Expression of Green Fluorescent Protein in Competent [Escherichia Coli.](https://www.scielo.br/j/gmb/a/T7SGC8JyGC7xyvnB6wwt9NS/?lang=en)

¹Ahmad Haris Khan, ²Izaz Hayat, ³Bello Abdullahi, ⁴Sulaiman Shah, ⁵Allah Madad

1,2,3,5School of Life Sciences, Department of Biology, Shiihezi University, Xinjiang, China. 4 School of Agriculture, Department of Horticulture, Shihezi University, Xinjiang, China Corresponding Author's email address: hariskhan@stu.shzu.edu.cn Authors' Contribution: Ahmad Haris Khan, Bello Abdullahi and Izaz Hayat, Allah Madad, Sulaiman Shah conducted the experiments in Shihezi University, Xinjiang China.

Abstract:The main purpose of gene cloning for biotechnological tenacities is to express the cloned gene in the selected organism, in this case, competent E. coli cells. Green fluorescent protein (GFP) was isolated using PCR and used as the gene of interest to transform competent E. coli cells. The PCR-amplified gene of interest (GFP) was ligated into the pGM-T expression vector to generate a recombinant plasmid. Competent E. coli cells were obtained using the CaCl2 method, and competent cells were transformed with our recombinant plasmid. The blue-white screening method was used to screen for E. coli cells that were successfully transformed using our recombinant plasmid. Transformed bacterial colonies were white, whereas untransformed E. coli colonies remained blue. Abstraction of the recombinant plasmid was performed to detect the presence of a green fluorescent protein. The plasmid was confirmed to carry our foreign gene based on the increase in its length upon gel electrophoresis. To create a restriction map, double enzymatic digestion of the recombinant plasmid was performed; however, the digestion was unsuccessful. In this study, GFP was successfully isolated, replicated into an appearance vector, and transmuted into proficient Escherichia coli DH5α cells.

Key Words: GFP, PCR, Competent cells, Plasmid, E. coli

Date of Submission: 21-03-2024 Date of acceptance: 03-04-2024

I. INTRODUCTION:

The main purpose of gene cloning for biotechnological purposes is to express the cloned gene in a selected organism for research or commercial purposes [1, 2]. To achieve this goal, vectors with high protein expression rates are developed or selected [2]. Genetic elements with high protein expression are used, which control transcription, translation, stability and secretion of the protein [2]. Elements frequently manipulated include promoter and terminator sequences, ribosome binding site (RBS), copy number of the cloned gene, vector type (plasmid-borne or integration into the host genome), cellular location of the foreign protein synthesized, host translation efficiency, and internal stability of the cloned gene protein within the host cell [3]. The level of recombinant protein expression also depends on the host organism, and both prokaryotic and eukaryotic organisms can express foreign genes. The minimum requirement for an efficient gene expression system is the presence of a strong and regulated promoter sequence upstream of the cloned gene. A strong promoter is one that has high affinity for RNA polymerase. The ability to regulate the promoter allows the cell to precisely control the extent of transcription [4, 5]. Strongly regulated promoters were used for the construction of expression vectors. Protocols have been developed to isolate the promoter regions from several organisms [6]. In addition to the promoter, the second major factor required for the expression of cloned DNA is the effective ribosome-binding site. In E. coli, the ribosome binding site includes an initiation codon (ATG) and a 3–9 nucleotide sequence located 3–11 nucleotides upstream of the initiation codon sequence (Shine-Dalgarno-SD), which is complementary to the 3rd codon. 'end of E. coli 16S rRNA [7]. Ribosome binding to mRNA is believed to be facilitated by base pairing between the SD sequence in the mRNA and a sequence at the 3' end of 16S rRNA. When a prokaryotic gene is expressed, the ribosome-binding site of this gene may be sufficient. Eukaryotic and prokaryotic genes with weak ribosome binding sites require an efficient ribosome binding site [6, 7].This can be achieved by:

