Issue 2 Vol 2 January 2021

Website: http://www.mpai.co.in/
Email: molpathindia@gmail.com

eNewsletter

AREA IN FOCUS

Common Techniques in
Molecular & Cytogenetic Diagnosis

Highlights:

Dear MPAI Members,

Hope you all are safe and healthy in this rapidly changing world. The role of molecular pathologists, microbiologists and biologists has increased tremendously over the last 6 months. As we are facing the pandemic, rapid and accurate molecular diagnostic technologies have become crucial for the screening, prevention and control of COVID-19.

I'm happy to note that in this period of pandemic also, the academic activity of the association is continuing so as to benefit everyone. In continuation with our online newsletter academic activity started in 2020 we are releasing the second issue this year. The first issue focused on BRCA1/BRCA2 whereas the theme of the second issue is **Techniques in Molecular biology** including Next Generation Sequencing. The need of cost effective molecular strategies and technologies have been never felt more.

The editor of this issue Dr Moushumi Suryavanshi and other authors of national and International repute have put together this focused letter for all of you. I am sure all of you shall benefit from their hard work.

Keep safe and healthy,

With best wishes

Sincerely

Dr Neeraj Arora

Secretary MPAI

Dear MPAI members,

Wishing you a very Happy and Healthy New Year 2021 !

The field of molecular has arisen from the convergence of work by geneticists, physicists, and structural chemists. Historically it began with the geneticist finding it difficult to analyse beyond Mendel's laws of segregation and independent assortment, the actual mechanisms of gene reproduction, mutation and expression. This gave birth to the term "molecular biology" in a 1938 by Warren Weaver, then the director of the Natural Sciences section of the Rockefeller Foundation. This new branch of science began to uncover many secrets concerning the ultimate units of the living cell….in which delicate modern techniques are being used to investigate ever more minute details of certain life processes (quoted in Olby 1994: 442).This branch has grown from just being "Molecular" to "Genomics". This fast track journey has been accomplished with the help of important "TOOLS". The list of these tools has grown in number and technological sophistication over the years. With this background I introduce the topic of our second eNewsletter "Common Molecular Techniques".

These are Corona Times. Suddenly science as a branch is driving the shifting paradigms of the new world order. The role of Molecular tools is perspicuous. This eNewsletter gives an overview of commoner techniques used in Molecular practice. These manuscripts are portrayal**s** by experts who have toiled with the nuisances of these technologies. The journey of instrumentation moves from Cytogenetic Techniques - Karyotyping and Fluorescent In situ Hybridization to the basic pillars of Molecular artistry- Polymerase Chain Reaction, Sanger Sequencing and Next Generation Sequencing.

I take this opportunity to thank the authors for prompt response in spite of these difficult times.

Warm Regards

Editor

Dr. Moushumi Suryavanshi MD PhD

Senior Consultant, Molecular Diagnostics

Rajiv Gandhi Cancer Institute and Research Centre, Delhi, India.

Bone marrow karyotyping – An Overview

Introduction

Cytogenetics is the study of morphology, structure, pathology, function, and behavior of chromosomes. Chromosomes are best studied at metaphase stage of cell cycle and can be obtained from either spontaneously dividing cells or from cells that may be induced to divide by using appropriate mitogens. Spontaneously dividing cells commonly used in routine diagnostic practice are derived from bone marrow, lymph node, peripheral blood with blasts, chorionic villi of placenta, and often in fetal or newborn blood. Cells from lymphoma or myeloma being mature cells need appropriate stimulation for them to enter cell division (1).

Dividing cells are arrested in metaphase stage and processed using hypotonic solution before being dropped on a glass slide for staining with appropriate. These stained slides are screened, and metaphases captured using a manual or automated microscope scanning system. Computer-generated images are used to arrange the chromosomes in pairs on karyogram. This procedure is called as karyotyping (2,3).

Procedure

Specimen requirement:

Bone marrow aspirate (BMA) is the sample of choice to study acquired cytogenetic abnormalities in hematological malignancies. BMA sample should be collected taking strict aseptic precautions in Na-heparin vacutainer. Cytogenetics should receive the first or the second draw as it is most cellular and least diluted. BMA specimens should never be exposed to extremes of temperature and transported at room temperature. In certain circumstances BMA sample may not be available like dry tap, patient not willing for a bone marrow procedure etc. In such situations peripheral blood sample may be used for looking at acquired cytogenetic abnormalities. The abnormal clone can be identified in such specimens, albeit not as often as in bone marrow (4).

Proper labelling of the sample is of paramount importance with sample being labelled with two patient identifiers. The quality of the sample received by the laboratory needs to be verified and any deviations like presence of clots, hemolysis documented. All the deviations in the quality of sample that may compromise the results need to be captured in the final report. BMA samples should be processed as soon as possible and preferably within 24 hours of collection.

Workflow:

Once the sample is received by the laboratory and is found to be optimal for processing, the sample undergoes various stages of processing as depicted in figure 1 before result generation. The various stages include – culture, hypotonic treatment, harvest, slide making, slide staining, metaphase screening and capturing, metaphase analysis or karyotyping (2).

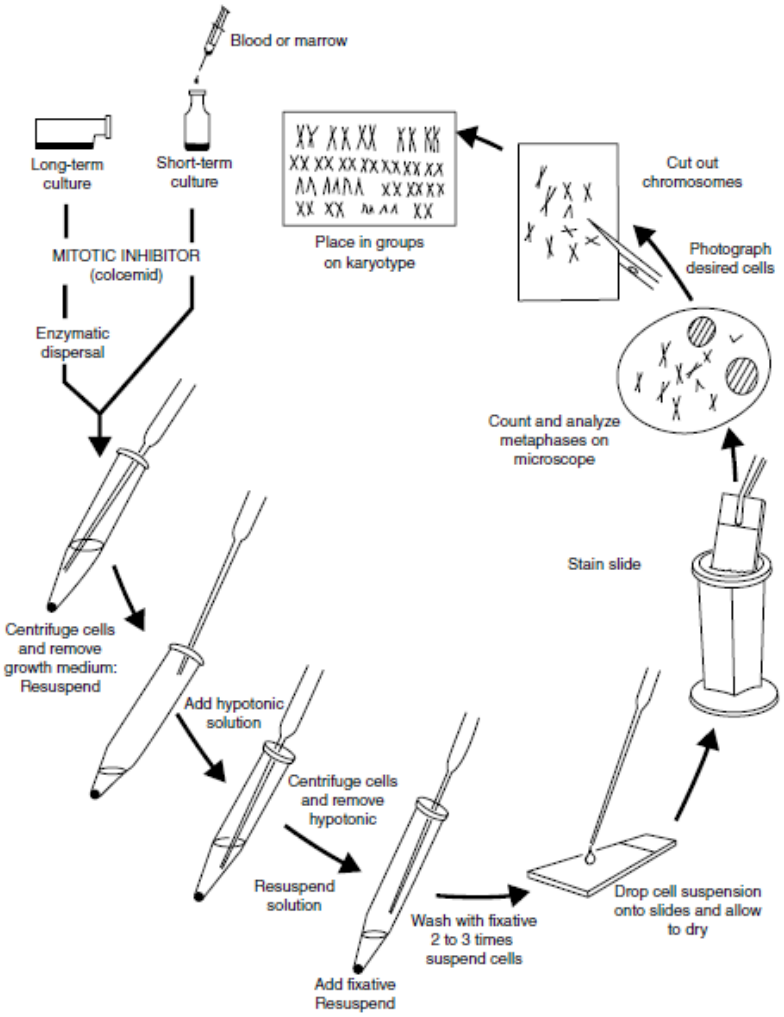


Figure 1: Steps in karyotyping studies

Culture:

BMA sample needs to be cultured to get metaphases for analysis. The duration of culture varies from 0 to 5 days depending upon the laboratories standard operating procedures, validation studies etc. Proper culture duration is critical to obtaining the results because abnormal tumor cell lines may be lost over time or may appear only after a few days in culture. Culturing is performed using strict aseptic techniques as microbes will cause culture failure.

BMA cultures being suspension cultures do not attach to the culture vessel and are of short duration. Most commonly, 15 ml conical bottom sterile tubes are used for setting up the cultures as they can be centrifuged during the harvest step. For acute leukemia and myeloid neoplasms, unstimulated BMA cultures are set up while stimulated cultures using mitogens like lipopolysaccharide, CD40L, immunostimulatory CpG-oligonucleotide DSP30 and IL-2 are used for mature cell neoplasms like lymphoma (5).

BMA cultures are open cultures as these allow exchange of gases with the atmosphere surrounding the culture and release gas phase byproducts of metabolism, which may be toxic

to the cells. The pH of the cultures is maintained in an incubator with a constant, controlled flow of gases, 5% CO2 in ambient air comprising of about 15–18% oxygen, or 5% CO2 in 2–5% O2 with the balance (93–95%) inert N2, from external gas tanks. BMA cultures are best grown at 37–37.5 °C, as growth curves fall off sharply over 38 °C, often resulting in cell death (2).

Of the scores of culture media available RPMI 1640 is the most used media supplemented with fetal bovine serum (2 to 25%) to provide proteins and growth factors. Antibiotics are used to prevent growth of Bacteria. As the BMA is a precious sample obtained by an invasive and painful procedure, these cultures should ideally be put in duplicate with one culture in each one of the two CO2 incubators with different supply of CO2 and UPS backup. The culture room may be fitted with UV lights to maintain room asepsis. All the culture set up work should be handled in a class II biosafety cabinet and the technical team should be well versed with the use and maintenance of BSC and CO2 incubators.

Harvesting of cultures

Harvesting cells for cytogenetic analysis is a stepwise procedure, that involves mitotic arrest, hypotonic treatment, and fixation of the cells.

Mitotic arrest:

The first step is to arrest cells in the mitotic stage, metaphase, by use of Colcemid which prevents formation of the spindle fiber apparatus required to pull the sister chromatids to opposite poles for incorporation into the two daughter cells (1).

Centrifugation

Next step involves centrifugation of the suspension cultures aiding in the process of change from culture media to hypotonic solution. Centrifugation is never complete, and several cells remain in the supernatant or are lost to damage from shear forces between cells or with the centrifuge tube walls. Therefore, the number of centrifugation steps should be kept to a minimum, hence a single centrifugation step of 10 minutes at 1200 rpm is used. Post centrifugation, most of the cells settle to the bottom of the 15 ml conical tubes and the supernatant is discarded.

Hypotonic treatment

This is followed by addition of hypotonic solution to the cell pellet. Hypotonic treatment is the most crucial step in the harvest process as it affects metaphase spreading, chromatid width and separation, and, in bloods and bone marrows, the elimination of red blood cells. Most used hypotonic agent is 0.56% or 0.075 M KCl (potassium chloride) (1). The KCl is maintained at 37^0 Cat the time of addition to the cell pellet to avoid cold shock that can disrupt and damage cells. Prewarming the hypotonic solution to 37 °C also increases effectiveness by speeding up water transport across the cell membrane and possibly by softening the cytoplasmic membrane, giving it more stretching capability and swelling cells. Most laboratories use 10 to 25 minutes time for hypotonic treatment and that needs to be standardized at individual laboratories.

Fixation:

Fixation is the process which removes water from the cells, killing and preserving them, hardening membranes and chromatin and preparing the chromosomes for the banding as banding patters are only possible in 3:1 methanol–acetic acid solution, Carnoy's fixative (1,2).

Addition of first few drops of fixative creates turbulence, hence they are added slowly, as drops one by one to avoid cell rupture, followed by more quick addition of the fixative. The cell pellet is mixed thoroughly to avoid any cell clumping. Cold fixative improves chromosome morphology, and in our laboratory the fixed cell pellets are kept in freezer for 1 to 24 hrs depending upon the need of slide making. As the concentration of methanol and acetic acid changes with time it is recommended to use freshly prepared fixative at least daily for the harvest procedure.

Slide Making

Once cells have been well-fixed in 3:1 methanol–acetic acid, they are dropped onto glass slide and air dried for optimal chromosome spreading and morphology.

