AL-FARABI KAZAKH NATIONAL UNIVERSITY FACULTY OF BIOLOGY
AND BIOTECHNOLOGY DEPARTMENT OF BIOTECHNOLOGY

Lecture 9

METHODS OF MOLECULAR BIOLOGY
CORE TECHNIQUES

PhD, Acting
Associate Professor,
Doktyrbay Gulina

Lecture Aim

To provide students with fundamental knowledge of molecular biology techniques and their applications in biological research, focusing on the principles, procedures, and practical significance of core laboratory methods used to study DNA, RNA, and proteins.

Expected Learning Outcomes

Upon completion of this lecture, students will be able to:

1. Describe the basic principles of molecular biology and its major laboratory techniques.
2. Explain the procedures and applications of DNA and RNA extraction.
3. Understand the principles of Polymerase Chain Reaction (PCR) and gel electrophoresis.
4. Identify the applications of DNA sequencing and molecular analysis methods.
5. Compare different molecular biology techniques and evaluate their advantages and limitations.
6. Select appropriate molecular methods for solving basic biological research problems.

Agenda & Structure

Foundations & PCR

Slides 3–11: Central Dogma, nucleic acid isolation, gel electrophoresis, PCR principles, thermal cycling, qPCR, RT-PCR, and specialized PCR techniques.

DNA Sequencing Technologies

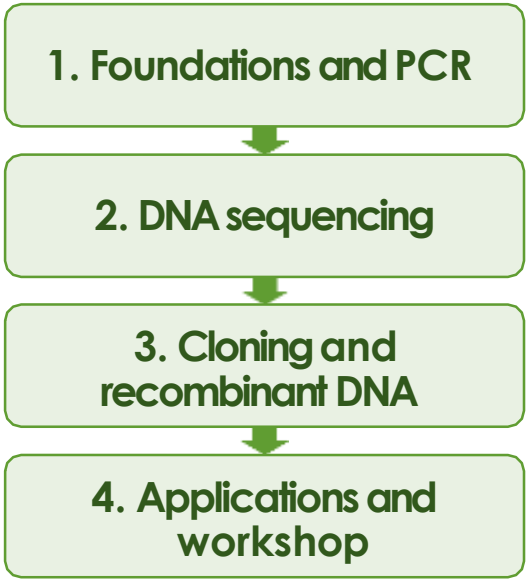
Slides 12–17: Sanger sequencing, next-generation sequencing (NGS), long-read platforms, comparative analysis, and bioinformatics workflows.

Cloning & Recombinant DNA

Slides 18–23: Restriction enzymes, vector architecture, DNA ligation, transformation strategies, and recombinant protein expression.

Applications & Workshop

Slides 24–30: Gene editing, medical diagnostics, synthetic biology, forensics, single-cell genomics, practical workshop, and references.



Central Dogma & Molecular Tools

Central Dogma

DNA → RNA → Protein. Genetic information flows from DNA via transcription to mRNA, then translation produces functional proteins. This foundational principle underpins all molecular biology techniques.

DNA Polymerases

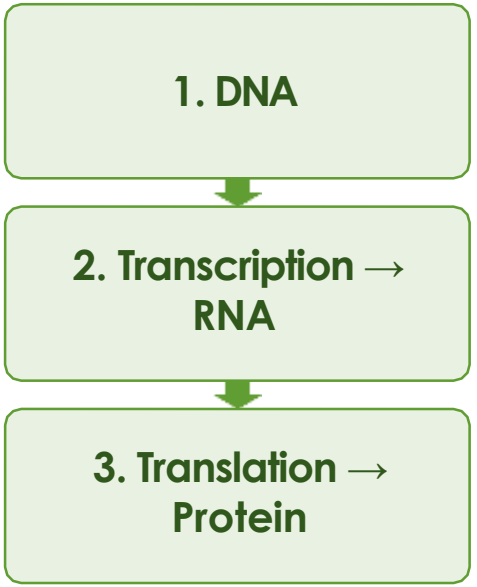
Enzymes that synthesize new DNA strands using a template. Essential for PCR amplification, DNA replication, and repair. Taq polymerase is the cornerstone of in vitro amplification methods.

Restriction Enzymes

Endonucleases that recognize specific palindromic DNA sequences and cleave them. They generate blunt or sticky ends, enabling precise DNA cutting for cloning and recombinant DNA construction.

DNA Ligases

Enzymes that catalyze phosphodiester bond formation between DNA fragments. T4 DNA ligase joins compatible ends in cloning workflows, enabling assembly of recombinant DNA constructs.