2. Primer repair and

^{1.} Synthesis of DNA encoding the amino terminus

3. Development of a restriction site.

Typically, recombinant DNA clones are constructed using restriction enzymes/ligases. DNA ligase, a DNA joining enzyme called DNA ligase, catalyzes the formation of a phosphodiester bond between the 52 phosphate nucleotide of one DNA fragment and the 32-hydroxyl group of another [8]. The joining of linear DNA fragments by covalent bonds is called ligation. An alternative to restriction enzyme/ligase cloning is PCR cloning [9, 10]. Recombinant DNA or DNA cloning is now more widely carried out using PCR. Traditional PCR cloning involves introducing restriction sites using primers into the PCR product or cloning the blunt ends of PCR products [9, 10]. The DNA fragment and vector of interest can be amplified by PCR and ligated without the use of restriction enzymes [10, 11]. PCR cloning is a rapid gene cloning method that is often used in projects that require higher throughput than traditional cloning methods [9]. This makes it possible to clone DNA fragments that are not available in large quantities. Typically, a PCR reaction is performed to amplify the sequence of interest, which is then attached to the vector via blunt or single base ligation prior to transformation [10]. Early PCR cloning often used Taq DNA polymerase to amplify the gene. The result was a PCR product with a single template-independent base, an adenine (A) residue added to the 3' end of the PCR product due to the normal action of the polymerase [10]. These A-tailed products were then ligated to a complementary Ttailed vector using T4 DNA ligase, followed by transformation [8]. High-fidelity polymerases are now widely used to amplify DNA sequences using a PCR product that does not contain 3' extensions [8]. The blunt-ended fragments are attached to the plasmid vector through a typical ligation reaction or by the action of an "activated vector that contains a covalently attached enzyme, usually topoisomerase I, which facilitates insertion of the vector [9, 10]. PCR cloning of blunt-ended fragments was unsuccessful. Some PCR cloning systems contain engineered "suicide" vectors that include a toxic gene to which the PCR product must be successfully ligated to allow propagation of the strain that takes up the recombinant molecule during transformation. 10]. There are several PCR cloning methods and protocols, including TA cloning, ligation-independent cloning (LIC), gateway recombinant cloning, sequence- and ligation-independent cloning, one-step cloning, overlapping PCR cloning, and rapid cloning.

1.2 OBJECTIVES: The goal of this study was to isolate, clone, and express green fluorescent protein in competent E. coli DH5α.

II. MATERIALS AND METHODS:

2.2 MATERIALS:

2.1 Flow Chart:

PCR Material Mix, Tiangen MIDI Purification Kit, 70% ethanol (room temperature), 1X TAE EDTA, gold view, agarose powder, restriction enzymes, LB medium, ampicillin, IPTG, X-gal, primers, gel loading dye, distilled water, RNase free water, Pasteur pipette, absorbent paper, weighing boat, cotton wool, pipette tips, Para film, sterile 1.5 ml micro centrifuge tubes, PCR tubes, hand gloves, scalpel blades, electrophoresis kit,

primer, micropipettes, stand, centrifuge, PCR machine. /thermal cycler, UV trans illuminator, spatula, analytical balance, conical flask, graduated cylinder, microwave oven, hotplate and magnetic stirrer, beaker, vortex machine, water bath, freezer and orbital mixer.

2.3 METHODOLOGY:

2.3.1 Acquisition of target gene

2.3.1.1 PCR amplification of the target gene

The PCR method was used to obtain the green fluorescent protein (GFP) gene using the primers provided for the experiment. PCR is a method developed by Kari Mullis to amplify a specific section of DNA. Many different PCR assays have been developed over the years, and the method has a wide range of applications. PCR amplification was performed in a reaction volume of 25μl using the provided master mix. The reaction mixture (25μl) consisted of 2x PCR master mix (12.5μl), 10μM forward and reverse primers (1μl each), 9.5μl RNase-free water, and 1μl DNA. Amplification conditions were set at 94°C for 4 minutes followed by 30 cycles of 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 45 seconds. The final extension was set at 72°C for 8 minutes. It was then kept at 4°C until further analysis.

2.3.1.2 Agarose gel electrophoresis: Agarose gel electrophoresis was performed to detect the presence of the target gene (GFP) after PCR. PCR products were analyzed on a 1% agarose gel, and a 2-kb marker was used to determine the size of the PCR product. 5μl of marker and PCR products were dispensed into separate wells, and the electrophoresis reservoir was filled with 1x TAE buffer. The electrophoresis chamber was then connected to an electrical current for 30 min at a constant voltage of 120V. The gel was then removed from the chamber and casting tray and the fragments were viewed using a UV trans illuminator.