Slide making is a "technical art" as it involves constant intervention by a competent technologist to decide optimum techniques for metaphase spreading as drying varies with humidity and temperature of the slide making room as well as the rate of drying. The humidity of the slide making area varies from 40% to 65% with the temperature range of 21^0 C to 25^0 C. Ideal temperature and humidity needs to be determined by the laboratory and these values change almost daily depending upon the climate and environmental temperature.

The cell pellet may be dropped on to a glass slide in different ways. Most laboratories place the glass slide at an angle of 30^0 from the surface. The dropping method varies, with some laboratories dropping the pellet on the glass slide dipped in chilled alcohol and water, from an height of around 1 feet (range from 0.1 meter to 0.31 meter) while others spread the pellet along the long side of the glass slide and allow the pellet to roll on to the film formed between alcohol and water (6). Post dropping the slide is flushed with fixative to remove excess water and help in drying.

Once the cells are dropped onto a glass slide, chromosomes are contained in enlarged cells with much thinner cell membranes and supported by the layer of fixative. As the fixative evaporates, the layer of fixative becomes thinner and the meniscus pushes down on the top of the cell, enlarging the area of the cell and pressing the metaphase chromosomes between the upper and lower membranes, spreading them out. This physical relaxation and collapse of the cell membrane takes some time, stretching the chromosomes and determining the extent of banding level. If the fixative dries before the cytoplasm has relaxed and cells have spread out, there will be a visible cytoplasmic background and the cells and chromosomes will be thicker shorter, and often poorly spread out (7).

Once the methanol–acetic acid-fixed chromosomes are dried onto a glass slide, they stick fast until physically scraped-off, and staining may be accomplished without losing cells from the slides. We at our laboratory use a slide-warming table at 37^0 C to manipulate the speed of drying and thus the spread and banding levels of the chromosomes.

The aim of the slide making is to achieve medium gray to dark gray chromosomes, with little chromosome scattering, minimum chromosome overlaps, thin or absent cytoplasmic

background, minimal chromatid splits, adequate density of metaphases and cells and no debris in the slide background.

Once the slides are done, they are labelled with the case / patient identifier and are ready for next steps of aging, banding, and staining.

Slide aging, banding, and staining

Slide aging

The slides are either left to dry for a period of 3 to 4 days or are heated at 40–60 °C for 12 to 24 hrs or 20 minutes at 90–95 °C before they are banded and stained as this allows for crisp staining of the chromosomes. This process is called aging. Till the slides are aged they do not stain and band well. It is postulated that aging may cause the remaining water to dry away and oxidize the protein sulfhydryl groups (8). The ideal temperature and duration of aging should be determined by individual laboratories.

Banding and staining

There are three stains capable of differentiating chromosomes: G-, R-, and Q-banding, however in this discussion we will restrict ourselves to G-banding which is the most commonly used staining technique in India and across the world.

The G-banding using Geimsa was introduced by Sumner et.al in 1971 (9). Most used banding and staining technique is the GTG "G banding by trypsin and geimsa". In this technique the glass slide with metaphases is exposed to trypsin enzyme in buffer at 37 °C for a couple of seconds and the action of the enzyme is stopped by snap chilling the slide in cold water or phosphate buffered saline. This step is followed by staining with geimsa stain for 2 to 5 minutes. The concentration of enzyme used, duration of exposure to trypsin and the staining duration affects the final banding pattern of the chromosomes. The duration of enzymatic action and staining also varies from day to day as well as sample to sample and is challenging. The Geimsa positive bands i.e. G – dark bands are rich in AT late replicating and gene poor while Geimsa negative bands i.e. G – light are GC rich, early replicating and relatively gene rich (10). A banding level of around 400 bphs (bands per haploid set) should be achieved for analysis while staining (2).

Metaphase analysis / microscopy

Once chromosome preparations have been banded, the slide is scanned in a methodical, field-by-field manner, either by traveling up and down or right to left on the slide using a low power lens (10×), searching for metaphase cells. This can be achieved manually by a trained technologist or by automated scanning systems. The metaphases of interest are then captured under oil immersion at 100x magnification.

The stage coordinates of each of the captured and analyzed metaphases needs to be recorded so that the metaphases may be reidentified later for any additional testing like FISH or by another technologist. To ensure that the coordinates are captured correctly every time when a slide is scanned and captured, the orientation of the slide loading on to the microscope stage should be kept uniform and the same may be achieved by ex. keeping the slide label on the left hand side while loading the slide.

Once the metaphases are captured, they can be analyzed by using commercially available softwares. The chromosomes are identified, separated, and then arranged in a karyogram. Depending upon the recommendations of international bodies, national bodies as well as the choice of the laboratory minimum number of metaphases that need to be scored, analyzed, and karyotyped needs to be defined. This also takes into consideration the type of sample that is being analyzed. For BMA studies a minimum of 20 metaphases needs to be scored and analyzed.

The term "Analyzed cells" is defined as banded metaphase cells in which the individual chromosomes are evaluated in their entirety, either at the microscope or from intact digitized images or photographic prints of intact cells (4).

Term "Karyogrammed cells / Karyotyped cells" are defined as the cutout and paired chromosomes from photograph(s) or computer-generated image(s) from a single cell following the format in An International System for Human Cytogenomic Nomenclature (ISCN) (4).

Term "Scored cells" refer to cells evaluated for the presence or absence of a specific cytogenetic feature, usually indicated by either a particular clinical history or by the finding of one or two abnormal cells during the course of a study (4).

Chromosomes are always arranged on the karyogram with the short arms uppermost. Groups are arranged in alphabetical order. The sex chromosomes may be placed together at the end or separated to their groups; either way is acceptable (2). Eye karyotyping refers to analysis of the chromosomes in a metaphase band by band under the microscope. In the authors opinion, this is a very tedious process especially in laboratories with decent load of BMA samples may require a greater number of microscopes.

Indications of Bone marrow karyotyping

Bone marrow karyotyping is an integral part of work up of a patient with hematological malignancies, including acute myeloid and lymphoid leukemia, myelodysplastic syndrome, myeloproliferative neoplasm, Myeloid/lymphoid neoplasms with eosinophilia and gene rearrangement, Myelodysplastic/myeloproliferative neoplasms, aplastic anemia, Chronic lymphocytic leukemia and myeloma, when appropriately stimulated. Bone marrow karyotyping has very limited value in mature lymphoid neoplasm

Interpretation of results

A karyogram gives complete overview of the whole genome of a cell in one go as all the chromosomes are immediately available for analysis. This provides a holistic approach of testing rather being directed as in FISH, PCR or Sanger sequencing-based testing. The later techniques require prior knowledge of the abnormality of interest, are very specific and limited to the abnormality in question.

Karyotyping allows for detection of many numerical and structural abnormalities both previously known as well as unknown and is currently the only available routinely used method with such a capability. Karyotyping allows for detection of abnormalities like hyperdiploidy, hypodiploidy, monosomal karyotype, balanced and unbalanced translocations, complex karyotypes. When used judiciously along with FISH or PCR it makes a powerful diagnostic and prognostic tool in hematological malignancies work-up. Clonal evolution

which is a known phenomenon in cancers can best be identified by karyotyping rather than any other technique.

The cytogenetic abnormalities as detected by karyotyping should be reported as per the current ISCN (International system for Human cytogenetic nomenclature) 2016 (3). Clonal origin is inferred by the presence of at least two cells containing the same extra chromosome(s) or structural chromosome abnormality or by the presence of at least three cells that have lost the same chromosome (4). Term “mosaicism” should be reserved for constitutional genetics.

During analysis if a constitutional abnormality is observed, analysis of a mitogen-stimulated peripheral blood sample during remission should be strongly recommended to confirm that the abnormality is constitutional and not clonal. Also at the time of initial diagnosis, if a single abnormal metaphase cell cannot be used as evidence of clonality, even if potentially significant, unless there is strong supporting evidence of clonality for the same abnormality by either FISH or other molecular technique.

The final cytogenetic report of hematological acquired chromosomal abnormalities should contain the information as per the recommendations of national (NABL) and international bodies (ACMG) (4,11).

Sources of error in analysis and reporting

As cytogenetics is an integral part of prognosis, diagnosis, and treatment decision making, it is of paramount importance to avoid errors. Almost all stages of a cytogenetic study involve subjectivity and individual expertise on the part of the cytogeneticist, it is important to be aware of some potential sources of errors (2).

1. Case mix up at any stage of processing
2. Handling multiple case simultaneously at any step of processing.
   We practice culture set up procedure being done in the presence of two technologist, with one setting up and the other guiding and checking the labelling and details of each case, culture tubes and vacutainer.
3. Poor growth of cultures and poor banding leading to poor quality of metaphases analyzed
4. Missing of a clone or a sub clone – due to poor morphology or inadequate number of metaphases being examined. In our personal experience many times the normal cells over grow the abnormal ones and inadequate number of metaphases analyzed missed the abnormal clone.
5. Cryptic abnormalities – Ex. t(12;21) ot t(11;v) will be picked by FISH.

Conclusion: BMA karyotyping is a tedious but integral part of work up of hematological neoplasms giving the most holistic view of the cancer genome.

References

1. Howe B, Umrigar A, Tsien F. Chromosome preparation from cultured cells. J Vis Exp. 2014;(83):e50203. Published 2014 Jan 28. doi:10.3791/50203

2. Lawce H.J, Brown M.G. (2017). Cytogenetics: An overview. In M.S. Arsham, M.J. Barch, Lawce H.J. The AGT Cytogenetics Laboratory Manual (pp 25-85). New Jersey: John Wiley & Sons, Inc.
3. Jean Mcgowan-Jordan, Alex Simons, Manuel J. Schmid. Medicine 2016. ISCN 2016: An International System for Human Cytogenomic Nomenclature (2016)
4. Mikhail FM, Heerema NA, Rao KW, Burnside RD, Cherry AM, Cooley LD. Section E6.1-6.4 of the ACMG technical standards and guidelines: chromosome studies of neoplastic blood and bone marrow-acquired chromosomal abnormalities. Genet Med 2016; 18: 635–642, doi: 10.1038/gim.2016.50.
5. Haferlach C, Dicker F, Schnittger S, Kern W, Haferlach T (2007): Comprehensive genetic characterization of CLL: a study on 506 cases analysed with chromosome banding analysis, interphase FISH, IgV(H) status and immunophenotyping. Leukemia 21:2442–2451.
6. Lambson B, Mendelow B, Bernstein R. Metaphase spreading in chromosome preparations is minimally affected by mechanical disruption. Karyogram 1986; 12:30–32.
7. Hliscs R, Muhlig P, Claussen U. The spreading of metaphases is a slow process which leads to a stretching of chromosomes. Cytogenet Cell Genet 1997; 76(3–4):167–171.
8. Evans HJ. Some facts and fancies relating to chromosome structure in man. Advances in Human Genetics 1977; 8:347–438.
9. Sumner AT, Evans HJ, Buckland RA. New technique for distinguishing between human chromosomes. Nature New Biol 1971;232:31–32.
10. Sumner, A.T., The nature and mechanisms of chromosome-banding. Cancer Genetics and Cytogenetics, 1982. 6(1): p. 59-87.
11. National Accreditation Board for Testing and Calibration Laboratories, NABL 112, Specific Criteria for Accreditation of Medical Laboratories. Downloaded from https://www.nablindia.org/nabl/index.php?c=publicaccredationdoc&m=index&docType=both&Itemid=199, on 1st Apr 2016.

Author

Dr. Anurag Gupta

MD (Path), Fellowship in Clinical and Cancer

Cytogenetics Consultant Molecular Pathology and

Head Cytogenetics Dept. Of Cytogenetics

AMPATH, Hyderabad, Telangana

Fluorescence In-situ Hybridization (FISH) and its applications

Background/ Introduction: In 1968, cytogenetics boomed the molecular era with In-situ hybridization (ISH) which used radioactivity to study chromosomes (mouse and drosophila) (1). However, with issues concerning safety, stability and for ease of detection, radioactivity was critically replaced by fluorescence and Fluorescence in situ hybridization (FISH) was conceptualized (Fig.1). By definition, FISH is a molecular cytogenetic technique for the spatial detection and quantification of nucleic acids in their cellular environment (in-situ). It represents a methodology that involves hybridization of a single-stranded probe (fluorescently labeled) to a single-stranded DNA target by high-affinity base complementation (A-T and G-C).

