CRIME SCENE DO NOT CROSS CRIME SCENE DO NOT CROSS

Lab Manager

NGS FOR FORENSIC INVESTIGATIONS

Next-generation sequencing (NGS) is used to sequence millions or billions of DNA strands in parallel. Forensic analyses require high accuracy and reproducibility, however templates of low copy number, and degraded or contaminated samples pose challenges. NGS offers some advantages to current techniques used for forensic DNA analysis, and could potentially be a valuable technique for forensic investigations.

CRIME SCENE DO NOT CROSS CRIME SCENE DO NOT CROSS

NEXT-GENERATION SEQUENCING: BASIC METHODOLOGY

LIBRARY PREPARATION

An NGS library consists of similarly sized DNA fragments with known adapter sequences on the 5' and 3' ends.

Different preparations all share four basic steps:

- DNA Fragmentation
- Addition of adapter sequences
- Size selection
- Library quantification and QC

SEQUENCING

Pyrosequencing:
The reaction is monitored through the release of a pyrophosphate during nucleotide incorporation.

Sequencing by synthesis:
Incorporation of reversibly-fluorescent and terminated nucleotides.

Sequencing by ligation:
Instead of DNA polymerase, this method relies on 16 8-mer oligonucleotide probes, each with 4 fluorescent dyes.

Ion semiconductor sequencing:
Relies on the release of H+ ions during sequencing reactions to detect the sequence of a cluster.

DATA ANALYSIS

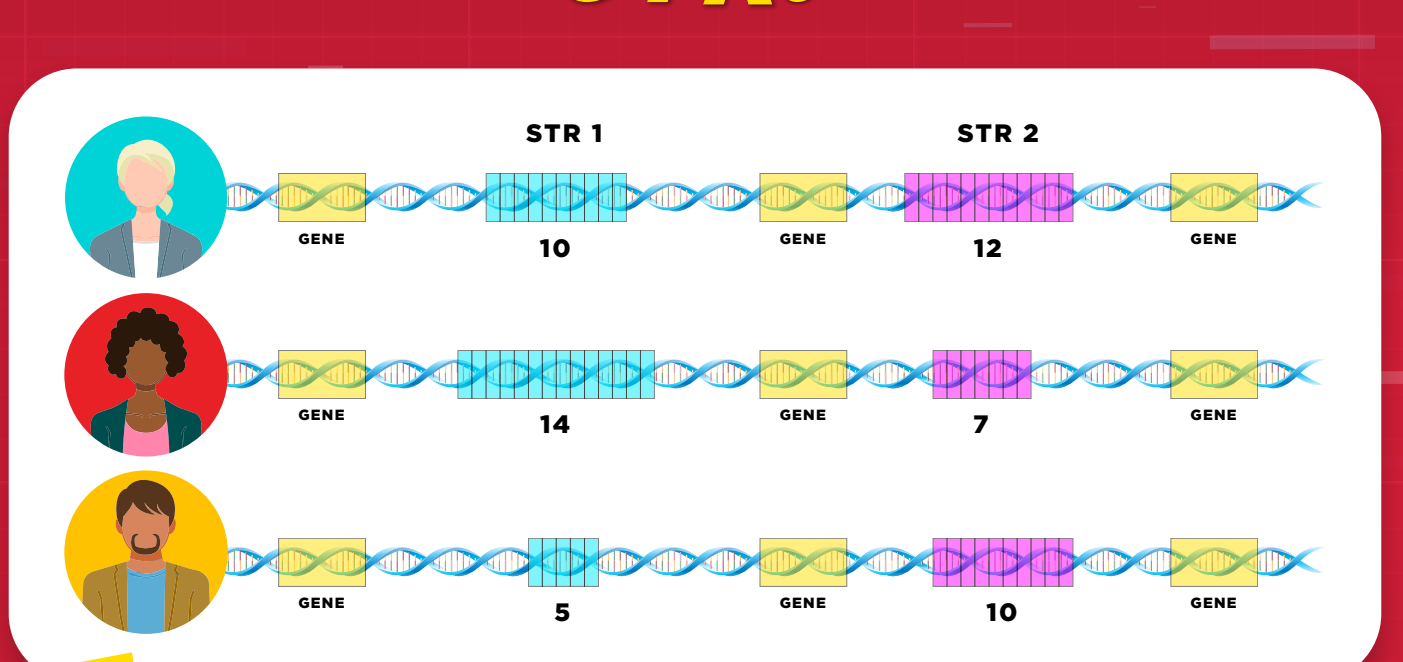
Sequencing produces an enormous amount of raw data for processing. There are several computational tools available for dry-lab data anlysis, including BWA, Bowtie, Galaxy, and SanGeniX. Data analysis involves:

- raw data quality assessment
- read alignment to a reference genome
- variant identification
- variant annotation
- data visualization

CRIME SCENE DO NOT CROSS CRIME SCENE DO NOT CROSS

NEXT-GENERATION SEQUENCING FOR FORENSIC INVESTIGATION

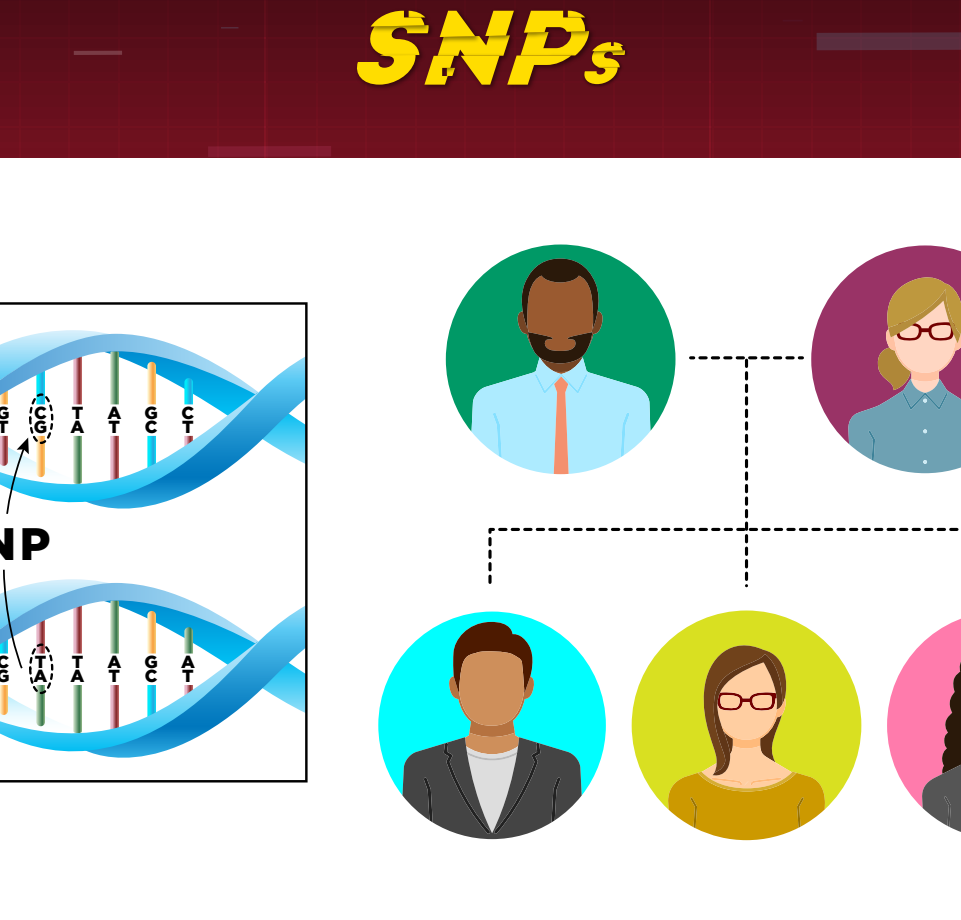
STRs



Short tandem repeats (STR) are stretches of DNA with approximately 30 repeating units of 1-6 bps. The number of repeats within the STR is referred to as an allele. STRs have a high degree of polymorphism, making them ideal genetic markers.

STR analysis is one of the most frequently used methods for human identification, with allele classifications identified using capillary electrophoresis (CE). Several countries have forensic DNA databases based on STRs. The combined DNA Index System (CODIS) includes 13 Core STR Loci: CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11. The potential advantages of using NGS for STR analysis include higher throughput, and the ability to analyze degraded samples, multiplex more markers, and distinguish alleles with similar length or digital read count.

MtDNA



Mitochondrial DNA (mtDNA) is the genetic information encoded in the mitochondrion. It is maternally inherited, and unlike nuclear DNA, has a high copy number making it more likely that a sufficient quantity can be recovered from compromised samples.