2.3.1.3 PCR purification: The PCR purification concept was developed to recover the PCR product from the agarose gel. For experiment; The tiangen MIDI cleaning kit was used for cleaning. Everything was done according to the protocol specified in the cleaning kit.

Protocol:Before use, 100% ethanol was added to the PW buffer. Column equilibration: 500 µl of buffer BL was added to the CA2 spin column (CA2 spin column was placed in a collection tube). The mixture was then centrifuged for 1 min at 12,000 rpm $(-13,400 \times g)$ in a benchtop microcentrifuge. The flow was discarded and the CA2 spin column was placed back into the collection tube. The DNA fragment was excised from the agarose gel using a clean sharp scalpel. A piece of the gel was weighed in a clean microcentrifuge tube. 150μL of PN buffer was added to the gel (if the gel weighs 0.1 g, the default volume was 100μL, then add 100μL of PN buffer was added. The tubes were then incubated at 50°C and turned up and down until the agarose gel was completely dissolved. The solution was transferred to a CA2 spin column. The column was maintained at room temperature for 3 min and then centrifuged for 60 s at 12,000 rpm (approximately 13,400 \times g) in a benchtop microcentrifuge. The flow was discarded and the CA2 spin column was placed back into the collection tube. The CA2 spin column was washed with 600_{μL} PW buffer and centrifuged for 60 s at 12,000 rpm (~13,400 × g). The flow column was discarded and the CA2 spin column was placed back into the collection tube (this step was performed twice). The CA2 spin column was placed back into the collection tube and centrifuged at 12,000 rpm $(\sim 13,400 \times g)$ for 2 min to remove residual PW buffer. The flow-through liquid was discarded, and the column was placed with the lid open for a few minutes to air-dry the membrane. The CA2 spin column was transferred to a clean 1.5 mL micro centrifuge tube. 30μL of EB buffer was added to the center of the membrane, incubated at room temperature for 2 min, and centrifuged at 12,000 rpm $(\sim 13,400 \times g)$ for 2 min. After purification, 30μl of PCR product was transferred to a microcentrifuge tube. 5μl The PCR product was mixed with 1μl of DNA loading dye. A final volume of 6μL was run on agarose gel to ensure successful reconstitution.

2.3.2 Cloning and gene expression:

2.3.2.1 Construction of the recombinant plasmid:The gene fragment product of interest for PCR amplification was ligated into the pGM-T expression vector. Typically, the idea behind ligation is to insert the foreign gene of interest into the multiple cloning site of our pGM-T expression vector. The ligation reaction was carried out in a reaction volume of 10μl. The reaction mixture contains 1μl of DNA ligase buffer, 1μl of T4 DNA ligase, 1μl of pGM-T expression vector, 5μl of the gene of interest (GFP, obtained from PCR purification), and 2μl of nuclease-free water. The solution was stirred and stored at 4℃ overnight.

2.3.2.2 Obtaining competent E. coli cells: A 10mL aliquot of E. coli broth was placed on ice for 15 minutes. It was then centrifuged at 6000 rpm for 10 minutes at 4℃. The supernatant was discarded and both pellets were carefully resuspended in 2 ml of ice-cold 0.1mol/L CaCl₂ solution. The mixture was left to cool on ice for 20 minutes. It was then centrifuged at 6000 rpm for 10 minutes at 4℃. After removing the supernatant, 0.4 ml of a cold CaCl₂ solution with a concentration of 0.1 mol/L was added to the sediment. The cells were carefully suspended in CaCl₂ solution after cooling on ice for 2 min. The solution was then stored at 4℃ for 24 hours to increase the transformation efficiency by 400-600%.

2.3.2.3 Transformation of competent cells

1.Preparation of LB solid middle plates for conversion.: 50 mg/ml 16μl IPTG and 20 mg/ml X-gal 40μl were added to LB solid media plates containing 100μg/ml ampicillin (Amp). The solutions were evenly distributed using a sterilized glass rod and stored upside down at 37° C in the dark to prevent photo degradation of the reagent.