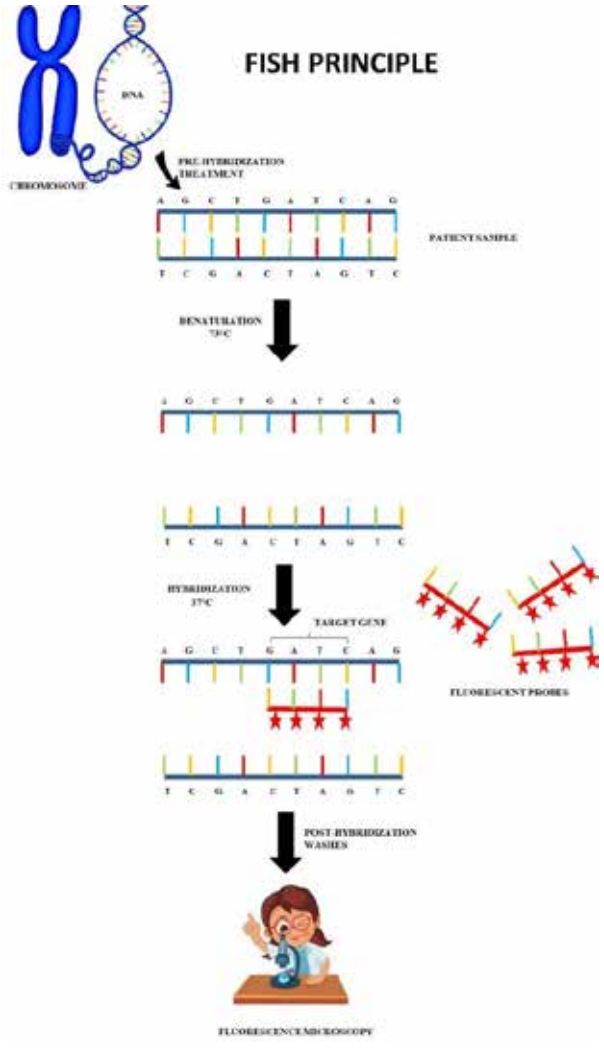


Figure1: Principle of Fluorescence in-situ hybridization (FISH).

It has thus become a powerful tool in identifying the location of a cloned DNA sequence on a metaphase chromosome. According to literature, chromosomal abnormalities occur in approximately 1 per 60 live

births (2,3), 60-80% of all miscarriages (4,5), 10% of stillbirths (2,6), 13% of individuals with congenital heart disease (7) and 3-6% of infertility cases (2). Cytogenetic analysis is also a routine and has shown to have both diagnostic and prognostic significance in malignancy cases. In recent years, applications of FISH have grown vastly and are expanding further with the development of newer commercial probes (binary colour combinations or up to 6-colour assay). The most common FISH probes in the clinical setup include fusion, break apart, centromere enumeration and deletion probes. Fusion probes detect gene translocations involving 2 known genes (e.g. *BCR* and *ABL1* fusion in CML), translocations between a known gene and an unknown partner is identified by break-apart probes (e.g. *MLL* translocation in AML and ALL), monosomies/trisomies are recognized by centromere enumeration probes (e.g. Trisomy 4, 10, 17 in B-ALL) gene deletions by deletion probes (e.g. 13q deletion in MM) (Fig.2) and whole chromosome paints and probes specific to alpha satellite and sub-telomeric regions are used for confirmation of abnormalities. Given that FISH can detect submicroscopic cytogenetic deletions, simple or complex translocations, inversions, gene amplifications and numerical abnormalities on both metaphase and interphase cells, it has been incorporated in the diagnostic algorithm for various diseases.

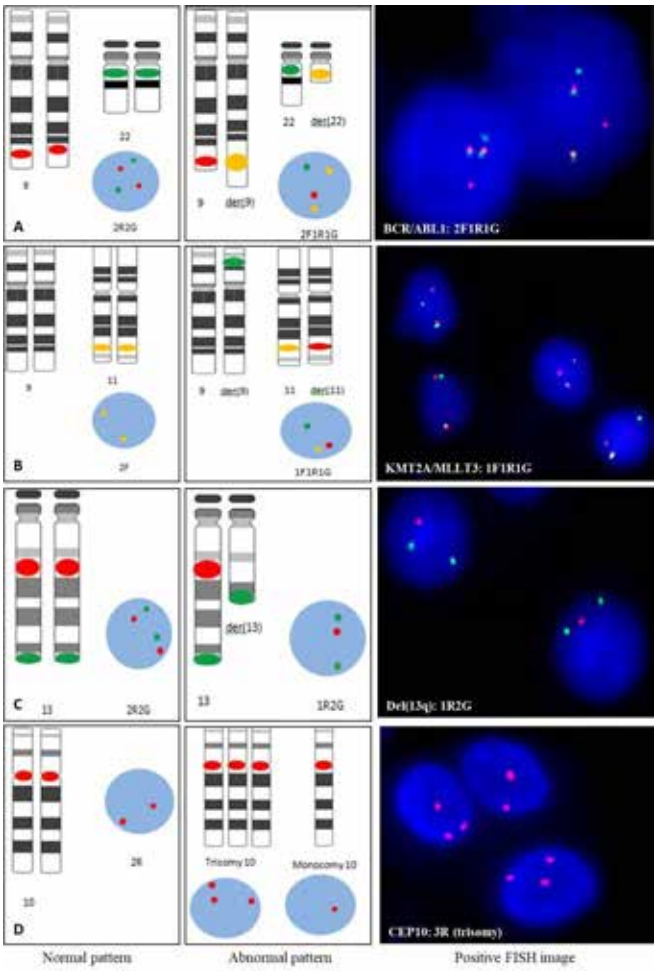


Figure 2: Types of FISH probes and their signal patterns

A. Dual colour dual fusion probe, B. Dual colour break-apart probe, C. Dual colour deletion probes,

D. Centromere enumeration probes.

Sample Type: FISH can be performed on blood, bone marrow (BM), amniotic fluid, product of conception, CVS, buccal smears /BM aspirate smears, urine sediments, biopsies or archival fixed tissue. The sample may be fresh, harvested, acid-alcohol fixed, flow-sorted, formalin-fixed or paraffin-embedded tissue (8). Sodium heparin is preferably used for sample collection. Of utmost importance, however, is the choice of sample e.g. blood is an inappropriate source in leukemia if it does not contain abnormal blast cells. Likewise, column or flow-sorted CD138+ plasma cells are essential for a correct diagnosis of multiple myeloma- a plasma cell disorder.

Sample processing and Methodology: Culturing is the first step in the FISH protocol. The sample is planted/cultured in appropriate media considering cellularity (especially for BM samples) followed by colcemid treatment to arrest cells in metaphase (for metaphase FISH). Cells are centrifuged into a pellet, harvested using warm KCl and fixed with Carnoy's fixative i.e Methanol:Glacial acetic acid (3:1). The KCL provides a hypotonic environment (causing cells to swell) and fixation removes cytoplasmic material, both in turn making the nuclear DNA more accessible to fluorescent probes. The FISH procedure includes a labeled DNA probe and a target DNA sequence (sample). Labelling of the probe is usually achieved by nick translation, random priming or PCR using either the direct (nucleotide modified to contain fluorophore) or the indirect labelling (through haptens) strategy. If hapten associated probes are utilized, an extra step for visualization is required using an enzymatic or immunological detection system. Although FISH may seem faster with directly labeled probes, indirect labelling offers the advantage of signal amplification. However, at present with the availability of commercial probes, in-house probe labelling has become less common.

The FISH procedure is composed of 3 elements- pre-hybridization treatment, denaturation followed by hybridization and post hybridization washes. Pre-hybridization involves treatment with pepsin to enhance probe penetration followed by dehydration using alcohol grades (70%, 85% and 100%). Denaturation happens at 73°C followed by annealing. Most probes hybridize well at 37°C, although some protocols call for reannealment at 42°C to decrease cross-hybridization (especially for probes specific to alpha satellite or other highly repetitive sequences). To prevent drying, humidified chambers are generally used. After hybridization (few hours to overnight), the slides are placed into Coplin jars with post-hybridization buffer to remove the unbound and non-specifically bound probe. This includes treatment with SSC solutions (0.4X) at 72°C followed by room temperature 2X SSC or formamide washes (usually 50% formamide in 2X SSC) at 45°C followed by a rinse in 2X SSC. The last step involves counterstaining with DAPI- 4',6-diamino-2-phenylindole which stains the background to enhance contrast. For FFPE specimens an extra step for deparaffinization (most commonly Xylene) is incorporated before hybridization. Through all the variations, basic steps in FISH include specimen processing, pre-treatment to harden chromatin, denaturation of probe and target, reannealing, followed by post-hybridization washes under stringent conditions and counterstaining.

Applications: FISH is used in clinical cytogenetics, most extensively in cases of microdeletion syndromes or when mitotic cells/ metaphases are not available for chromosomal analysis (eg Product of Conception). It has also been used to improve outcomes in assisted reproductive technologies by aneuploidy testing in high-risk indicators- maternal age, previous history of aneuploidies, multiple miscarriages, recurrent implantation failure and translocation status when one parent is a carrier of a balanced translocation (8).

A set of probes for most specific abnormalities seen at term (aneuploidy of 13, 16, 18, 21, 22, X, Y), in spontaneous abortions and in day 3 embryos are used for aneuploidy screening. Munne S. suggests FISH testing to detect 70% of aneuploidies found in spontaneous abortions (9). It has become all the more significant in detecting low mosaics and in sex developmental disorders. Likewise, FISH on the pre-implantation cleavage-stage embryos is the technique of choice to detect chromosome imbalances and select female embryos in families with X-linked disease. Most infertility clinics prefer doing interphase FISH on sperm nuclei to study sperm aneuploidies which may be a reason for male infertility. Some examples of diseases diagnosed by FISH include microdeletion syndromes like Prader-Willi syndrome, Angelman syndrome, DiGeorge syndrome, Cri-du-chat, William syndrome, Smith-magenis syndrome and aneuploidies like trisomies of 13, 16, 18, Down syndrome (trisomy 21), Klinefelter syndrome (XXY) and Turner syndrome(X) (10).

FISH has especially revolutionized the study of oncology specimens; classification, prognosis and treatment. It is faster than conventional cytogenetics, which can take up to several days considering culture and harvest. It has paved the way for alteration specific therapeutic strategies- one good example is the use of Herceptin in *HER2* amplification for an improved response (11). Similarly, FISH has become the gold standard test for the detection of anaplastic lymphoma kinase (*ALK*) gene rearrangements in lung adenocarcinoma which shows a good response with *ALK* inhibitors (12). Of recent, focus on *ROS* testing by FISH is growing given the availability of an effective *ROS* inhibitor- Crizotinib (13). It is also routinely being used to detect aneuploidies for chromosomes 3, 7, 17 and loss of 9p21 loci in Transitional cell carcinoma (TCC) which constitutes 90% human bladder cancers (14). Besides, it is a useful adjunct in the diagnostic assessment of Ewing sarcoma (*EWS*), synovial sarcoma (*SS18/SSX*), rhabdomyosarcoma (*FOXO1*), and other soft tissue sarcomas (*MDM2*) (15). Identifying co-deletions of chromosomal arms of 1p and 19q (favourable prognostic markers) a characteristic and early genetic event in oligodendroglioma tumours is another application of FISH (16).