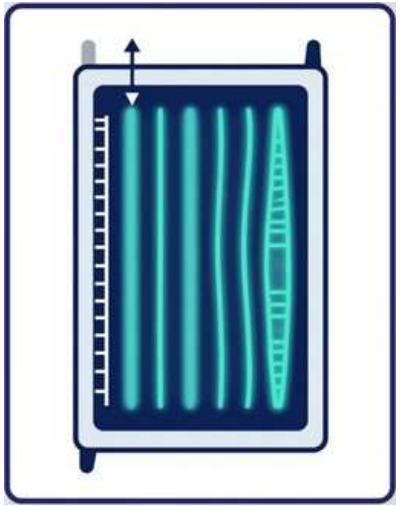
Nucleic Acid Isolation

1. **Cell Lysis**
Mechanical, chemical, or enzymatic disruption of the cell membrane releases genomic DNA or total RNA. Detergents such as SDS solubilise lipid bilayers while proteinase K digests associated proteins.
2. **Binding & Separation**
Silica-column chromatography selectively adsorbs nucleic acids under chaotropic salt conditions. Phase-separation methods (e.g., phenol-chloroform) partition DNA or RNA into the aqueous phase away from proteins and lipids.
3. **Washing & Inhibitor Removal**
Sequential wash buffers remove residual salts, humic acids, and PCR inhibitors. Thorough washing is critical to ensure high-quality nucleic acids suitable for sensitive downstream assays such as qPCR and NGS.
4. **Elution & Quantification**
Purified DNA or RNA is eluted using low-salt buffer or nuclease-free water.
5. **Quality Assessment**
Concentration and purity are measured by spectrophotometry (A260/A280 ratio), while integrity is assessed by agarose gel electrophoresis.



Gel Electrophoresis

1. Separation PrincipleAgarose gel electrophoresis is a laboratory technique used to separate DNA and RNA fragments according to their size. Smaller fragments migrate faster through the gel matrix, while larger fragments move more slowly.
2. DNA Charge and MigrationDNA molecules carry a negative charge due to their phosphate backbone. When an electric field is applied, DNA fragments migrate toward the positive electrode (anode). This allows separation based on fragment length.
3. DNA Ladder and Fragment AnalysisA DNA ladder containing fragments of known sizes is used as a molecular size marker. By comparing sample bands with the DNA ladder, researchers can estimate the size of unknown DNA fragments.
4. VisualizationAfter electrophoresis, nucleic acid bands are visualized using appropriate fluorescent dyes under UV or blue-light illumination.



PCR Principles

History & Discovery

PCR was invented by Kary Mullis in 1983, revolutionizing molecular biology. The technique enables in vitro amplification of specific DNA sequences and earned Mullis the Nobel Prize in Chemistry in 1993.

Reaction Components

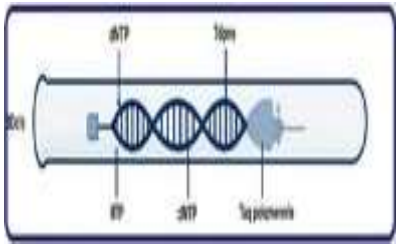
A standard PCR reaction requires template DNA, sequence-specific forward and reverse primers, deoxynucleotide triphosphates (dNTPs), thermostable Taq polymerase, and an optimized reaction buffer with $MgCl_2$.

Taq Polymerase

Taq polymerase, isolated from Thermus aquaticus, is heat-stable and active at ~72°C. It synthesizes new DNA strands by extending primers along the template, enabling repeated thermal cycling without enzyme degradation.

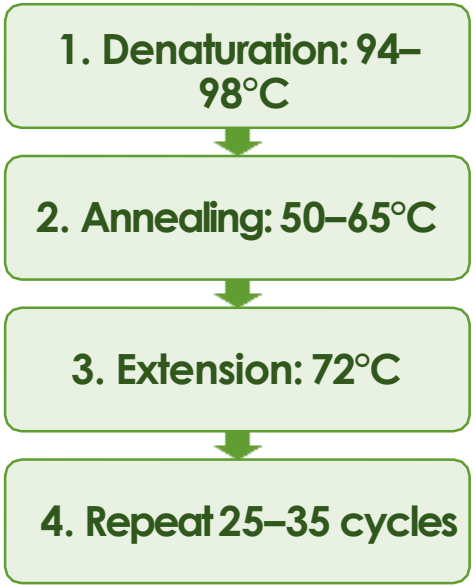
Exponential Amplification

Each PCR cycle doubles the number of target DNA copies. After 30–40 cycles, a single template molecule yields over one billion copies, enabling detection and analysis of trace DNA quantities in research and diagnostics.



The Thermal Cycling Mechanism

Denaturation (~94-98°C)
The double-stranded DNA template is heated to 94-98°C, breaking hydrogen bonds between complementary bases and separating the two strands into single-stranded templates.
Annealing (~50-65°C)
The reaction is cooled to 50-65°C, allowing short oligonucleotide primers to bind specifically to their complementary sequences on each single-stranded DNA template.
Extension (~72°C)
Taq polymerase synthesizes new DNA strands at 72°C, extending from each primer in the 5' to 3' direction and producing two copies of the target sequence.
Cycle Amplification (30-40 Cycles)
Each cycle doubles the number of target DNA copies. After 30-40 cycles, a single template molecule yields over one billion copies, enabling detection and downstream analysis.



Quantitative Real-Time PCR

Fluorescence Detection

SYBR Green intercalates into double-stranded DNA, emitting fluorescence proportional to amplicon quantity. TaqMan probes offer sequence-specific detection via 5' exonuclease cleavage, increasing signal specificity in complex samples.