2. Ligation products transform competent E. coli DH5α cells.

5μl of the ligation product was added to 100μl of competent E. coli DH5 $α$ cells. The solution was mixed thoroughly and then cooled on ice for 20 minutes. Competent cells were heat shocked by placing the tubes in a 42°C water bath for 90 seconds. Immediately after this, it was cooled on ice for 5 minutes. 900μL of antibioticfree LB liquid medium was added to a centrifuge tube and mixed thoroughly to ensure uniform growth. It was then incubated at 37°C for an hour in an incubator shaker at 200 rpm. Then centrifuged at 6000 rpm for 2 minutes. 800μl of the supernatant was discarded. The remaining 100μl of supernatant was used to sediment the bacteria. It was then inoculated onto an LB agar plate (solid medium) containing 100μg/ml ampicillin (Amp) and X-gal using a sterile glass rod. The plates were then incubated at 37° C for 16 hours.

3. Control conversion coating: 100μl of DH5α competent cell suspension was added to 2 sterile centrifuge tubes. Then 5μl of sterile water was added to the centrifuge tube, and 5μl of plasmid DNA solution was added to the second tube. The solution was stirred gently and then placed on ice for 30 minutes. It was then heat shocked in a 42°C water bath for 90 seconds and rapidly cooled on ice for 5 minutes. 100μL of liquid LB medium was added to the tubes and then incubated at 37°C with shaking for 30 min. A 50-μL aliquot from tube 1 was inoculated onto antibiotic- and non-antibiotic-containing plates. Another 50μL and 100μL aliquot from tube 2 were inoculated into the plates containing the antibiotic. The bacteria were allowed to dry for approximately 10 minutes before incubation.

2.3.3 Screening and identification of recombinant plasmids

2.3.3.1 PCR identification of recombinant colonies: Five white colonies were collected using a sterile pipette tip and transferred to a 1.5 mL Eppendorf tube containing liquid LB medium with antibiotics. The broth was cultured for 4 h, and colony PCR was performed with 1μl of bacterial broth. PCR amplification was performed in a reaction volume of 25μl using the master mix. PCR amplification was performed in a reaction volume of 25μl using the provided master mix. The reaction mixture (25μl) consisted of 2x PCR master mix (12.5μl), 10μM forward and reverse primers (1μl each), 9.5μl RNase-free water, and 1μl DNA. The amplification conditions were set at 94°C for 4 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 40 s. The final extension step was performed at 72°C for 8 min. This temperature was then maintained at 4°C until further analysis.

PCR products were analyzed on a 1% agarose gel to detect the presence of recombinant DNA.

2.3.3.2 Extraction and detection of recombinant plasmid DNA: Plasmid DNA extraction was performed using a TIANprep Mini Plasmid Kit. The TIANprep Mini Plasmid Kit is based on alkaline lysis technology followed by the adsorption of DNA onto a silica membrane in the presence of high salt content. Absolute ethanol was added to the PW buffer prior to use.

2.3.3.3 Enzymatic cleavage of the recombinant plasmid: Restriction digestion was performed using Thermo Scientific FastDigest enzymes. The restriction enzymes used in this experiment were NdeI (**5'...CA↓TATG...3' 3'...GTAT↑AC...5'**) and RhoI (**5'...C↓TCGAG... .3' 3'...GAGKT↑С...5'**). Thermo Scientific FastDigest Enzymes are an advanced line of restriction enzymes for rapid DNA digestion. All FastDigest™ enzymes are 100% active in FastDigest and FastDigest Green universal buffers are capable of digesting DNA in 5-15 minutes. This allows any combination of restriction enzymes to run simultaneously in a single reaction tube and eliminates the need for sequential digestions.

Enzymatic digestion was performed at room temperature in a reaction volume of 20μL. The reaction mixture contains 13 µl nuclease-free water, 2 µl 10X FastDigest, 4 µl plasmid DNA, 0.5 µl NdeI and 0.5 µl RhoI. The reaction mixture was gently stirred and centrifuged for 5s. then incubated at 37°C for 5 minutes in a water bath. The enzymes were then inactivated by heating for 5 min at 80°C.

The resulting restriction digest was then run on a 1% agarose gel to detect the presence of recombinant DNA.

III. RESULTS

3.1 ACQUISITION OF THE TARGET GENE

Figure 1 is an agarose gel electrophoresis of the PCR-amplified GFP gene. It can be seen that there is a clear bright band of about 750bp in size, which is the target product. Far left is the manufacturer, with no contamination in the sample hole and no scattered streaks; indicating that the overall enhancement was successful.