Seeing hematolymphoid malignancies, cytogenetic diagnosis has become the standard of care for most leukemias and lymphomas (17). The WHO 2016 prognostically classifies B-ALL and AML based on presence of chromosomal abnormalities e.g. AML with *RUNX1/RUNX1T1* translocation (favourable risk), B-ALL with *TCF3/PBX1* translocation (poor-intermediate risk) etc. Moreover, FISH is helpful in detecting cryptic abnormalities that go unidentified by G-banding e.g. *ETV6/RUNX1* rearrangement. A wide spectrum of genetic aberrations found in 39% to 91% of Ph-like ALL affecting genes encoding kinases (*JAK1*, *JAK2*, *JAK3*, *ABL1*, *ABL2*, and *TYK2*), cytokine and growth factor receptors (*CRLF2*, *PDGFRB*, *EPOR*, *IL7R*, *CSF1R*, *NTRK3*, and *FLT3*), and signalling mediators and regulators (*KRAS*, *NRAS*, *BRAF*, *PTPN11*, *NF1*, and *SH2B3*) are identified by FISH. Majority of these alterations are amenable to targeted therapy (tyrosine kinase inhibitors) and thus their identification becomes essential. Likewise, deletion of 13q14 region, found in more than 50% of CLL patients, is the most common cytogenetic abnormality detected by FISH and has historically been associated with good prognosis (18). Intrachromosomal amplification is yet another abnormality that is identified solely by FISH. Besides, FISH has become the go-to test where quick diagnosis (e.g. APL) can be a matter of life or death. For these cases, FISH may be done in 4-5 hours using a short hybridization time. Interphase FISH has especially become useful for indolent cancers such as CLL and MM that proliferate slowly and for cells that are difficult to culture. FISH is routinely used for CML and in fact variant CML is primarily identified by FISH. A recent advanced FISH strategy in leukemia diagnosis is the use of probe panels to query multiple disease-specific loci. These include the most recurrent abnormalities associated with the disease e. Lymphomas- *IgH/CCND1* (Mantle cell lymphomas), *MYC/IGH* (Burkitt's lymphoma), *IGH/BCL2* (Follicular lymphoma), *BCL6*, *BCL3*; monosomy/trisomy 8, monosomy and deletion of 5q, 7q and 20q in MDS, wherein each abnormality is prognostically relevant. Moreover,

spatial transcriptomics has leveled up with improved sequential FISH approach allowing imaging expressions of 10,000 genes in single cells (19).

Future Perspective/Conclusion: Currently, FISH studies are essential for molecular classification and predicting prognosis in many diseases. Although, FISH may have paved way for newer hybridization technologies- comparative genome hybridization, spectral karyotype, M-FISH and SNP arrays it still stands important given the fact that it can detect low-level mosaics, balanced translocations, gene inversions, microdeletions etc. Moreover, availability of robust commercial probes, low cost and a shorter turn-around time justifies its inclusion in diagnostics.

References:

1. Gall J. The origin of in situ hybridization- A personal history. Methods. 2016; 98:4-9.
2. Driscoll D, Gross S. Prenatal screening for aneuploidy. New Engl. J. Med. 2009; 360:2556-2562.
3. Nussbuam, R., McInnes R, Willard H. Genetics in Medicine. 7th Edition. 2007, Saunders/Elsevier, Philadelphia.
4. Ljunger E, Cnattingius S, Lundin C, et al. Chromosomal abnormalities in first-trimester miscarriages. Acta Obstet. Gynecol. Scand. 2005;84(11):1103-1107.
5. Kwinecka-Dmitri, B, Zakrzewska M, Latos-Bieleńska A, et al. Frequency of chromosomal aberrations in material from abortions. Ginecol. Pol. 2010; 81(12):896-901.
6. Reddy U, Goldenberg R, Silver R, et al. Stillbirth classification-Developing an international consensus for research: Executive summary of a national institute of child health and human development workshop. Obstet. Gynecol. 2009; 114(4):901-914.
7. Pierpoint M, Basson C, Benson DW Jr, et al. Genetic basis for congenital heart defects: Current knowledge. Circulation. 2007; 115(23):3015.
8. Arsham M, Barch M, Lawce H. The AGT Cytogenetics Laboratory Manual. 4th edition. 2017, John Wiley & Sons, Inc., Hoboken, New Jersey.
9. Munne S. Predictability of preimplantation genetic diagnosis of aneuploidies and translocations on prospective attempts. Reprod BioMed Online. 2004; 9(6):645-651.
10. Halder A, Jain M, Chaudhary I, et al. Fluorescence *in situ* hybridization (FISH) using non-commercial probes in the diagnosis of clinically suspected microdeletion syndromes. Indian J Med Res. 2013; 138(1): 135–142.
11. Hwang K, Weedin J, Lamb D. The use of fluorescent in situ hybridization in male infertility. Ther Adv Urol. 2010; 2(4):157–169.
12. Kipp B, Halling K. Clinical FISH Testing for the Diagnosis of Solid Tumors. In: Highsmith, Jr. W. (eds) Molecular Diagnostics. Mol Transl Med. 2014, Humana Press, New York.
13. Martin V, Bernasconi B, Merlo E, et al. ALK testing in lung adenocarcinoma: technical aspects to improve FISH evaluation in daily practice. J. Thoracic Oncology. 2015; 10(4):595-602.
14. Bubendorf L, Buttner R, Al-Dayel F, et al. Testing for ROS1 in non-small cell lung cancer: A review with recommendations. Virchows Arch. 2016; 469(5): 489-503.
15. Dodurga Y, Avcı C, Yılmaz S, et al. UroVysion fluorescence in situ hybridization (UroVysion FISH) assay for detection of bladder cancer in voided urine of Turkish patients: a preliminary study. Contemp Oncol (Pozn). 2013; 17(2):156–160.
16. Asif A, Mustaq S, Hassan U, et al. Fluorescence in situ hybridization for differential diagnosis of soft tissue sarcomas. Asian Pac J Cancer Prev. 2018; 19(3):655-660.
17. Woehrer A, Dander P, Haberler C, et al. FISH-based detection of 1p 19q codeletion in oligodendroglial tumours: procedures and protocol for neuropathological practice-a publication under

the auspices of the research committee of the European confederation of neuropathological societies (Euro-CNS). Clin Neuropathol. 2011; 30(2):47-55.
18. Swerdlow S, Campo E, Harris N et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasm. Blood. 2016; 127(20):2375–2390.
19. Puiggros A, Blanco G, Espinet B. Genetic Abnormalities in Chronic Lymphocytic Leukemia: Where We Are and Where We Go. 2014; 2014:1-13.
20. Sansone A. Spatial transcriptomics levels up. Nat. Methods. 2019; 16:458.

Authors : Dhanlaxmi Shetty and Elizabeth Talker

Cancer Cytogenetics Department, Advanced Centre for Treatment, Research and Education in Cancer(ACTREC), Tata Memorial Centre, Kharghar, Navi Mumbai- 410210

Polymerase Chain Reaction: Tool for Genomics Analysis

Abstract

Polymerase Chain Reaction (PCR) is a biotechnological breakthrough technique developed in 1983 by Kary Mullis and was patented in 1985. This technique continues to play significant and major roles in science even today. Since characterization of diverse species in the ecosystem remains as a major scientific interest among the researchers to understand the critical functioning of these species, PCR is considered as an indispensable tool for understanding the functioning of these biological systems. This method utilizes the deoxyribonucleic acid (DNA) polymerase enzyme which specifically amplifies the target nucleic acid sequences providing a sufficient amount of product that can be used for further downstream methods such as visualization of target nucleic acids on an agarose gel. This technology basically works with the aim of exponentially amplifying the target sequence to millions of copies. There are huge number of applications of PCR as it has now become an essential technique in cellular and molecular biology. Also, PCR has wide applications in diagnostic purposes to detect specific DNA sequence of a particular biological organism or fluid. PCR, also essentially used for site-directed mutagenesis. PCR has wide applications in genotyping, microarray analysis, RNA sequencing, analyzing single nucleotide polymorphism, allele analysis and gene expression analysis, DNA methylation analysis, DNA fingerprinting, virus detection etc. The PCR techniques have been modified over time and thus different variants of this technique has been developed such as real-time PCR, PCR Arrays, Chip PCR, Reverse transcriptase PCR, Touch down PCR, Digital PCR, Inverse PCR, Colony PCR, Nanoparticle assisted PCR, Methylation specific PCR, RNase H-dependent PCR and many more.

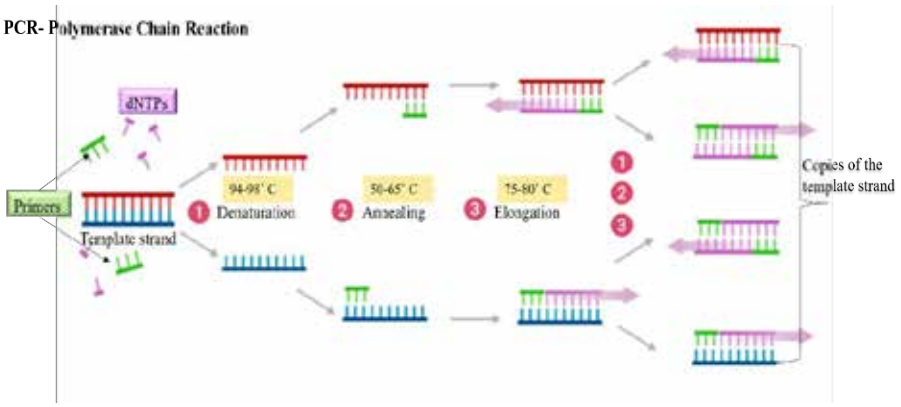
The principle behind PCR reaction

The basic aim of PCR reactions is to amplify target nucleic acid sequences and this purpose requires key players such as the DNA polymerase, primers and nucleotides. The template required for a PCR reaction i.e., the target nucleic acid sequence may be of any source such as DNA, RNA or cDNA (complementary DNA). Primers are 15-40 bases long nucleotide sequences with the application of annealing the opposite strands of the target nucleotide sequence. PCR uses the thermostable DNA polymerase commonly known as the Taq DNA polymerase (Taq-Thermus aquaticus). This polymerase helps in the addition of deoxyribonucleoside triphosphates (dNTPs) onto the ends of the primer sequence as to extend the nucleic acid sequence based on the target sequence. Most of the steps in PCR rely on thermal cycling consisting of a series of 20-40 repeated temperature changes. This allows the reactant mixture to undergo repeated cycles of heating and cooling to undergo different temperature dependent reactions. This temperature and length of time applied depend on different factors such as the enzyme used for DNA synthesis, melting temperature of primers (*Tm)* and concentration of dNTPs in the reaction. The most common steps in all types of PCR are:

Initialization: This step is mostly required for DNA polymerase required by hot-start PCR where the reaction chamber is heated to a temperature of 94-96 ^{o}C and is held for 1-10 minutes. This high temperature used in PCR reactions make the need of thermostable DNA polymerase.

Denaturation: This step melts or denatures the double stranded template DNA at a temperature of 94-98 ^{o}C. this step yields two-single stranded DNA molecules.

Annealing: In this step the reaction temperature is lowered to 50-65 ^{o}C allowing the annealing of primers to each of the single stranded DNA template strands. Two different primers which are single stranded themselves complement to 3' end of each template strand. It is very critical to determine the annealing temperature as it strongly affects the efficiency and specificity. The typical annealing temperature is 3-5 ^{o}C below the *Tm* of the primers used.



Extension: The temperature at this step depends on the temperature at which optimum activity of DNA polymerase is observed. This temperature approximately lies in the range of 75-80 ^{o}C. this crucial step elongates the nucleotide sequence based on the complementary strand by adding free dNTPs.

The process of denaturation, annealing and elongation make a single cycle. To form multiple copies of the target sequence multiple cycles are required. The number of copies is calculated on the basis of 2^{n} where n is the number of cycles.

***Final hold*:** This final step after the completion of the cycles cools the chamber to 4 ^{o}C for an indefinite time for the short time storage of PCR products.

One of the further downstream application of the product generated is to check the target DNA generated by PCR via agarose gel electrophoresis.