Ct (Threshold Cycle)

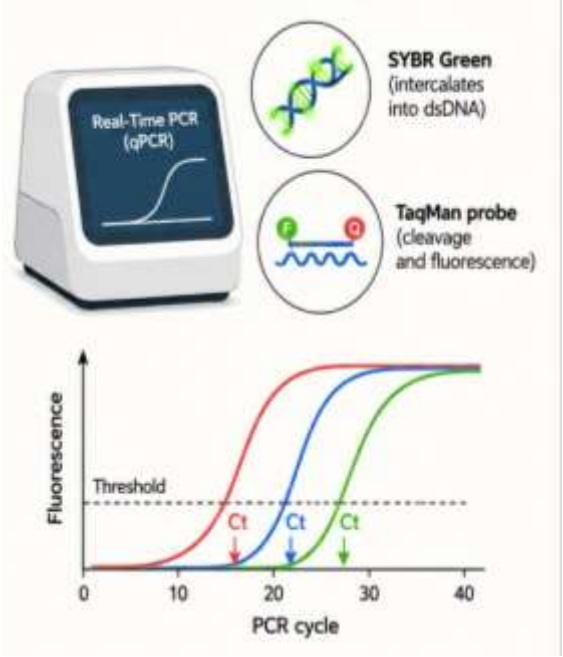
The Ct value is the cycle at which fluorescence exceeds background threshold. Lower Ct indicates higher initial template quantity. Relative quantification uses the ΔΔCt method against a reference gene for normalisation.

Gene Expression Analysis

qPCR quantifies mRNA levels by combining RT-PCR with real-time detection. Enables comparison of gene expression across conditions, tissues, or time points with high sensitivity and a dynamic range of up to 10 orders of magnitude.

Diagnostic Applications

qPCR is the gold standard for pathogen detection, including viral RNA quantification (e.g., SARS-CoV-2). Used in oncology for minimal residual disease monitoring and in clinical microbiology for rapid, quantitative pathogen load assessment.



Reverse Transcription PCR (RT-PCR)

Reverse Transcriptase Enzyme

Reverse transcriptase converts single-stranded RNA into complementary DNA (cDNA), enabling downstream molecular analysis of RNA molecules.

RNA to cDNA Conversion

mRNA is isolated and used as a template. Reverse transcriptase synthesizes a stable cDNA copy, preserving the gene expression profile of the sample.

PCR Amplification of cDNA

The synthesized cDNA is amplified by standard PCR or quantitative qPCR, enabling detection and quantification of specific gene transcripts.

Applications in Research

RT-PCR is widely used in transcriptomics, gene expression studies, and detection of RNA viruses such as SARS-CoV-2 and influenza.

Specialized PCR Techniques

Method / concept	Principle and applications
Multiplex PCR	Amplifies multiple distinct target sequences simultaneously in a single reaction using several primer pairs. Reduces time and sample usage; widely applied in pathogen panels and genetic screening.
Nested PCR	Employs two successive primer sets: an outer pair in round one, then an inner pair amplifying within the first product. Greatly enhances specificity and sensitivity for low-abundance targets.
Digital PCR	Partitions the sample into thousands of micro- reactions; each contains zero or one template molecule. Poisson statistics yield absolute quantification without a standard curve.
Comparative Applications	Multiplex PCR suits high- throughput diagnostics; nested PCR excels in forensic and clinical trace detection; digital PCR is preferred for rare mutation detection and copy-number variation analysis.

PCR Troubleshooting

Non-Specific Bands & Primer-Dimers

Non-specific amplification occurs when primers bind off-target sequences. Primer-dimers form when primers self-anneal. Increase annealing temperature and redesign primers with higher specificity to resolve these issues.

Mg^{2+}

Concentration Optimization

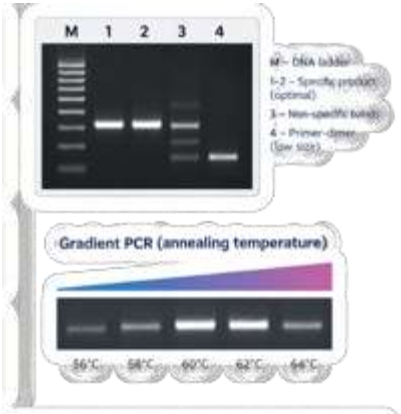
Magnesium ion concentration critically affects polymerase activity and primer binding. Too low yields no product; too high increases non-specific amplification. Titrate $MgCl_2$ between 1.0–4.0 mM to find optimal conditions.

Annealing Temperature Gradients

Gradient PCR systematically tests a range of annealing temperatures across a thermal block. Identifies the optimal temperature that maximizes specific product yield while minimizing non-specific bands and primer-dimer formation.

Hot-Start Enzymes

Hot-start polymerases remain inactive below activation temperature, preventing non-specific amplification during reaction setup. Antibody-mediated or chemically modified enzymes activate only at high denaturation temperatures, improving specificity and yield.



DNA Sequencing Fundamentals

Nucleotide Order

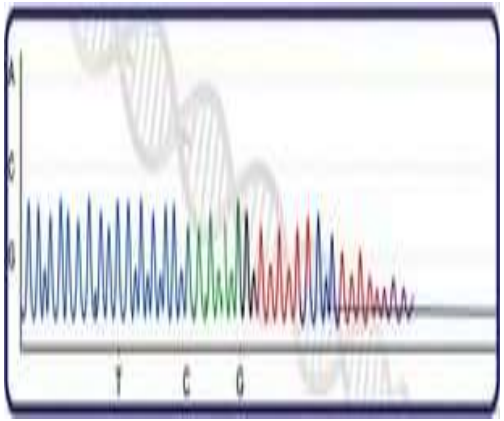
Sequencing determines the exact order of bases (A, T, C, G) in a DNA strand, providing the complete genetic blueprint essential for biological and medical research.

