of DNA can be sequenced. (Thermofisher et al. 2016).

Design Principals

Centrifuge:

Design of the centrifuge was based around the Dremelfuge design (Fig. 1), using a modified Dremel 3000 series connected to the central Raspberry Pi to modulate speed (desired at 14k rpm for each cycle). Dremelfuge project notes were referenced in fabricating the ABS plastic tube-holder body of the centrifuge using a Makerbot 3-D printer.

PCR Thermo-Cycler:

PCR design was based on two versions of an Open-PCR project (Perfetto et al. 2015), and modified to suit the constraints of this project. Eppendorf tubes were used as a guideline for measurements of sample-holding units (Fig. 2). An 85mm x 85mm x 35mm computer heat sync was used to hold samples and provided effective heat distribution, with sample vials being embedded into holes drilled into the top portion of the heat sync. Denaturing was accomplished using a 12W Peltier element attached to the bottom of the heat sync (Perfetto et al. 2015) The other side of the element was attached with thermal epoxy to a 90 degree CPU cooler, which was used for respective annealing cycles. Temperature was detect by a thermocouple drilled into the

heat sync (Fig. 3), which was connected to a MAX31855 breakout board (K, Stacey et al. 2012), which in turn, along with the CPU cooler and Peltier element, was connected to a Raspberry Pi used to control timing and amount of PCR cycles. Code for cycling and user interface was primarily based on OpenPyCR by Garvey et al. (2014).

Capillary Electrophoresis System:

Electrophoretic system was determined to be solution easiest to miniaturize for the purposes of this project, as opposed to a larger hydraulic injection system. (Voeten et al. 2018) A 0 to -6000V high-voltage module (E60, XP-Emco) was utilized in combination with two platinum electrodes for solution injection into a short-run fused-silica capillary system (10cm length, fully transparent, beg. ID 51.4µm, end ID 52.8µm, TSP050192, Polymicro Technologies). The initial end of the capillary was to be kept stationary, while an X-axis mini translation stage (Ebay) would be used in combination with a 12V DC motor to transition between the buffer-filled vial and the sample-filled vile. (Pan et al. 2018) A waste vile collected processed buffer and solution at the other end of the capillary, with a design based on Thermofisher's Sanger Sequencing unit (Fig. 4-6). A 300nm and 505nm laser diode (Lee's Electronics) were placed 5cm from the beginning of the capillary and directed vertically upward through the capillary. The LIF detection system composed of a collimating lens, a bandpass filter, diffraction grating, and a photodiode. The collimating lens was specified to be a moulded aspheric condenser lens with 0.56 numerical aperture (88286, Edmund Optics) placed at a 45 degree angle to the capillary. This angle was favoured based on the results of Fang et al. (2006), as opposed to the traditional setup 90 degrees to the capillary. After fluorescence collection, a 510nm bandpass filter (CB510, Tangsinuo) physically reduced noise of the incoming signal, a thin plastic film from the top of a compact disk was used as a diffraction grating, which was finally placed over a PIN photodiode (SG108S-18, SGLux) for signal collection. (flourescane intensity CE citation) The photodiode was connected to an ADC (AD623DNZ, Lee's Electronics), which was in turn connected to the central Raspberry Pi. (Gus et al. 2016) All of the components of the capillary injection and LIF

detection systems were installed within a fabricated ABS encasing. Finally, data visualization was accomplished in the Raspberry Pi using Matplotlib. (Johal et al. 2018)

Experimental Procedure:

Hypothesized experiment proceeds as follows: (Palumbi et al. 2002)
Extract 0.1 - 0.5 grams of the the desired tissue type and place in a 1.5 ml micro-centrifuge tube (Eppendorf tube) Add an approximately equal volume of Lysis Buffer (100mM EDTA 10mM Tris (pH 7.5), 1% SDS) to the tissue swatch. Macerate/grind the tissue, then microcentrifuge this solution for approx. 3 min at 14,000 rpm, decant supernatant and save. Add an approximately equal volume of buffer-equilibrated phenol to the decanted supernatant and invert gently, then microcentrifuge for 10 min at 14,000 rpm. After centifuging, the DNA should be suspended within the upper aqueous layer. Immediately under this layer and above the bottom phenol layer there may be a whitish layer that contains proteins and carbohydrates. Pipet off only the topmost layer and save. Add an approximately equal volume of 1:1 phenol/chloroform to the decanted liquid and invert gently. Microcentrifuge for 5 min at 14,000 rpm. Pipet off and save the clear/whitish DNA containing layer off the top of the phenol/chloroform. Add an approximately equal volume of chloroform to the decanted liquid and invert gently. Microcentrifuge for 3 min at 14,000 rpm. Pipet off the top layer into well labeled, pre-weighed microcentrifuge tubes. Weigh each sample and determine the volume of each by assuming 1 gram of the liquid = 1 ml. Add a volume of 7.5 M NH_4, and add to this solution either (1) enough 2-propanol to make the solution 50% 2-propanol, or (2) enough ethanol to make the solution 66% ethanol. Invert gently and let the solutions sit for at least 10 min at room temperature. Microcentrifuge for 10 minutes at 14,000 rpm and decant liquid, keeping pellet intact, then almost fill each sample tube with 70% ethanol and let the solution sit for 10 min. Microcentrifuge for 10 min at 14,000 rpm and decant remaining liquid. Vacuum dry the pellet and add 100µl of sterile distilled H_2O to each sample and freeze until needed. PCR was theorized to be preformed using Promega PCR Master Mix (Fig. 7), Promega GoTaq reaction buffer (Fig. 8), Promega dNTP mix (Fig. 9), Invitrogen primers (Fig. 10, 13), Roche ddNTPs