Variants of PCR

Depending on the same ideology of PCR steps, PCR technology has been modified over the time for different applications. Some of the types of PCR are:

Real-time PCR

Real-time or RT-PCR is also known as quantitative PCR and offers several advantages over the common method of PCR. This method allows to determine the quantity of the amplified DNA by comparing the samples to the standard curve of known quantities, abolishing the requirement of agarose gel electrophoresis to visualize the product. Also, RT-PCR can be

both quantitative and specific. This specificity is achieved with the help of specific nucleic acid probes. This technique uses RNA as the starting material and is first transcribed into cDNA with the help of enzyme reverse transcriptase. This cDNA is then used as the template for qPCR reactions. RT-PCR is used in different applications such as gene expression analysis, microarray validation, pathogen detection, and disease diagnosis.

Droplet digital PCR

Droplet digital PCR technology is a kind of digital PCR which utilize a droplet system containing water-oil emulsion. Small nanodroplets are formed in this emulsion which basically separates the template DNA molecule. The droplets here functions as individual wells or tubes where the PCR reaction takes place. This digital system separates the nano-sized droplets where the PCR amplifications takes place. The one advantage of this kind of PCR is the requirement of a smaller sample as compared to other PCR systems which reduces the cost and sample loss.

Chip PCR

Chip PCR is a technique of chromatin immunoprecipitation combined with quantitative PCR that helps in the determination of protein-DNA interaction at known genomic binding sites. This technique is used in studies that focus on specific genes and regulatory proteins that differ across the experimental conditions. Also, studies like cellular differentiation, tumor suppressor, gene silencing and histone modifications can be done through Chip PCR.

Reverse Transcriptase PCR

Reverse Transcriptase PCR involves the enzyme reverse transcriptase which has the capability to synthesize complementary DNA (cDNA) from mRNA transcripts, and further amplifying the gene of interest through PCR. A major procedure in the RT-PCR is purifying mRNA from the cell extracts. This method is used in variety of applications such as gene expression analysis, RNAi validation, microarray validation, pathogen detection, genetic testing, and disease research.

Nested PCR

Nested PCR is mainly used in situations which requires increased sensitivity or specificity of PCR. This technique involves two sequential amplification reactions involving two different pair of primers. The product of initial amplification reaction is used as the template for the second PCR. This is then primed by oligonucleotides that are placed internal to the first primer. This involvement of two primer oligonucleotides increases the number of cycle as well as sensitivity of the PCR reaction. Nested PCR reduces the non-specific amplification of the target sequence. Nested PCRs are mostly used for microbial detection especially where they are present in low quantities.

Multiplex PCR

Multiplex PCR is majorly used to amplify multiple target sequences in a single PCR experiment. This can be done by using multiple primer pairs in a reaction mixture. The primer design has to be optimized for all the primer pairs to work at the single annealing temperature during PCR. In 2020, Center for Diseases and Control designed RT-PCR multiplex assays which combines multiple gene targets in a single reaction to increase molecular testing accessibility and throughput for SARS-COV-2 diagnostics. Some of the applications of multiplex PCR are pathogen identification, mutation analysis, gene detection analysis, RNA detection, Forensic studies.

Other PCR methods

There are several other PCR methods in use for biological research. Colony PCR screens the presence of a genomic insert from bacteria which negotiates the need of bacterial culture or plasmid purification. Hot Start PCR uses hybrid polymerases which are inactive at ambient temperature. This type of PCR uses 95 ^{o}C of high temperature before adding the polymerase in the reaction mixture. Inverse PCR amplifies the DNA sequence surrounding the target DNA sequence. Methylation specific PCR detects the pattern of DNA methylation and characterize the methylation state. Multiplex PCR involves the use of multiple primers to target various sequences simultaneously. This technique is specifically useful in analyzing genetic mutations and finger printing.

Applications of PCR

Depending on the type and protocol of PCR method, it can be used for different sets of applications.

Disease diagnosis

One of the major breakthrough applications of PCR is in the disease diagnosis and screening. PCR can detect diseases that are inherited, non-inherited as well as infectious diseases. It can be analyzed through direct PCR or in combination with other techniques such as restriction digestion, sequencing, blotting etc.

DNA fingerprinting

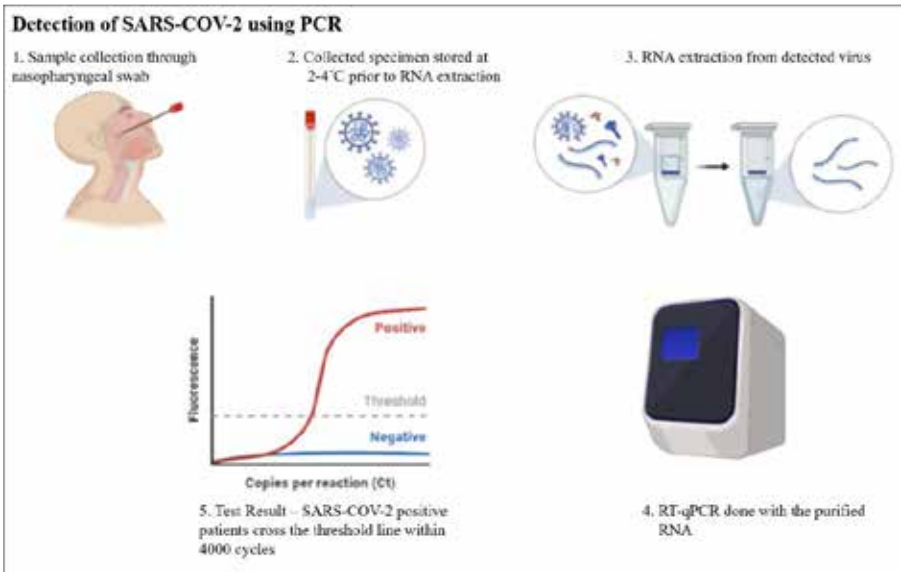
Researchers are able to analyze unique DNA fingerprints between individuals using PCR based genetic markers such as short tandem repeats and variable numbers of tandem repeats. These genetic markers with different repeat numbers and length are unique DNA sequences present in a genome.

Virus detection

Viral DNA can be detected using PCR by using specific primers to the targeted DNA sequence of a virus. The high sensitivity of PCR allows the detection of virus at a very early stage of infection. The amount of virus can be quantified by PCR methods such as the RT-PCR.

Mentioning about the ongoing pandemic of COVID-19 across the globe PCR has been a kind of COVID warrior in the field. The COVID-19 RT-PCR test is widely being used for the quantitative detection of COVID-19 virus. It enables the specific detection of nucleic acid from corona virus originating in the upper and lower respiratory samples collected from suspected individuals. This test uses three sets of primers and probes detecting the nucleocapsid (N) gene and the human RNase P in the human sample.

Cancer



PCR can be used to analyze different chromosomal changes or mutations within the genome to detect certain diseases such as cancer. Molecular markers are targeted to accurately analyze and detect the oncogenic conditions. PCR such as the RT-PCR can identify small mutations, gene duplications, or deletions.

PCR is turning into a powerful technique in today's science field as it is fairly simple to use and gives the result rapidly. The technique is highly sensitive with the potential to generate millions of copies of a nucleic acid sequence using just a sample template strand. As its application stands in a variety of disease diagnosis and detection PCR will continue to have a prominent place in clinical laboratories for years to come.

Authors
Sibi Raj, Dhruv Kumar

Amity Institute of Molecular Medicine & Stem Cell Research (AIMMSCR), Amity University Uttar Pradesh, Sec-125, Noida-201313, India

Email: dhruvbhu@gmail.com, dkumar13@amity.edu

Sanger Sequencing- The Gold Standard

History:

Pioneering X-ray crystallography work done by Rosalind Franklin, Watson and Crick led to the double helix structure of the DNA. However, DNA sequencing was not discovered until 1974. Two teams, one American and other English independently developed methods to sequence the DNA. The English team led by Frederick Sanger used "chain termination method" for DNA sequencing of bacteriophage ϕX174, while the American team led by Maxam and Gilbert used a "chemical cleavage protocol". (1,2) Both the teams shared the 1980 Nobel Prize in Chemistry. However, due to practicality of Sanger's method, it was widely accepted. Subsequently, two laboratories the European Molecular Biology Laboratory (EMBL) and Pharmacia-Amersham introduced first automated DNA sequencing which was commercialized by Applied Biosystems (ABI) in 1986. In 1990, for the first time, the *HGPRT1* gene, that encodes for Hypoxanthine-guanine phosphoribosyl transferase enzyme was sequenced using bidirectional sequencing. (3) In 1996, ABI Prism 310 was introduced which was an automated DNA sequencer that utilized slab gel electrophoresis. The latter was a labour intensive technique and was subsequently replaced by automated reloading of capillaries with polymer matrix used by ABI Prism 3700 with 96 capillaries. The automated DNA sequencing was used by the Human Genome Project over a period of 13 years to decipher the entire euchromatic human genome. A draft of the genome comprising of ~ 3 billion bases was announced in 2000 and a more complete draft was published in 2003. The approximate cost of sequencing of the human genome was $2.7 billion. Taking into account the cost and throughput of automated sequencing further advancement lead to modern DNA sequencing with high throughput next generation sequencing.

Principle

Sequencing is the process by which the order of nucleotides in a chosen portion of DNA is deciphered. The chain-termination technique makes use of chemical analogues of the deoxyribonucleotides (dNTPs) that are the monomers of DNA strands. Dideoxynucleotides (ddNTP) are fundamentally different from deoxynucleotides (dNTP) with respect to a hydroxyl group at 3' carbon of the sugar moiety (Figure 1). The phosphodiester bond formation occurs between 5' carbon atom of one sugar molecule and 3' carbon atom of another. For the formation of phosphodiester bond, 3' carbon of sugar should have hydroxyl group. Hence the incorporation of ddNTP in a reaction leads to chain termination (Figure 2). The quantity of dideoxynucleotides is limited so as ddNTP may cause chain termination only once in a while. In each tube, the primer anneals to the template strand and during extension, either dNTP or ddNTP gets randomly incorporated. When a dNTP gets incorporated, further elongation is possible, however, in case a ddNTP gets incorporated, chain termination occurs. This results in the formation of products of different lengths of DNA fragments due to termination of elongation at places where ddATP gets incorporated in place of dATP in ddATP tube (Figure 3a). Once the reaction is complete, the contents of each tube are electrophoresed by application

of an electrical field in separate lanes of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The function of sodium dodecyl sulphate is to neutralize the charge so that mobility during electrophoresis is dependent only on the molecular weight of the fragments. Depending on the method used to label the ddNTP, autoradiography or UV light is used to decipher the sequence (Figure 3b).

Several improvements were made to Sanger sequencing in the following years, which primarily involved the replacement of phospho- or tritrium-radiolabelling with fluorescent dye labelled bases and improved resolution of detection through capillary electrophoresis. This technology was first commercialized by Applied Biosystems in 1986. Capillary gel electrophoresis, which used soluble polymers, had subsequently replaced the slab gel electrophoresis. (4) This enabled application of high voltage (in kilovolts) and led to a considerable increase in sample throughput. Each ddNTP is fluorescently labelled a unique dye of different fluorescent emission wavelengths in a single reaction. The cycle sequencing product is then processed by capillary electrophoresis. A high voltage electrical field is applied over the capillary so that cycle sequencing products move from negative charged electrode to positively charged electrode. The gel polymer provides a frictional resistance. Hence, the DNA strand with shortest length faces least resistance and reaches the far capillary length fastest. Before the DNA strand reaches the positively charged end, they pass through a window through which a laser beam enters the capillary. The fluorescence emitted is detected by an optical system. The emitted fluorescence is converted in digital data and is represented in the form of a chromatogram. Schematic representation of automated capillary electrophoresis is shown in Figure 4.

Figure 1. The difference between deoxynucleotide (dNTP) and dideoxynucleotide (ddNTP) is shown in the figure with respect to hydroxyl and hydrogen base at 3' carbon position of sugar moiety.