Technology Generations

Three generations of sequencing: 1st gen (Sanger), 2nd gen (NGS/Illumina short reads), and 3rd gen (PacBio, Nanopore long reads), each advancing throughput and read length.

Genetics & Diagnostics

Sequencing enables mutation detection, disease diagnosis, pathogen identification, and population genetics studies by revealing variants at single-nucleotide resolution. Applied in whole-genome sequencing, transcriptomics, epigenomics, and metagenomics to advance understanding of gene function, evolution, and complex biological systems.



Sanger Sequencing

Chain Termination Principle

Dideoxynucleotides (ddNTPs) lack a 3'-OH group, terminating DNA synthesis at specific bases. Each ddNTP is fluorescently labelled, generating fragments of varying lengths.

Capillary Electrophoresis

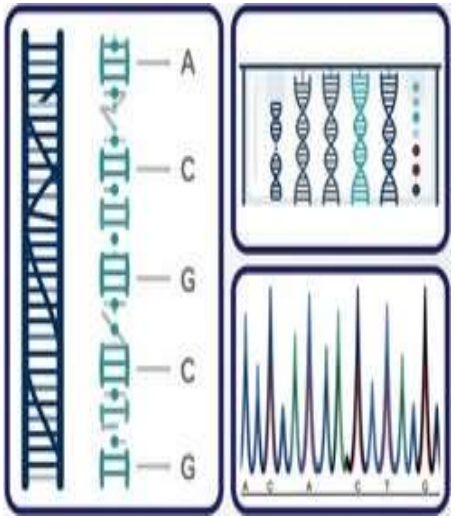
Terminated fragments are separated by size through capillary electrophoresis. Smaller fragments migrate faster, producing a size-ordered sequence of fluorescent signals.

Chromatogram Output

Fluorescence detection generates a chromatogram with colour-coded peaks corresponding to each nucleotide (A, T, C, G), enabling direct reading of the DNA sequence.

Read Length & Accuracy

Sanger sequencing produces highly accurate reads of 800–1000 base pairs per run. It remains the gold standard for validating cloned inserts and targeted sequence verification.



Next-Generation Sequencing (NGS)

Illumina Platform

The Illumina platform dominates short-read NGS. Sequencing-by- synthesis uses reversible dye terminators to identify each base as it is incorporated, enabling massively parallel reads across millions of clusters simultaneously.

Bridge Amplification

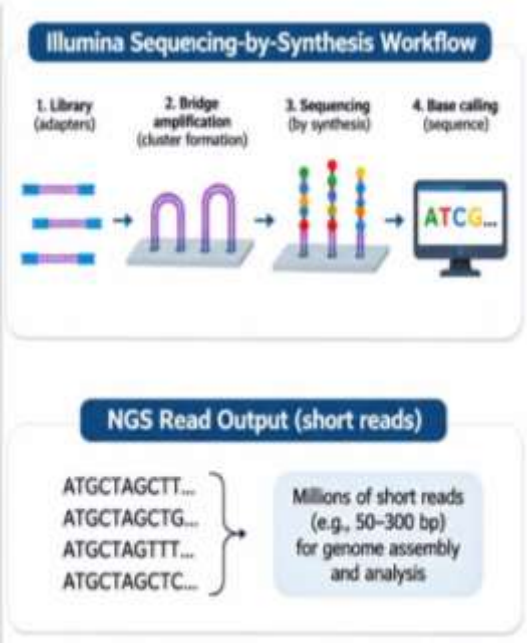
DNA fragments are ligated to adapters and loaded onto a flow cell. Each fragment bends and hybridizes to adjacent surface-bound oligos, forming bridges. Repeated cycles generate dense clusters of identical copies for robust signal detection.

Reversible Terminators

Fluorescently labeled, reversible chain-terminating nucleotides are incorporated one at a time. After imaging to identify each base, the terminator is chemically cleaved, allowing the next cycle to proceed with high accuracy and low error rates.

Genome-Scale Output

NGS enables whole genome sequencing (WGS), whole exome sequencing (WES), and transcriptome profiling (RNA-seq) at scale. High throughput and reduced cost make it the standard platform for large-cohort and clinical genomics studies.



Third-Generation Single-Molecule Sequencing

PacBio Platform

Single-molecule real-time (SMRT) sequencing observes DNA synthesis in zero-mode waveguides. Generates reads exceeding 10,000 bases with high consensus accuracy.

Oxford Nanopore

Nanopore sequencing detects ionic current changes as DNA translocates through a protein pore. Enables ultra-long reads and real-time base calling without PCR amplification.

Long-Read Advantages

Reads >10,000 bases resolve complex genomic regions including repetitive elements, structural variants, and segmental duplications unresolvable by short-read platforms.

Epigenetics & Modifications

Long-read platforms directly detect base modifications such as 5-methylcytosine and 6-methyladenine without bisulfite conversion, enabling native epigenome profiling.

Sequencing Technologies Compared

Technology	Characteristics and uses
Sanger (1st generation)	800–1000 bp; >99.9% accuracy; low throughput. Targeted sequencing and clone verification.
NGS (2nd generation)	75–300 bp reads; >99.9% with coverage; very high throughput and low cost per Mb. Genome and transcriptome studies.
Long-read (3rd generation)	>10,000 bp, up to Mb; improving accuracy (~99% in the source). Repeats, structural variants, epigenetics, and full-length transcripts.
Select by application	Sanger: verification. NGS: discovery and diagnostics. Long reads: complex assembly and isoform analysis.

