



Applications of Next-Generation Sequencing in Forensic Field

23

María Saiz, Luis Javier Martínez-Gonzalez, Jose Antonio Lorente, and Maria Jesus Alvarez-Cubero

Abstract

Next-generation sequencing (NGS) technology has developed and improved rapidly in the recent years, and many research fields use this methodology for its application. In the forensic field, new strategies and opportunities have emerged by the use of NGS which could be useful in analyzing multiple loci in certain samples and scenes. NGS overtakes the limitations of using the most common technologies such as capillary electrophoresis, STR, or SNP markers analysis for human genetic DNA typing. Nowadays, there are more equipment, analysis strategies, or even regions of analysis such as microbiome, miRNAs, or other markers. Here we report the evolution of NGS in forensic field, the workflow, and the opportunities of these methodologies in forensic research.

M. Saiz

Laboratory of Genetic Identification, Department of Legal Medicine, Toxicology and Physical Anthropology, Faculty of Medicine, University of Granada, Granada, Spain

L. J. Martínez-Gonzalez

Centre for Genomics and Oncological Research, GENYO (University of Granada-Andalusian Government—Pfizer), Granada, Spain

J. A. Lorente (✉)

Laboratory of Genetic Identification, Department of Legal Medicine, Toxicology and Physical Anthropology, Faculty of Medicine, University of Granada, Granada, Spain

Centre for Genomics and Oncological Research, GENYO (University of Granada-Andalusian Government—Pfizer), Granada, Spain

e-mail: jose.lorente@genyo.es

M. J. Alvarez-Cubero

Centre for Genomics and Oncological Research, GENYO (University of Granada-Andalusian Government—Pfizer), Granada, Spain

Department of Biochemistry & Molecular Biology III, Faculty of Medicine, University of Granada, Granada, Spain

KeywordsNGS · STRs · Epigenetic · miRNA · Microbiome

23.1 Introduction**23.1.1 What Is It?**

Within the area of biomedicine, the techniques associated with molecular genetics, highlighting DNA sequencing, have enormously advanced in the last two decades. This big progress in molecular biology techniques led to next-generation sequencing (NGS) development. New sequencing technologies are characterized by performing the massive parallel sequencing (MPS), where thousands of reads are produced at the same time, reads coming from DNA fragment (Williams 2012).

Main technologies of NGS equipment are principally massive parallel chemical reactions, detection systems, and computational approaches to analyze sequences. These groundbreaking scientific advances have drastically decreased sequencing cost and reduced the time to a few days. Now, the amount of data derived from NGS platform is facing another challenge associated with the problem of millions of confusing results (Wong et al. 2013). Nevertheless, these know-hows have the potential to endorse another revolution in sequencing and its use (directly sequencing single molecules like DNA, RNA or even protein) and the great challenge to the substantially increase of data storage and processing (Han et al. 2014; van Dijk et al. 2014).

NGS technologies are principally applied in biomedical areas which have allowed finding the bases of some rare syndromes. Even so, developments in prenatal diagnosis are not comparable with the ones in genetics identification where there is a great variability of samples and not enough NGS data for each evidence type. Moreover, they also help to improve research in cancer diagnosis and genetically heterogeneous disorders as well as common diseases. Also, NGS has been frequently applied into the whole mitochondrial sequencing in forensic sciences. Furthermore, when a low amount of DNA is available, NGS has also found great achievements, such as in noninvasive prenatal diagnosis and in many forensic cases. So, the advances that have occurred in prenatal diagnosis can be transferred to the forensic field. Fetus DNA can be found in maternal blood, even though it is of low quality, quantity, and its difficulty to be distinguished from maternal DNA. So, the huge power of NGS seems striking for noninvasive prenatal testing. It is possible to detect trisomies with noninvasive sample collection (Buermans and den Dunnen 2014). In the last decade, advanced research approaches to circulating cell-free fetal DNA have made it possible to determine fetal rhesus D (RhD) genotype, fetal sex, aneuploidies, detection, and microdeletions of inherited paternal monogenic disorders. NGS will allow detection from hereditary mutations that cause genetic diseases (Breveglieri et al. 2019).

We should consider some ancient DNA (aDNA) reviews where the developments of DNA isolation methods combined with and short sequences, generated by NGS, have helped to increase its accessibility (Hofreiter et al. 2015). The elevated grade of sample degradation and the limited amount of DNA are still the principal obstacle for applying NGS systems in forensic fields. However, there are some studies that have also considered enriching aDNA (ancient DNA) using custom project-specific SNP panels or detecting specific genomic areas (Alvarez-Cubero et al. 2017).

23.1.2 Evolution

In 1977, Frederick Sanger announced a new method to study the DNA molecule known as Sanger sequencing; this new sequencing method was based on a chain-termination technique (Sanger et al. 1977). This was the primary technology of sequencing or the “first generation” for research or diagnostic sequencing applications (Collins et al. 2004; Liu et al. 2012). In 2003, the “Human Genome Project” concludes all the human genome thanks to semiautomatic capillary electrophoresis sequencers and Sanger technology (Collins et al. 2004; International Human Genome Sequencing Consortium 2004). The “Human Genome Project” verified that whole-genome sequencing (WGS) could be performed, but not routinely, due to its high economic and time costs. NGS has accelerated the development of WGS, decreasing the effort and high cost associated (Liu et al. 2012).