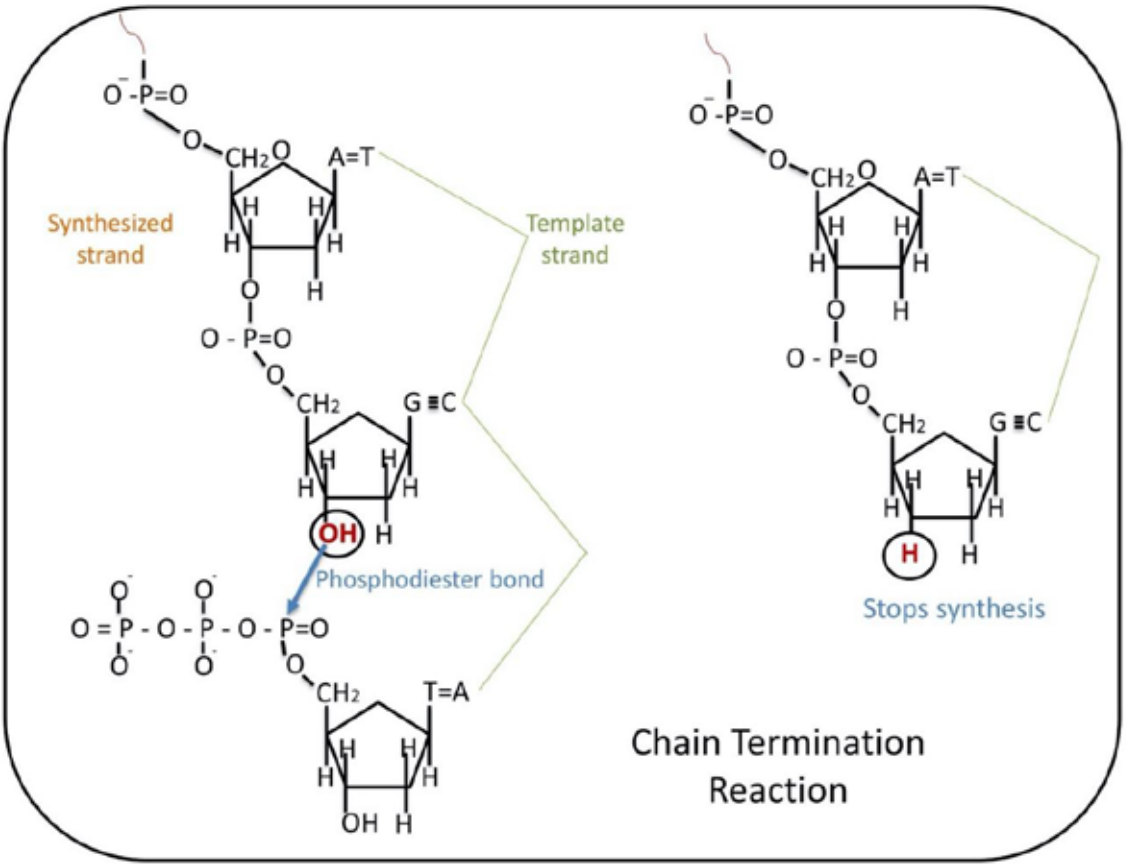


Figure 2. Principle of chain termination reaction is that the incorporation of ddNTP leads to chain termination.

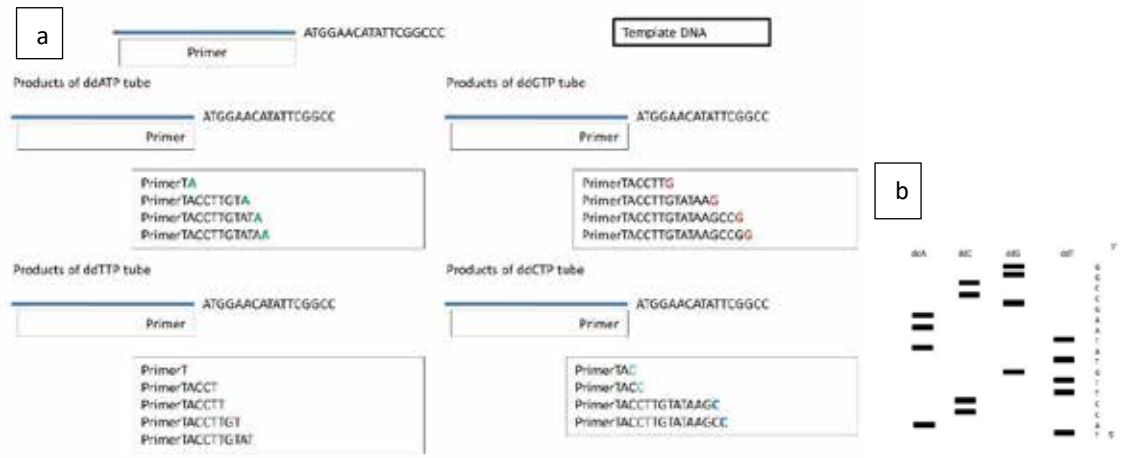


Figure 3a. Schematic representation of chain termination on incorporation of ddNTP. Figure 3b. Schematic representation of gel electrophoresis for deciphering the sequence.

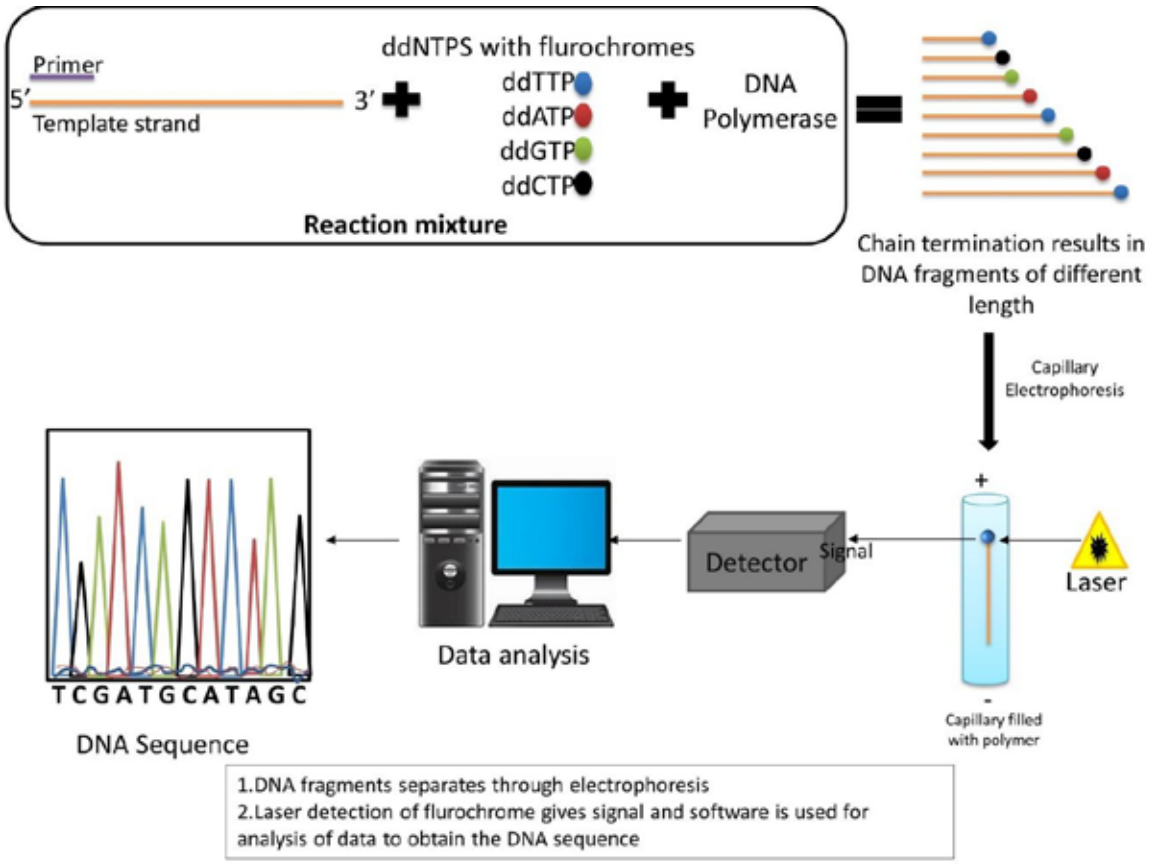


Figure 4. Schematic representation of automated DNA sequencing.

Methodology

The workflow for Sanger sequencing is given in Table 1.

1. **Nucleic acid extraction:** It is important to use good quality nucleic acid that is pure and free of contaminants. Choice of DNA versus RNA depends on what is the objective of the experiment. DNA is more widely used for Sanger sequencing. Various methods of DNA/RNA extraction are available. Methods may vary according to individual laboratory preferences. RNA extraction should follow cDNA synthesis. The yield of DNA and RNA can be checked using various methods including absorbance (optical density), agarose gel electrophoresis, or use of fluorescent DNA-binding dyes. The most common technique to determine DNA yield and purity is measurement of absorbance. Optical density allows measuring both quantity and quality of nucleic acid to ensure it is intact and free of contaminants before proceeding to PCR.
2. **Polymerase chain reaction:** Amplification of the target region is important before sequencing it. Polymerase chain reaction (PCR) uses multiple cycles of three steps (denaturation, annealing and extension) to produce billions of copies of template. For amplifying the region of interest, appropriate primer designing is the key. Primer designing can be done using online tools such as Primer3. During primer designing, it is important to keep in mind that the first 30-40 nucleotides and nucleotides after 800 bases show background noise and hence either the region of interest should not fall in

such area or primers should be designed with overlapping areas. A good PCR for Sanger sequencing produces a single product of correct size.

3. **Purification of PCR products:** The unused reagents including unused primers, unincorporated dNTPs, salts, proteins etc during PCR need to be cleaned up as they may interfere with sequencing. Various methods available for purification of PCR products are gel purification, enzymatic purification, bead cleanup, ethanol purification and column purification. Ethanol purification is cost-effective, however, inexperienced users are at high risk of losing template. Commercial kits are using various principles and individual laboratories have their preferences.
4. **Cycle sequencing:** Many cycle sequencing kits are commercially available. The most frequently used kit is BigDye Terminator (BDT) version 3.1 from Life Technologies. Good template quality yields good results. However, too much template makes data appear top heavy, with strong peaks at the beginning of the run that fades rapidly and too little template or primer reduces the signal strength and peak height. The trick for getting good results is to add the right amount of template. The table 2 below shows the recommended amount of template for BGT version 3.1. In addition to the template, sequencing primer, BigDye Sequencing Buffer and reaction premix are mixed and cycle sequencing is done in conditions listed in Table 3. The template to be taken and conditions to be followed varies according to the kit being used and the details provided here are an example of BDT 3.1.
5. **Purification of cycle sequencing product:** This purification step is essential to remove both unlabeled and dye-labelled components that can interfere with the electro-kinetic injection, electrophoresis and data analysis. The purification can be done using various methods which include bead-cleanup, BigDye Xterminator Purification (Life Technologies), spin-column purification (e.g. Qiagen) and ethanol precipitation.
6. **Capillary electrophoresis:** The purified cycle sequencing product is loaded in a 96 well plate for automated capillary electrophoresis and sequencing.

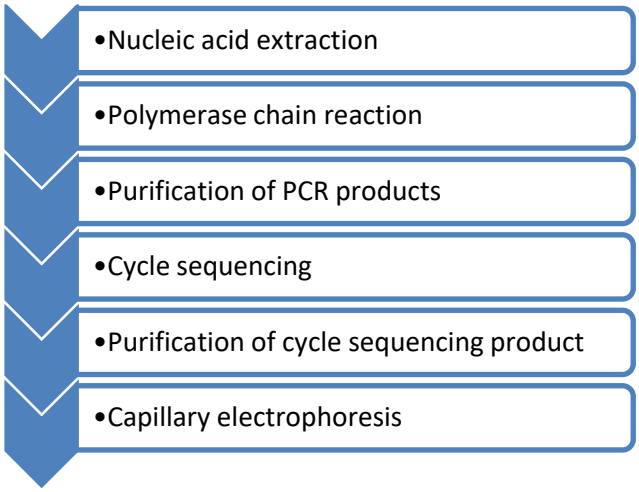


Table 1. Workflow for Sanger sequencing.