Bioinformatics & Data Analysis

Quality Control & Filtering

Raw sequencing reads are assessed using tools such as FastQC. Low-quality bases, adapter sequences, and short reads are trimmed to ensure downstream accuracy.

Reference Genome Alignment

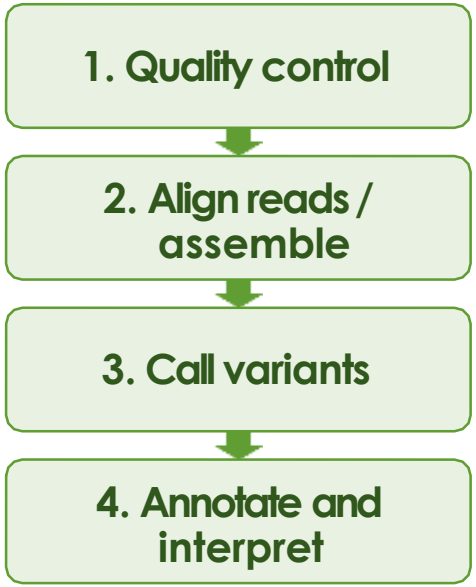
Filtered reads are aligned to a reference genome using tools like BWA or HISAT2, producing sorted BAM files that record each read's genomic position.

De Novo Assembly

When no reference exists, reads are assembled into contigs and scaffolds using assemblers such as SPAdes or Velvet, reconstructing the genome from overlapping sequences.

Variant Calling & Annotation

SNPs and indels are identified from aligned BAM files using GATK or FreeBayes, then annotated with functional impact using tools such as ANNOVAR or SnpEff.



Recombinant DNA Technology

Gene Isolation & Replication

Target genes are isolated from genomic DNA using restriction enzymes or PCR amplification. Cloning allows indefinite replication of the gene of interest within a host organism.

Vector Cloning

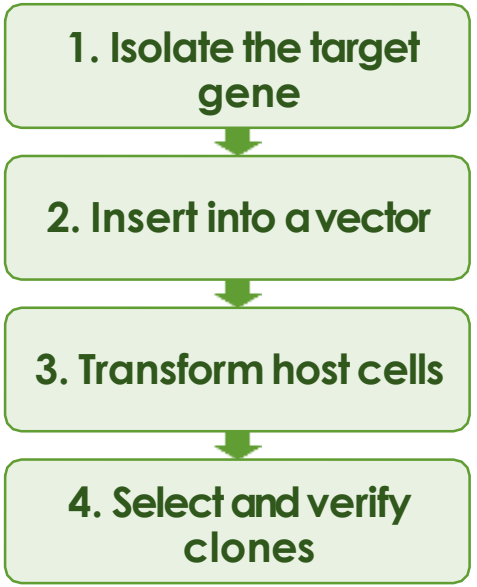
Insert DNA is ligated into plasmid vectors and transformed into competent bacterial host cells. Recombinant colonies are selected via antibiotic resistance markers.

Gene Expression Studies

Cloned genes are expressed using promoter-driven vectors to produce recombinant RNA or protein. This enables functional characterization and protein structure-function analysis.

Biotech Foundation

Recombinant DNA technology underpins modern biotechnology: therapeutic protein production, transgenic organisms, and gene therapy all rely on core cloning principles.



Molecular Scissors: Restriction Endonucleases

Palindromic Sequence Recognition

Restriction enzymes recognize specific palindromic DNA sequences, typically 4-8 base pairs long. The sequence reads identically on both strands in the 5' → 3' direction, ensuring precise and reproducible cleavage at defined sites.

Sticky End Generation

Manyrestriction enzymes cut asymmetrically within the recognition site, producing short single-stranded overhangs called sticky ends. These complementary overhangs facilitate efficient ligation of compatible DNA fragments in cloning workflows.

Blunt End Cleavage

Certain restriction enzymescut both strands at the same position, generating blunt-ended fragments with no overhangs. Blunt-end ligation is less efficient but allows joining of any two DNA fragments regardless of sequence compatibility.

Role in DNA Manipulation

Restriction endonucleases are foundational tools in molecular biology. They enable precise excision of target genes, construction of recombinant vectors, restriction mapping, and are essential for cloning, library preparation, and diagnostic applications.

Vector Architecture

Origin of Replication

The ori sequence allows autonomous plasmid replication within the host cell. Copy number is determined by ori type: high-copy (pUC) or low-copy (pBR322) origins suit different experimental needs.

Multiple Cloning Site

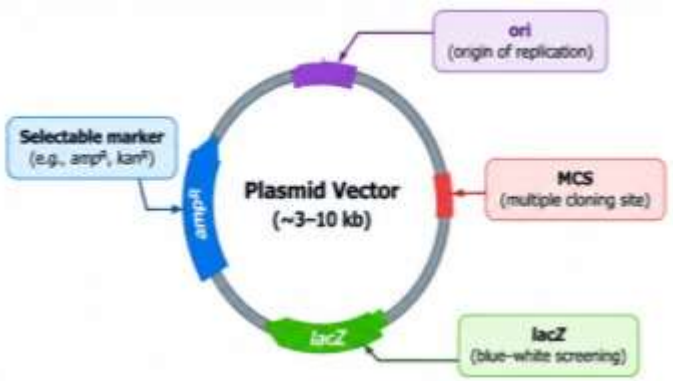
The MCS (polylinker) contains clustered unique restriction sites for directional insert cloning. Flanking promoter sequences enable in vitro transcription and expression of the cloned insert.