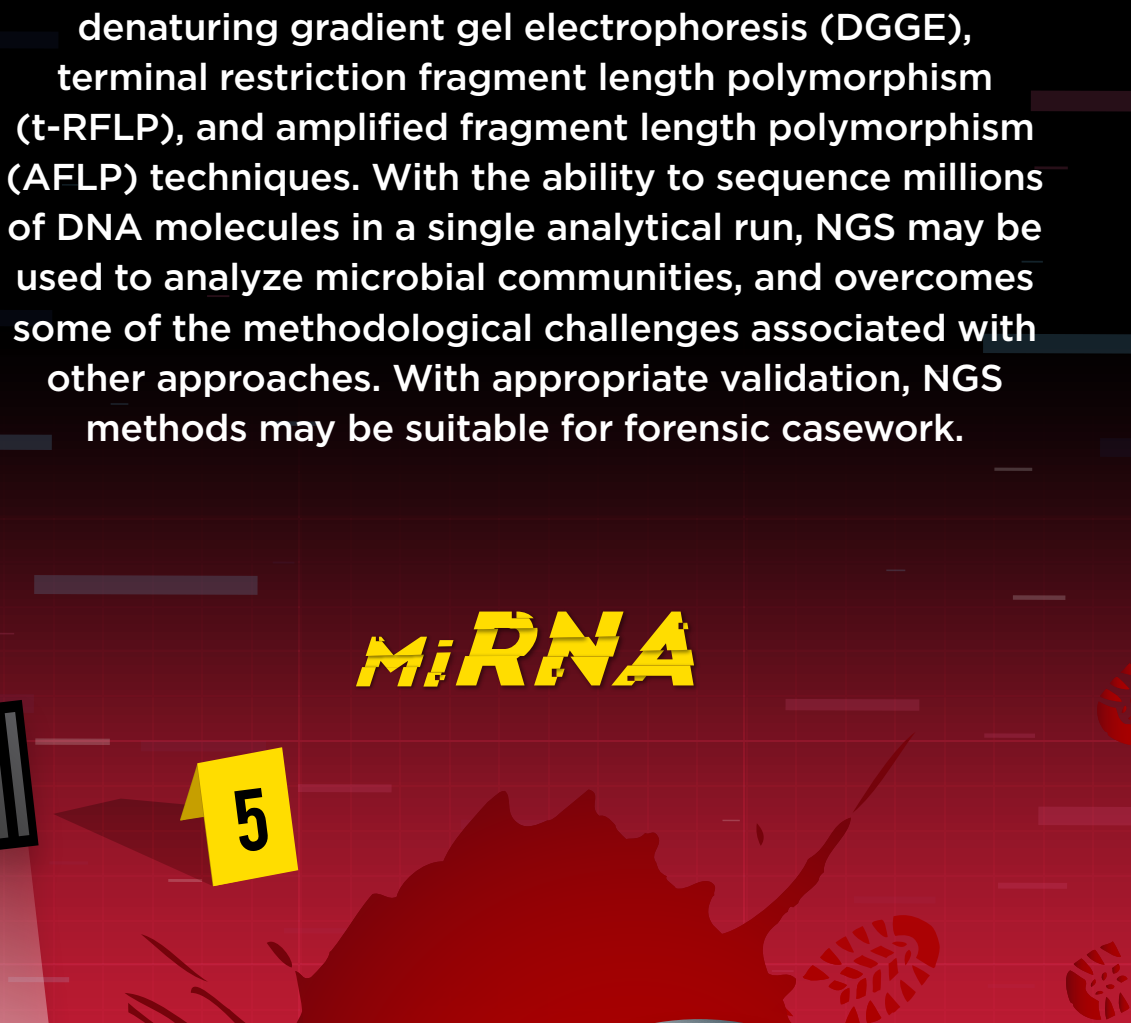
mtDNA analyses are valuable for missing persons investigations as well as mass disasters. Analyses often focus on the highly polymorphic Control Region with hypervariable regions I and II via Sanger sequencing. In some cases, common mtDNA types result in inconclusive identifications. NGS enables analysis of the entire mtDNA genome, potentially providing additional polymorphic loci to increase discriminatory power. NGS may also enable detection of mtDNA heteroplasmy- the presence of more than one mtDNA type.

SNPs

Single nucleotide polymorphisms (SNPs) are alleles that differ by a single base, via substitution, insertion, or deletion. SNPs may be favorable over STRs due to their high abundance, low mutation rate, short amplicon length, and presence in degraded DNA.

There are established databases for STRs, making them the primary forensic markers. However, there may be a potential role for SNPs in the future of forensic DNA typing. Forensically relevant SNPs include: identity-testing SNPs, lineage informative SNPs (used for kinship analyses), ancestry informative SNPs (biogeographical ancestry for inference of phenotypic characteristics), and phenotype informative SNPs (to determine the probability of a specific phenotypic characteristic like eye color). Single base extension capillary electrophoresis enables SNP analysis, however NGS workflows may achieve higher throughput at lower cost.