Template (PCR product)	Quantity
100-200bp	1-3ng
200-500bp	3-10ng
500-1000bp	5-20ng
Plasmid	150-300ng

Table 2. Recommended amount of template for cycle sequencing.

Stage	Description	Temperature	Time
1	Denaturation	96°	1 min
2	Amplification (25 cycles)	96°	10 sec
		50°	5 sec
		60°	4 min
3	Hold	4°	Indefinite

Table 3. Conditions for cycle sequencing for BDT 3.1.

Data Interpretation

Automated DNA sequencing generates output in the form of Analyzed ABI files (.ab1), text file of the sequence in ABI or FASTA files (.seq; all bases or only clear range bases), phred (.phd.1) files, standard chromatogram format (.scf) files, analysis reports (TXT, HTML, PDF, or XML files). Frequently, .ab1 file is uploaded in software provided by the automated sequencer or other paid (Geneious, Sequencher, Mutation Surveyor etc) or freely available software (FinchTV). The sequencing data appears as 4 colours chromatogram according to the nucleotide sequence (A-green, G-black, T- red and C- blue) (note that this is a user defined feature). In addition to the 4 peaks, the peaks which are not clear enough to be designated to either of the four peaks are designated as N. The peaks in the chromatogram should be evenly spaced. The peak height may vary three folds. To interpret the data, the sequence obtained is aligned to the reference sequence. A homozygous variant can be easily identified by the change in nucleotide. Heterozygous variant shows two peaks of two different colours at the same position. Though there are automated programs that point to variants, there are chances of error especially when the variant is heterozygous and the chromatogram shows the presence of both reference and alternate nucleotide peak. In addition, it is important to consider background noise before interpreting a heterozygous variant. Hence, manual checking of the sequence becomes essential for correct interpretation in Sanger sequencing data. The initial 15-40 nucleotides of the sequence are poor quality owing to site of primer binding and the sequence traces after 800 nucleotides are also of poor quality and hence interpretation at these areas should be done with caution. Homozygous deletion is easy to interpret as a particular length of the sequence is missing as compared to the reference sequence. Figure 6a shows a clean sequence with a heterozygous change in nucleotide. However, heterozygous deletion may lead to heterozygous peaks at a stretch of

sequence. Annotation of the variants should be done following The Human Genome Variation Society (HGVS) guidelines.

Trouble shooting

The problem in Sanger sequencing can be split into two parts: input template or sequencing reaction. To resolve the difference between the two it is important to run controls with each run. In each run, pGEM control DNA provided with the kit is run. Obtaining a proper sequence of pGEM control but not the samples indicates a problem with the template. Soft pointers where the problem may have occurred need observation of the trend. When all the samples in a particular run have failed, the current, voltage, temperature, and power throughout the electrophoresis run should be reviewed to determine whether an electrical problem occurred during the run. A few common artefacts include very low signals, dye blobs (Figure 6b) and background noise (Figure 6c). Low signals may occur due to inadequate amount of template or primers used. Background noise may occur due to carry over from contaminated septa, water or buffer. Dye blobs may occur due to insufficient cleanup and signal from unincorporated ddNTPs.

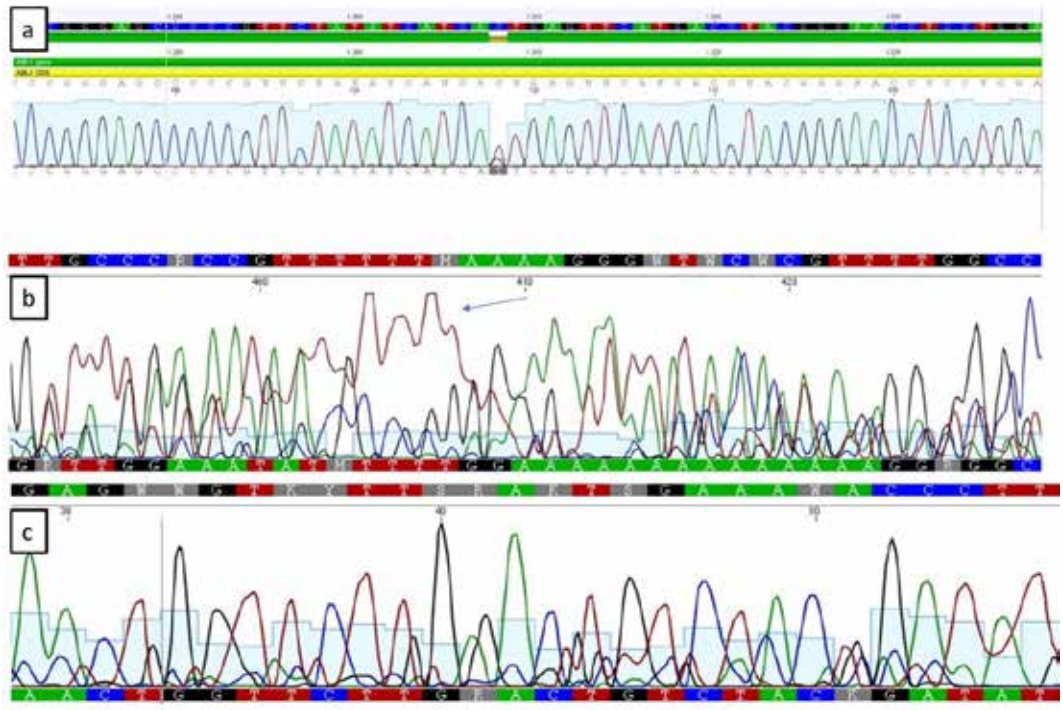


Figure 6a. Chromatogram showing a clean chromatogram with heterozygous change in nucleotide; Figure 6b Chromatogram showing dye blobs (arrow); Figure 6c Chromatogram showing background noise.

Indications of Sanger sequencing in the era of NGS

Sanger sequencing is a technique that can be used economically when a single small area of the gene needs to be screened especially when the expected mutation burden is high. For instance, taking an example of cystic fibrosis, in more than 90% patients, the mutations responsible for this disease are located in the coding region of CFTR gene (4,500 basepairs) which can be covered in a few Sanger sequencing reactions. In such scenarios, screening the suspected cystic fibrosis patients with Sanger sequencing makes sense in the terms of cost and ease of the procedure. However, for scenarios like acute myeloid leukemia where multiple genes need to be sequenced and the mutation burden may be at low levels, next generation sequencing (NGS) becomes the preferred modality in terms of overall cost as well as sensitivity of detection of mutations. However, variants detected by NGS are validated using Sanger sequencing and it still remains the gold standard.

Limitations

Sanger sequencing has limited throughput and can sequence upto 1000 nucleotides in a single sequence. Multiple reactions will be needed to sequence a gene which is large or multiple genes. This ultimately increases the cost of sequencing especially for large genes. All the more, the capital cost for Sanger sequencing is also high. Another disadvantage is its sensitivity as low level of variant (less than 20%) may be missed by Sanger sequencing.

References

1. Maxam AM, Gilbert W. A new method for sequencing DNA. Proc Natl Acad Sci U S A [Internet]. 1977 [cited 2020 Jul 5];74(2):560–4. Available from: https://pubmed.ncbi.nlm.nih.gov/265521/

2. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A [Internet]. 1977 [cited 2020 Jul 5];74(12):5463–7. Available from: https://pubmed.ncbi.nlm.nih.gov/271968/

3. Edwards A, Voss H, Rice P, Civitello A, Stegemann J, Schwager C, et al. Automated DNA sequencing of the human HPRT locus. Genomics [Internet]. 1990 [cited 2020 Jul 5];6(4):593–608. Available from: https://pubmed.ncbi.nlm.nih.gov/2341149/

4. Jorgenson JW, Lukacs KD. Free-zone electrophoresis in glass capillaries. Clin Chem. 1981 Sep;27(9):1551–3.

Authors:

1. **Dr Sweta Rajpal, MD, DNB (Pathology), DM Hematopathology, Assistant Professor, Department of Hematopathology, Tata Memorial Centre, Mumbai, India**

2. **Dr Anurodh Gupta, MD Fellow Molecular Hematopathology Tata Memorial Centre, Mumbai, India**

3. **Dr Nikhil Patkar (MD DNB, Pathology), Clinician Scientist & Associate Professor Wellcome-DBT Intermediate Fellow Hematopathology Laboratory Tata Memorial Centre, Mumbai, India**

4. **Dr P G Subramanian Professor and Officer Incharge Hematopathology Laboratory Tata Memorial Centre, Mumbai, India**

Game Changer- Next Generation Sequencing Technology

1. What is Next-Generation DNA Sequencing?

DNA sequencing is the process of determining the sequence of nucleotides in a section of DNA. The first commercialized method of DNA sequencing was Sanger sequencing. The massively parallel sequencing technology known as next-generation sequencing (NGS) has revolutionized the biological sciences. With its ultra-high throughput, scalability, and speed, NGS enables researchers to perform a wide variety of applications and study biological systems at a level never before possible. Today's complex genomic research questions demand a depth of information beyond the capacity of traditional DNA sequencing technologies. NGS has filled that gap and become an everyday research tool to address these questions. NGS, also known as high-throughput sequencing, is the catch-all term used to describe a number of different modern sequencing technologies. These technologies allow for sequencing of DNA and RNA much more quickly and cheaply than the previously used Sanger sequencing, and as such revolutionized the study of genomics and molecular biology (Van Allen et al, 2014). Such technologies include:

1.1. **Illumina (Solexa) sequencing**

Illumina sequencing works by simultaneously identifying DNA bases, as each base emits a unique fluorescent signal, and adding them to a nucleic acid chain. In NGS, vast numbers of short reads are sequenced in a single stroke. To do this, firstly the input sample must be cleaved into short sections. The length of these sections will depend on the particular sequencing machinery used. In Illumina sequencing, 100-150bp reads are used. Somewhat longer fragments are ligated to generic adaptors and annealed to a slide using the adaptors. PCR is carried out to amplify each read, creating a spot with many copies of the same read (Hedegaard et al, 2014). They are then separated into single strands to be sequenced. There are actually four steps engaged in this sequencing which is depicted in **figure 1.**