Selectable Markers

Antibiotic resistance genes (e.g., ampicillin, kanamycin) allow selection of transformed colonies. Blue-white screening via lacZ disruption identifies recombinant clones carrying an insert.

BACs & Large-Insert Vectors

Bacterial Artificial Chromosomes (BACs) accommodate inserts of 100-300 kb. Bacteriophages, cosmids, and fosmids extend cloning capacity beyond standard plasmid limits for genomic libraries.



Vector Types and Insert Capacity

Vector	Typical Insert Size	Example Use
Plasmid	~1–10 kb	General cloning, expression
Cosmid	~35–45 kb	Medium-size genomic inserts
Fosmid	~40 kb	Genomic library construction
BAC	~100–300 kb	Large DNA fragments, genomics

DNA Ligation & Assembly Methods

Method / concept	Principle and applications
T4 DNA Ligase	Catalyzes the formation of phosphodiester bonds between adjacent DNA fragments. Requires ATP as a cofactor. Joins both blunt and cohesive (sticky) ends generated by restriction digestion.
Phosphodiester Bond Formation	Ligation seals nicks in the DNA backbone by forming covalent 3'-5' phosphodiester bonds. Efficiency depends on fragment concentration, temperature (16°C overnight), and compatible end types.
Gibson Assembly	Restriction-free, seamless multi-fragment assembly in a single isothermal reaction. Uses 5' exonuclease, DNA polymerase, and ligase. Requires ~20-40 bp overlapping sequences between fragments.
Golden Gate Assembly	Uses Type IIS restriction enzymes (e.g., BsaI) that cut outside their recognition sequence. Generates custom 4-nt overhangs enabling scarless, ordered multi-fragment assembly in one-pot reactions.

Transformation & Selection

Competent Cell Preparation

Cells are made competent via chemical treatment ($CaCl_2$) or electroporation, transiently increasing membrane permeability to allow uptake of foreign plasmid DNA.

Heat-Shock & Electroporation

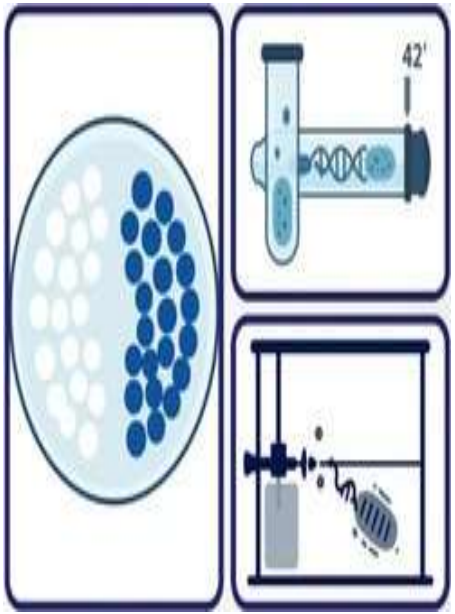
Heat-shock (42°C, 90 sec) or brief high-voltage pulse drives recombinant plasmid through the bacterial membrane into the cytoplasm for stable transformation.

Antibiotic Selection

Transformed cells are plated on selective media containing antibiotics (e.g., ampicillin). Only colonies harbouring the resistance-carrying plasmid survive and grow.

Blue-White Screening

Insertional inactivation of the lacZ gene distinguishes recombinant (white) from non- recombinant (blue) colonies on X-Gal/IPTG plates, enabling rapid identification of positive clones.



Recombinant Protein Expression

Expression Vectors

Strong promoters such as T7 drive high-level transcription of cloned genes. Expression vectors carry regulatory elements including ribosome binding sites and terminators optimised for the chosen host organism.

Affinity Tags

His-tag (6xHis) and GST-tag sequences are fused to the target protein, enabling selective capture on metal-chelate or glutathione resin columns for rapid, high-purity protein purification.

Induction & Cell Lysis

IPTG induction activates T7 RNA polymerase-driven expression. Cells are harvested and lysed by sonication or French press, releasing soluble recombinant protein for downstream purification steps.

Chromatographic Purification

Immobilised metal affinity chromatography (IMAC) or GST-affinity resin captures tagged protein. Subsequent size-exclusion or ion-exchange steps yield highly pure, functionally active recombinant protein.



Gene Editing & Functional Genomics

CRISPR-Cas9 Mechanism

CRISPR-Cas9 enables precise, targeted genome modification. The Cas9 nuclease is guided to a specific DNA locus by a single guide RNA (sgRNA), where it introduces a double-strand break at the target sequence.