Figure1: Agarose gel electrophoresis profile of GFP gene

Figure 2 is an agarose gel electrophoresis profile after DNA extraction and purification. The results of our experiments are shown in the red box, the manufacturer is on the right. From the gel, it can be seen that the brightness of the DNA bands is relatively weak, and compared with the group on the right, the effect is poor. The faint band may be due to too much glue being cut out during the cutting process, resulting in a low concentration of recovered and purified DNA. However, increasing the absorption and elution times and the volume of PN buffer can improve the recovery of PCR products.

Figure 2: Agarose gel electrophoresis profile of GFP obtained from cleanup (recovery). **3.2 TRANSFORMATION INTO COMPETENT CELLS**

3.2.1 Screening of recombinant plasmids

After transformation into competent E. coli cells and culture on LB agar containing antibiotics, the growth of many white colonies can be clearly seen (Figure 3), indicating relatively successful transformation. The appearance of white spots indicates that the recombinant plasmid has transformed into Escherichia coli, and the appearance of blue indicates that it has not transformed into Escherichia coli. In this group of experiments, all spots were white, and no obvious blue spots were observed.

Figure 3: Transformed *E. coli* colonies

3.2.2 Isolation of recombinant plasmid

Figure 4 shows the experimental result of isolating the recombinant plasmid. The first well in Figure 4 is a 2 kb marker, the second well contains our recombinant plasmid. Our Research result shows that the brightness of the sample 1 band is relatively deep, which may be due to the high concentration of plasmid DNA during extraction, resulting in a higher concentration of plasmid supercoil molecules in the supernatant. There are no other bands on the electrophoresis map and generally no bands extend, indicating that the recombinant plasmid was not contaminated. Finally, during plasmid extraction, the CP3 spin column was washed twice and centrifuged with PW buffer before adding elution buffer. Double washing and removal of residual PW buffer results in proper separation of plasmid DNA from digested bacterial chromosomal DNA. Simply put, this results in higher quality and purer isolation of plasmid DNA.

Figure 4: Electrophoresis profile of recombinant plasmid

3.2.3 Enzymatic cleavage of the recombinant plasmid

The restriction enzymes used in this experiment were NdeI (5'...**CA↓TATG**...3' 3'...**GTAT↑AC**...5') and RhoI (5'...**C↓TCGAG**...3' 3'...**GAGKT↑С**...5'). The combination of two restriction enzymes in one tube eliminates the need for sequential digestion. Thermo Scientific FastDigest Enzymes are an advanced line of restriction enzymes for rapid DNA digestion. The marked box in Figure 5 shows the results of my group's double restriction enzyme digestion of a recombinant plasmid. Lane 1 from the gel is our plasmid DNA, lane 2 contains the 5 kb marker, and lane 3 is our restriction digested plasmid DNA. The results of the gel analysis show that the length of the two sample bands is about 3000 base pairs, and there is no obvious DNA band of the target gene. This indicates that double restriction enzyme digestion is not very successful.

Figure 5: Agarose gel electrophoresis profile of plasmid DNA restriction digest.

IV. DISCUSSION

In this experiment, we achieved good results in cloning the target gene at an early stage. PCR technology is mainly a method in which nucleic acids are denatured and amplified under high temperature conditions [10]. The reason why the target gene cloning gave good results is that during the sampling process, we strictly followed the requirements and set various parameters in the PCR program, which resulted in good results. The reason why the glue recycling strip is dull is because a lot of empty glue was cut out during glue recycling, resulting in a low concentration of recycled product.

In the blue and white spot screening experiment, we operated in a standardized manner and strictly followed the instructions. The results were good: all white spots and blue spots appear.

The results of experiments with this group of plasmid DNA showed that the brightness of the bands was relatively high, which was due to the fact that we eluted twice and the elution was sufficient, resulting in a higher concentration of plasmid supercoil molecules in the supernatant; There are no other bands in the electrophoresis map, and in general the bands do not lag, indicating that the recombinant plasmid is not contaminated and the separation effect between Escherichia coli DNA and plasmid DNA is good.