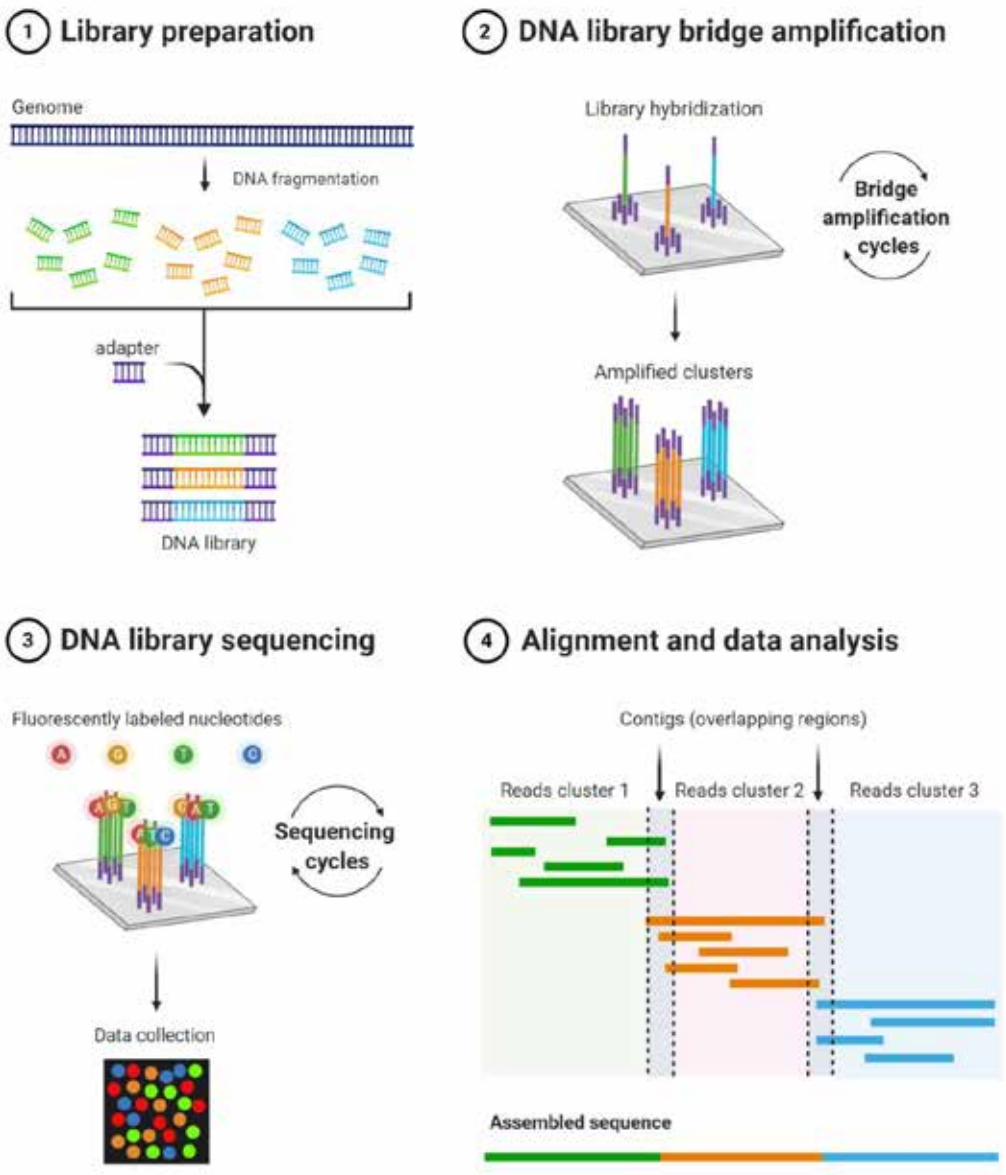


Figure 1. Steps involved in Illumina (Solexa) sequencing

1.2. **Roche 454 sequencing**

This method is based on pyrosequencing, a technique which detects pyrophosphate release, again using fluorescence, after nucleotides are incorporated by polymerase to a new strand of DNA.

1.3. **Ion Torrent: Proton / PGM sequencing**

Ion Torrent sequencing measures the direct release of H+ (protons) from the incorporation of individual bases by DNA polymerase and therefore differs from the previous two methods as it does not measure light.

2. **What NGS Can Do For Us?**

NGS technology has fundamentally changed the kinds of questions scientists can ask and answer. Innovative sample preparation and data analysis options enable a broad range of applications (Li and Freudenberg 2014). For example, NGS allows researchers to:

1. Rapidly sequence whole genomes
2. Zoom in to deeply sequence target regions
3. Utilize RNA sequencing (RNA-Seq) to discover novel RNA variants and splice sites, or quantify mRNAs for gene expression analysis
4. Analyze epigenetic factors such as genome-wide DNA methylation and DNA-protein interactions
5. Sequence cancer samples to study rare somatic variants, tumor sub-clones, and more
6. Study the human microbiome and discover novel pathogens

3. **The scope of NGS technology**

NGS can be performed at different levels. It can be used for whole-genome sequencing. At this level, almost all the nucleotide in the genome, including chromosomal DNA and the mitochondrial DNA, are interrogated. Whole genome sequencing is used more often in research and less common in clinical settings. When used in clinical settings, it is used more often for constitutional genetic diseases, rather than for cancer somatic mutations. It is especially useful for the diagnosis of some rare genetic diseases. For example, when a genetic disease is suspected but no specific mutation has been identified by other molecular tests. In such cases, whole genome sequencing may provide additional information in terms of disease-associated mutations. Whole genome sequencing is used less frequently for cancer somatic mutation because the average depth in whole genome sequencing is limited. In a certain tumor, allelic mutation frequencies may vary and the percentages of tumor cells in different specimens may also vary. Detection of different mutations with different allelic frequencies in such settings often need deep sequencing and that is very challenging for the whole genome sequencing method (Garraway and Lander 2013).

NGS assay can be used for whole exome sequencing. The entire coding region of all axons of an organism including any cell types can be sequenced. In human, that is about 1% of the human genome and is more often used in research.

NGS can also be performed at transcriptome level which includes entire assembly of RNA transcripts in a given cell type including mRNA, rRNA, tRNA, micro-RNA, and non-coding

RNA. Unlike DNA sequencing, this is called RNA sequencing. Specially designed mRNA sequencing is also often used to detect fusion genes (Van Allen et al, 2013).
The most commonly used NGS assay for cancer patients is targeted panel sequencing which usually interrogates dozens or hundreds of targeted genes. Such targeted NGS assays are usually designed for a disease or a category of diseases, for example, a panel designed for myeloid leukemia or a panel designed for carcinoma. Compared to whole genome sequencing, such targeted panel has only limited targets. Therefore, it allows a lot more depth in sequencing, which is necessary to cover different mutations with different allelic mutation frequencies.

4. Next-Generation Sequencing (NGS): stimulating the Next Generation of Cancer Diagnostics and Treatment

The advent of NGS in the past decade has made a tremendous impact on basic and clinical research. This high-throughput DNA sequencing technology can sequence an entire human genome within a few hours at a cost of just around one thousand US dollars (USD). Only 18 years ago, this feat took the International Human Genome Sequencing Consortium 13 years and three billion USD to accomplish using Sanger sequencing, demonstrating the sheer magnitude in technological development in the last two decades. Advances in NGS have catalyzed progress in countless areas of cancer research, heralding a "molecular age" of cancer. Traditionally, cancer diagnosis rested heavily on histological classification of the tumor. For instance, brain cancers were historically classified by the putative cell of origin, such as astrocytomas or oligodendrogliomas. However, the most recently revised 2016 World Health Organization (WHO) brain tumor guidelines strongly emphasized integration of molecular characteristics into diagnosis, underscoring the mounting importance of tumors' molecular and genetic properties on clinical practice (Smits et al, 2014).

5. Two facets of NGS in clinical cancer research: diagnostics and treatment

NGS analysis of tumor genomics, transcriptomics, and epigenomics is driving biomarker discovery for cancer diagnostics and tumor stratification. Clinical biomarkers, as defined by the Food and Drug Administration (FDA), can serve diagnostic, prognostic, predictive, and pharmacodynamics purposes in cancer and other diseases. Diagnostic biomarkers are characteristic of a disease and can be used to determine whether an individual has cancer. Prognostic biomarkers forecast the cancer's natural history without treatment (i.e., aggressive versus less aggressive phenotype by tumor stratification), whereas predictive biomarkers anticipate the patient's response to therapy. Finally, and importantly for the evaluation of clinical trial success, pharmacodynamics biomarkers indicate a treatment effect. In addition to biomarker discovery, NGS is ushering in an era of precision oncology, a paradigm shift in cancer management that aims to match a tumor's molecular characteristics with targeted drugs to improve patients' prospects (Forment et al, 2012). This approach has stimulated a

series of new oncological clinical trial designs that employ NGS to identity genetic vulnerabilities in patients' tumors, which inform treatment options.

6. Next-generation sequencing for SARS-CoV-2 research

During viral pandemics, NGS can be a valuable tool for the characterization and detection of viruses in the environment, animals, and humans (Kustin et al. 2019). Additionally, NGS can provide useful insights to help researchers and clinicians looking to develop the right treatment options. We offer a variety of RNA-sequencing kits for unmatched sensitivity and reproducibility for all of your demanding sequencing applications, including research on SARS-CoV-2.

7. Conclusion

NGS is inextricably intertwined with the realization of precision medicine in oncology. While it is unlikely to obviate traditional pathologic diagnosis in its current state, it allows a more complete picture of cancer etiology than can be seen with any other modality. However, precision cancer medicine and large-scale NGS testing will require novel approaches towards ensuring evidence-based medicine. Treating each genetic abnormality as an independent variable when hundreds or thousands are queried in every patient will require new trial designs and statistical methods to ensure the utility of these approaches. Broadly, clinicians and translational researchers will need to continue to engage in direct dialog, both within and across institutions, to advance the integration of genomic information and clinical phenotypes, and enable precision cancer medicine through NGS approaches.

8. References

Forment JV, Kaidi A, Jackson SP. Chromothripsis and cancer: causes and consequences of chromosome shattering. Nat Rev Cancer. 2012;12:663–70.

Garraway LA, Lander ES. Lessons from the cancer genome. Cell. 2013;153:17–37.

Hedegaard J, Thorsen K, Lund MK, Hein A-M K, Hamilton-Dutoit SJ, Vang S, et al. Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue. PLoS One. 2014;9, e98187.

Kustin, T. *et al.* A method to identify respiratory virus infections in clinical samples using next-generation sequencing. *Sci. Rep.* **9,** 1–8 (2019).

Li W, Freudenberg J. Characterizing regions in the human genome unmappable by next-generation-sequencing at the read length of 1000 bases. Comput Biol Chem. 2014;53:108–17.

Smits AJJ, Kummer JA, de Bruin PC, Bol M, van den Tweel JG, Seldenrijk KA, et al. The estimation of tumor cell percentage for molecular testing by pathologists is not accurate. Mod Pathol. 2014;27:168–74.

Van Allen EM, Wagle N, Levy MA. Clinical analysis and interpretation of cancer genome data. J Clin Oncol. 2013;31:1825–33.

Van Allen EM, Wagle N, Stojanov P, Perrin DL, Cibulskis K, Marlow S, et al. Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. Nat Med. 2014;20:682–8.

Author(s)

Dr. Niraj Kumar Jha

Assistant Professor
Department of Biotechnology
School of Engineering & T echnology (SET)
Sharda University, Plot No.32-34, Knowledge Park III,
Greater Noida, Uttar Pradesh-201310, India
Email: nirajkumarjha2011@gmail.com;
niraj.jha@sharda.ac.in

Dr. Saurabh Kumar Jha

Assistant Professor
Department of Biotechnology
School of Engineering & Technology (SET)
Sharda University, Plot No.32-34, Knowledge Park III, Greater Noida, Uttar Pradesh-201310, India
Email: saurabh.jha@sharda.ac.in

About MPAI

Molecular Pathology refers to the application of genetic knowledge to diagnose, monitor or prognosticate disease, or to guide its treatment. Of late, Molecular Pathology is being universally acknowledged as the modern face of Pathology and more frequently it is being defined as the Pathology of the future. However, the awareness about molecular diagnostics and its implication in clinical practice need to be strengthened in India.

In 2011, a group of experienced, enthusiastic experts in the field of molecular pathology & diagnostics in India, with the objective of cultivation and promotion of the study and practice of molecular pathology, came together to form a 'Molecular Pathology Association of India (MPAI). This association will also work as a platform for exchanging ideas, debating application of molecular assays amongst individuals from the field of biomedical sciences with special interest in practice and application of molecular pathology. The MPAI will also endeavour to create a strong bridge between the basic scientists working in various research institutes and the clinicians working in hospitals for achieving the best possible management of patients. The Association remains conscious of the responsibility it shoulders and its activities will not be restricted only for its members, societies and professionals. It will always try to reach out various other biomedical societies to contribute to overall growth and excellence in healthcare management of the country.

MPAI Objectives

The main objective of the Association is to further scientific and clinical interest by encouraging the study and practice of Molecular Pathology in India. To achieve this objective the Association shall:

- Establish standards for education and training in the field of Molecular Pathology.
- Develop and maintain Molecular Pathology Practice guidelines for Clinicians and Labo ratorians through collaboration with other professional organizations, review of evidences and consensus of experts.
- Facilitate translational research pertaining to development, adoption and appropriate use of advanced molecular assays for successful health management of various diseases.
- Encourage high standards of quality excellence by creating sound quality assurance and control policies, nodal centers for proficiency testing and validations of molecular assays.
- Foster research and innovations by providing a forum for young talents for the exchange of ideas and information on current practices and advances in the field of Molecular Pathology.
- Continually enhance scientific knowledge of Molecular Pathology by organizing sympo sia, meetings, discussion forum, reports and publications.