Guide RNA Design

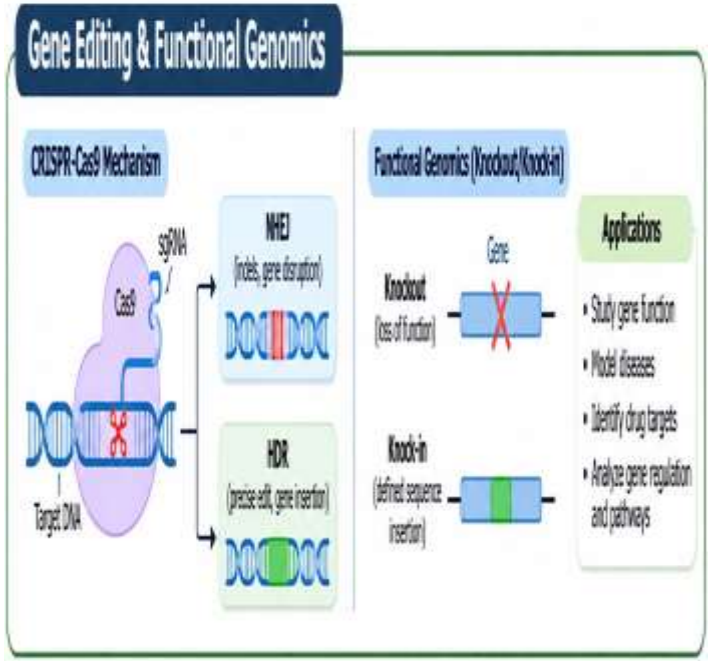
The sgRNA is engineered to match the target DNA sequence adjacent to a PAM motif (NGG for SpCas9). Specificity is critical: off- target analysis and bioinformatic design tools minimize unintended edits.

Double-Strand Break Repair

DSBs are repaired via NHEJ (non-homologous end joining), introducing insertions/deletions (indels) that disrupt gene function, or via HDR (homology-directed repair) to introduce precise sequence changes.

Knockout & Knock-in Studies

Functional knockouts abolish gene expression to study loss-of-function phenotypes. Knock-ins insert defined sequences or reporter genes, enabling mechanistic dissection of gene function in cell and animal models.



Medical Diagnostics

PCR Pathogen Detection

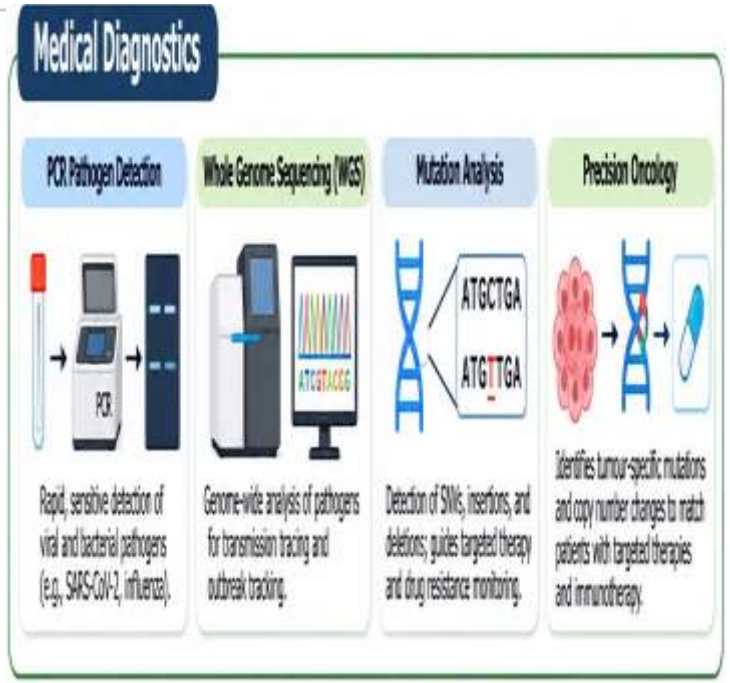
PCR enables rapid, sensitive detection of viral and bacterial pathogens from clinical samples. RT-PCR detects RNA viruses such as SARS-CoV-2 and influenza, providing results within hours for timely clinical decision-making. Outbreak Tracking
Whole genome sequencing (WGS) of pathogen isolates enables phylogenetic analysis to trace transmission chains. Real-time genomic surveillance identifies emerging variants and supports public health interventions during epidemics.

Mutation Analysis

NGS-based sequencing detects single nucleotide variants (SNVs), insertions, and deletions in pathogen or host genomes. Mutation profiling informs drug resistance monitoring and guides targeted therapeutic strategies in infectious disease.

Precision Oncology

Tumour sequencing identifies somatic driver mutations, copy number alterations, and fusion genes. Molecular profiling matches patients to targeted therapies and immunotherapies, improving outcomes and enabling liquid biopsy-based
monitor ing.



Research Applications: Synthetic Biology & Biotechnology

Modular Genetic Circuit Design

Synthetic biology uses standardized genetic parts (promoters, RBSs, terminators) assembled into circuits that regulate gene expression, enabling programmable cellular behavior for research and industrial applications.

Therapeutic Protein Production

Recombinant expression systems produce insulin, monoclonal antibodies, and vaccines. Host organisms (E. coli, yeast, CHO cells) are engineered with optimized expression vectors and affinity-tagged constructs.

Biofuels & Transgenic Crops

Metabolic engineering redirects cellular pathways to produce ethanol, butanol, and biodiesel. Transgenic crops are engineered for pest resistance, drought tolerance, and enhanced nutritional profiles.

Industrial Bioprocess Optimization

Fermentation scale-up, bioreactor design, and process analytical technology (PAT) ensure consistent yields. Directed evolution and systems biology tools refine microbial strains for industrial production.

Forensics & eDNA

STR Profiling

Short tandem repeat (STR) profiling via multiplex PCR amplifies specific loci simultaneously. Each individual has a unique STR pattern, enabling reliable identity verification in forensic casework.