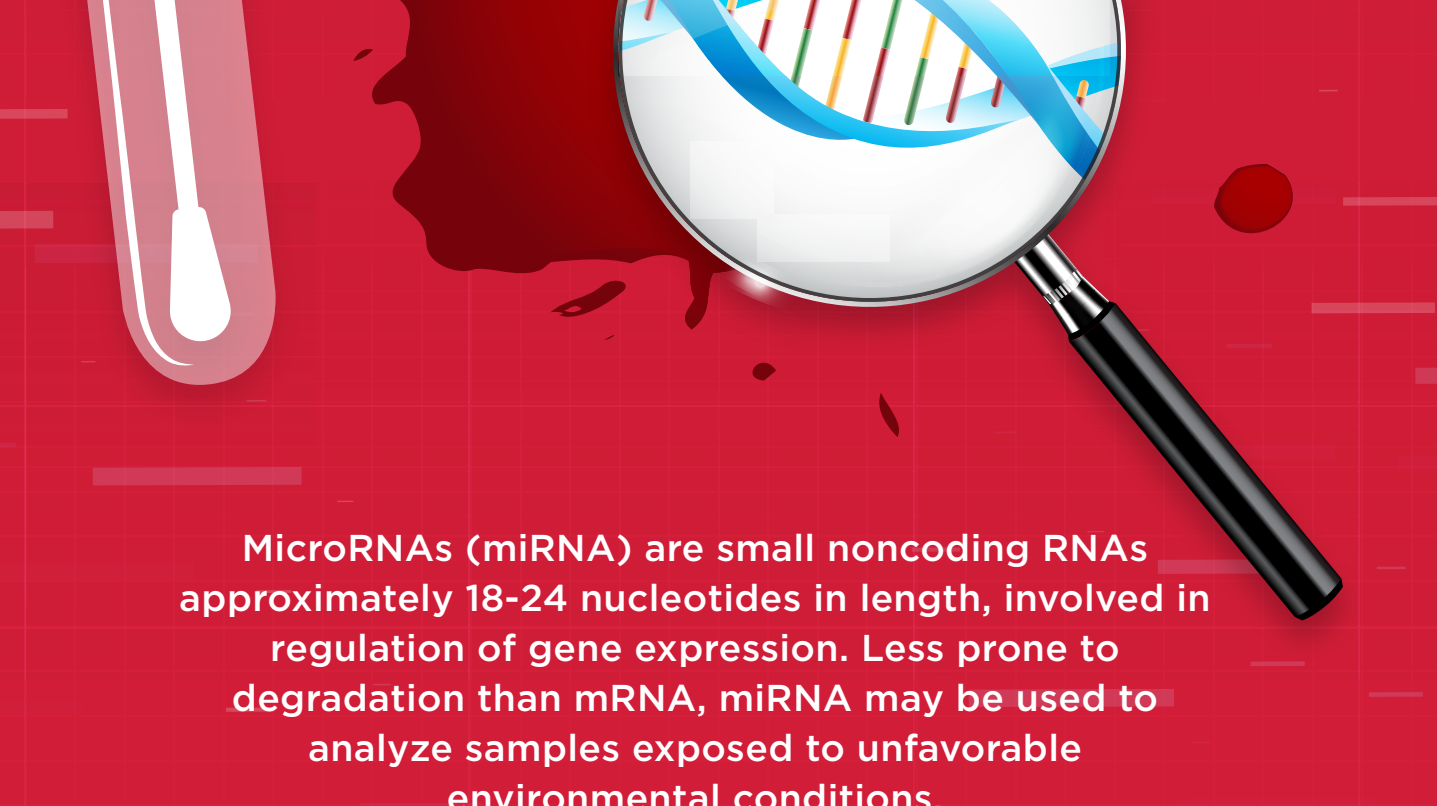
MICROBES



Microbial analyses are used in forensic investigations to compare soil samples, identify strains of infection, and for human identification and drowning cases.

Microbial population profiles have been generated using denaturing gradient gel electrophoresis (DGGE), terminal restriction fragment length polymorphism (t-RFLP), and amplified fragment length polymorphism (AFLP) techniques. With the ability to sequence millions of DNA molecules in a single analytical run, NGS may be used to analyze microbial communities, and overcomes some of the methodological challenges associated with other approaches. With appropriate validation, NGS methods may be suitable for forensic casework.

MiRNA



MicroRNAs (miRNA) are small noncoding RNAs approximately 18-24 nucleotides in length, involved in regulation of gene expression. Less prone to degradation than mRNA, miRNA may be used to analyze samples exposed to unfavorable environmental conditions.

miRNAs are often used for body fluid identification, as the level of miRNA present in blood, urine, saliva, and sweat may increases as a result of tissue insult. Known miRNA sequences may be analyzed with real-time PCR. NGS enables rapid, untargeted detection of all miRNA in a sample.

CRIME SCENE DO NOT CROSS CRIME SCENE DO NOT CROSS