(Fig. 11), and with two separate concentrations of Biotium GelGreen and GelRed dyes respectively (Fig. 12), with a 60s 95C burn in and 32 cycles of 20s 95C denature, 15s 65C anneal, 30s 72C extend, and 20s 4C chill, estimated to run approximately 2 hours. Finally, CE would be accomplished by first flushing capillary with Biobasic TBE buffer (Fig.14), and then completing cycles of CE until total amount of sample had been analyzed, with expected dye emissions of approximately 530nm (GelGreen) and 600nm (GelRed) (Fig. 15).

Conclusion:

In the attempted creation of a low-cost, ultra-portable Sanger Sequencing system for DNA sequencing, separate, miniaturized versions of a centrifuge, PCR thermo-cycler, and LIF detection capillary electrophoresis module were designed prototyped with relative success. Opposing this, the additional goal of integrating the user control analysis and system through a Raspberry Pi was not achieved, nor was the real-time fluorescence analysis, partially due to a lack of monetary resources, and mainly the result of a void of electronics and circuitry knowledge and human resources. Thus, the goal and hypothesis of the study was not achieved, but several advancements in the miniaturization and cost-efficacy of the Sanger Sequencing process were accomplished.

Appendix:

Google drive link to all images and references included in the study.
https://drive.google.com/open?id=12xRiAn5VhclFrgf6bwjVL17CNzEvZIvS

Citations:

Voeten, Robert L C, et al. “Capillary Electrophoresis: Trends and Recent Advances.” Analytical Chemistry, American Chemical Society, 6 Feb. 2018, www.ncbi.nlm.nih.gov/pmc/articles/PMC5994730/.

Scheeline, Alexander. “How to Design a Spectrometer.” SagePub, 2017, journals.sagepub.com/doi/pdf/10.1177/0003702817720468.

Perfetto, Josh. “Wild OpenPCR.” Wild OpenPCR - Hackteria Wiki, 2012, hackteria.org/wiki/Wild_OpenPCR.

Pan, Jian-Zhang, et al. “A Low-Cost Palmtop High-Speed Capillary Electrophoresis Bioanalyzer with Laser Induced Fluorescence Detection.” Scientific Reports, Nature Publishing Group UK, 29 Jan. 2018, www.ncbi.nlm.nih.gov/pmc/articles/PMC5789010/#CR7.

Palumbi, Steve. “THE SIMPLE FOOL’S GUIDE TO PCR.” University of Hawaii, 2002, palumbilab.stanford.edu/SimpleFoolsMaster.pdf.

Johal, Mischa. “How to Convert Analog Output of Photodiode into Digital Intensity Chart?” All About Circuits, 31 Aug. 2018, forum.allaboutcircuits.com/threads/how-to-convert-analog-output-of-photodiode-into-digital-intensity-chart.151719/.

Hickstein, Dan. “Matplotlib - Color under Curve Based on Spectral Color.” Stack Overflow, 1 June 1967, stackoverflow.com/questions/44959955/matplotlib-color-under-curve-based-on-spectral-color.

Gus. “Raspberry Pi ADC (Analog to Digital Converter).” Pi My Life Up, 18 Dec. 2019, pimylifeup.com/raspberry-pi-adc/.

Garvey, Cathal. “OpenPyCR.” GitLab, 2014, gitlab.com/cathalgarvey/OpenPyCR.

“DNA Sequencing by Capillary Electrophoresis.” Thermofisher Applied Sciences, 2016, www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/sequencing_handbook_FLR.pdf.

Carver, Rebecca Bruu, et al. “Young Adults' Belief in Genetic Determinism, and Knowledge and Attitudes towards Modern Genetics and Genomics: The PUGGS Questionnaire.” PloS One, Public Library of Science, 23 Jan. 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5256916/.

Special thanks to Jeff Spence (John Oliver Secondary), Carl Sommerfeld (John Oliver Secondary), Robert Kennedy (University of Michigan), Alex Chattwood (Genome BC)