5. Conflict of Interest: Authors have no conflict of interest.

6. Acknowledgment: We would like to thank Mr. Fahad Ullah, for their helpful guidance on experiments and data analysis. Ahmad Haris Khan, Allah Madad and Izaz Hayat, Bello Abdullahi, Sulaiman Shahconducted experiments in the laboratory. We would also like to thank many graduate students who are not listed as coauthors.

7. PROSPECTS FOR THE FUTURE

In the process of research in the field of genetic engineering technology in recent years, the widely used methods of cloning recombinant plasmids and screening empty plasmid bacteria have the most classical applications. α Additional principle for blue and white spot screening, as well as rapid plasmid detection, colony PCR. in situ colony hybridization, insertional inactivation, restriction endonuclease digestion reaction [13] and so on. These methods have some disadvantages such as cumbersome operation, high costs, narrow application range and high false positive rate. Therefore, we need to explore a new plasmid vector that could screen bacterial recombinant plasmid clones more easily, efficiently, quickly, and reliably [6, 14, 15]. When mentioning Wei Qun's gene recombination method, blue spots appeared in the single enzyme digestion pattern, which was due to the self-crosslinking of the vector; The reason why only white spots cannot grow blue spots in the double enzymatic cleavage scheme is that the carrier after double enzymatic cleavage cannot self-join, and non-recombinant linear carriers cannot transform [11]. Although screening and identification of recombinant plasmids are widely used, further research and studies can be conducted at a later stage to better serve human studies at the molecular biology level [12, 16].

BIBLOGRAPHY

- [1]. Lee, H.-M., et al., Identification of efficient prokaryotic cell-penetrating peptides with applications in bacterial biotechnology. Communications Biology, 2021. **4**(1): p. 205.
- [2]. Rosano, G.L., E.S. Morales, and E.A. Ceccarelli, New tools for recombinant protein production in Escherichia coli: A 5‐ year update. Protein science, 2019. **28**(8): p. 1412-1422.
- [3]. Enghiad, B., et al., PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction. Nature communications, 2022. **13**(1): p. 2697.
- [4]. Rouwendal*, G.J., et al., Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage. Plant Molecular Biology, 1997. **33**: p. 989-999.
- [5]. Meng, D.-M., et al., Expression, purification and characterization of a recombinant antimicrobial peptide Hispidalin in Pichia pastoris. Protein Expression and Purification, 2019. **160**: p. 19-27.
- [6]. Denis-Quanquin, S., et al., Protein expression from synthetic genes: selection of clones using GFP. Journal of biotechnology, 2007. **131**(3): p. 223-230.
- [7]. Chalfie, M., et al., Green fluorescent protein as a marker for gene expression. Science, 1994. **263**(5148): p. 802-805.
- [8]. Bhat, S., et al., A ligation and restriction enzyme independent cloning technique: an alternative to conventional methods for cloning hard-to-clone gene segments in the influenza reverse genetics system. Virology journal, 2020. **17**: p. 1-9.
- [9]. Park, K.-S., S.-W. Park, and S.-Y. Choi, Improved T-vector for the cloning of PCR DNA using green fluorescent protein. Journal of microbiology and biotechnology, 2000. **10**(2): p. 264-266.
- [10]. Zhu, H., et al., PCR past, present and future. Biotechniques, 2020. **69**(4): p. 317-325.
- [11]. Szeberényi, J., Problem‐ solving test: Digestion of a plasmid with restriction endonucleases. Biochemistry and Molecular Biology Education, 2013. **41**(3): p. 189-190.
- [12]. Seo, J.H. and J.E. Bailey, Effects of recombinant plasmid content on growth properties and cloned gene product formation in Escherichia coli. Biotechnology and Bioengineering, 1985. **27**(12): p. 1668-1674.
- [13]. Remington, S.J., Green fluorescent protein: a perspective. Protein Science, 2011. **20**(9): p. 1509-1519.
- [14]. Uddin, T.M., et al., Antibiotic resistance in microbes: History, mechanisms, therapeutic strategies and future prospects. Journal of infection and public health, 2021. **14**(12): p. 1750-1766.
- [15]. Zimmer, M., GFP: from jellyfish to the Nobel prize and beyond. Chemical Society Reviews, 2009. **38**(10): p. 2823-2832.
- [16]. Van der Auwera, G.A., et al., From FastQ data to high‐ confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics, 2013. **43**(1): p. 11.10. 1-11.10. 33.