Environmental DNA (eDNA)

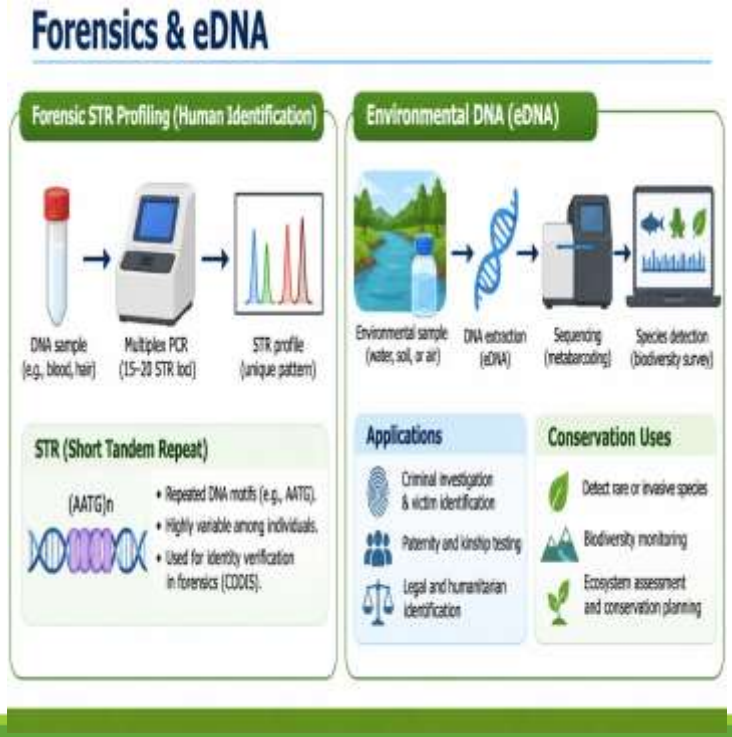
DNA shed by organisms into water, soil, or air is captured and sequenced. Metabarcoding of eDNA enables non-invasive biodiversity surveys and species detection without direct observation.

Forensic Applications

Multiplex PCR panels target 15–20 STR loci in a single reaction. Combined DNA Index System (CODIS) databases allow cross-referencing of profiles for criminal investigation and victim identification.

Conservation Biology

eDNA metabarcoding sequences bulk environmental samples to detect rare or invasive species. This supports ecological monitoring, conservation planning, and early detection of biodiversity loss.



Emerging Trends Single-Cell Genomics & Spatial Omics

Single-Cell Sequencing

Resolves cellular heterogeneity invisible in bulk assays. Each cell is individually barcoded and sequenced, revealing distinct transcriptional states and rare cell populations.

Cellular Heterogeneity

Tumours and tissues contain diverse cell subtypes. Single-cell RNA-seq uncovers subpopulations, lineage trajectories, and differential gene expression at single-cell resolution.

Spatial Transcriptomics

Maps RNA expression while preserving tissue architecture. Spatially barcoded arrays capture transcriptomes in situ, linking molecular profiles to anatomical context.

Integration & Applications

Combining single-cell and spatial omics advances understanding of development, disease, and drug response. Enables high- resolution atlases of complex biological systems.

Workshop Assignment

Scenario & Objective

Clone a 1.5 kb human kinase gene into an AmpR plasmid vector containing a HindIII restriction site. Apply PCR, cloning, transformation, and sequencing techniques to complete the workflow.

Task 1: PCR Amplification

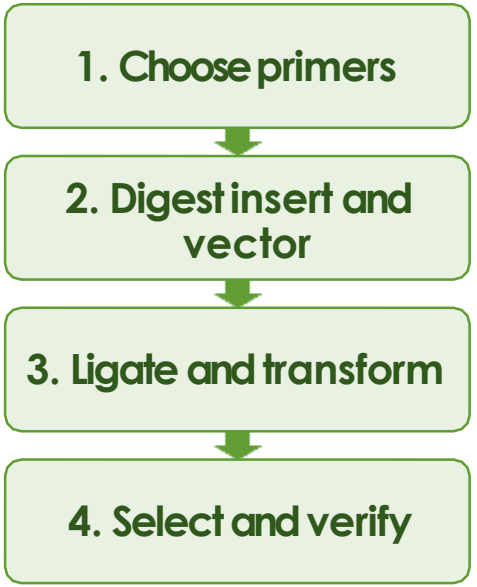
Design PCR thermal cycling steps and list all required components (template DNA, primers with HindIII overhangs, dNTPs, Taq polymerase, buffer) to amplify the 1.5 kb kinase gene insert.

Task 2: Transformation & Selection

Describe the transformation protocol (heat-shock or electroporation) and antibiotic selection strategy to identify positive recombinant colonies carrying the cloned insert on ampicillin plates.

Task 3: Sequencing Verification

Select and justify a sequencing method — Sanger, NGS, or Nanopore — to confirm an error-free insert. Consider read length, accuracy, cost, and turnaround time in your justification.



References & Core Textbooks

Green & Sambrook (2012), Molecular Cloning: A Laboratory Manual. Watson et al. (2007),

Recombinant DNA: Genes and Genomes.

Alberts et al. (2017),

Molecular Biology of the Cell.

Ausubel et al. (2003), Current Protocols in Molecular Biology. Core Techniques: PCR, Sequencing, Cloning