In 2005, the first NGS system appears as 454, it was carried out and launched by 454 and acquired by Roche in 2007, a year before Solexa developed a Genome Analyzer which would be purchased by Illumina afterwards. The same year Agencourt announced its new equipment, SOLiD (Sequencing by Oligo Ligation Detection). In 2011, Life Technologies introduced a new system based on the detection of hydrogen ions, the Ion Torrent Personal Genome Machine (PGM). These were, and still are, the most extended NGS sequencers in all fields, with Illumina instruments at the top. All of them share a satisfactory accuracy, a high throughput and a lower cost compared to Sanger sequencing (Liu et al. 2012). All these NGS platforms (categorized as Second Generation of sequencer, 2G) share the basic idea to perform a ligation of DNA fragments with specific adapter sequences. When DNA has adapters at both ends (libraries), they are linked to a surface (solid surface or microsphere), then amplified and sequenced by diverse methods, resulting in millions of reads.

Third-generation (3G) instruments are the evolution of NGS platforms. 3G systems use single-molecule, DNA or RNA, templates without PCR, due to requiring less starting material, and produce longer readings. 3G platforms include PacBio (Pacific Biosciences) and HeliScope (Helicos Biosciences) (Ku and Roukos 2013). An enzymatic template replication system is used to sequence individual clonal molecules. Genomic DNA template is captured, and the incorporation of a new base is controlled by breakup of fluorescent dye-linked pyrophosphate in the volume-limited observation window (Gut 2013). 3G instruments had some sensitivity problems, especially in some genome regions. PacBio had no sequencing problems

in regions with 100% GC content of CGG trinucleotide repeat expansions; these regions are very interesting in forensic studies (Loomis et al. 2013). Nowadays, long-read technologies are removing initial limitations in throughput and accuracy and also expanding their application domains in genomics.

4G platforms are also available such as Oxford Nanopore Technologies (ONT). Nanopore sequencing has evolved rapidly and recently. This technology is based on the straight reading of individual molecules incorporating nanopore technology to single-molecule-Seq (Ku and Roukos 2013). In forensic sciences, the precision is very important, for specific studies, low accuracy in 4G sequences can be enough to make a difference so we need higher accuracy technologies. PCR does not occur in single-molecule systems; therefore, 4G technologies do not produce artifacts like unbalanced amplification and GC-bias, and as a result, they produce homogeneous coverage and span GC-rich regions. Besides, they may span duplications or repeated structures that cannot be determined using short read sequencing (Buermans and den Dunnen 2014).

The possibility of in situ sequencing seems to be the next step in the evolution. Carrying out an experiment in an autopsy room or at a crime scene opens the doors to discovering RNA at a specific moment, cell, or tissue (Ku and Roukos 2013).

23.2 Workflow and Main Platforms of NGS

23.2.1 Workflow in NGS Sample Processing

NGS instruments need hard protocols for samples analysis. Library preparation goes from a multistep procedure starting with genomic template fragmentation, end repair, purifications, and adapters ligation, but it can become a very complex process like enrichment of targeted region of DNA or RNA. The common steps within the workflow of the library construction, which are described in Fig. 23.1, include (a) DNA/RNA fragmentation, (b) fragmented products end repairing and

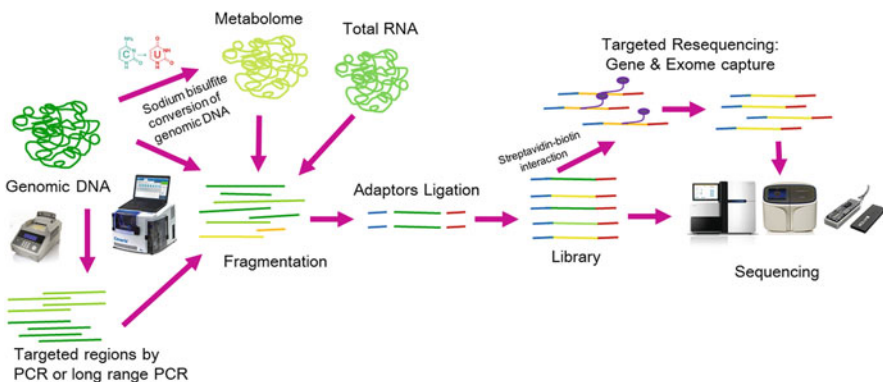


Fig. 23.1 Workflow in NGS sample preparation

A-tailing, (c) adapters ligation, (d) size selection, and (e) if necessary, amplification and, generally, after fragmentation, RNA library will require a new reverse transcription step.

23.2.2 Main NGS Platforms

Currently, there are a wide variability of NGS instruments, Ion Torrent (semiconductor sequencing technology) platform: PGM; Proton; S5, S5 plus; and S5 prime (managing 50Gb/day of sequences up to 600 bp). PacBio systems: RS II and Sequel System, Advance genomics with single-molecule real-time (SMRT) sequencing which is the longest read lengths available (average > 10,000 bp and some reads >60,000 bp), but both systems and sequences are expensive (Sequel System—PacBio [n.d.](#)). In recent years, DNBSEQ technology has broken into sequencers system, MGI company offers to world trade instruments like DNBSEQ-G400, DNBSEQ-G50, and DNBSEQ-T7. But after all, the highest percentage of sequencing data are produced by Illumina instruments: iSeq 100, MiniSeq, MiSeq Series, Nexseq 550 Series, HiSeq 2000, and Novaseq 6000 (ordered by instrument cost and generally, just the opposite, cost of sequencing). All of them produce short paired-end sequences (2×150 , except Miseq 2×300 pb) (Sequencing Platforms | Compare NGS platforms (benchtop, production-scale) [n.d.](#)). Among Illumina systems, we highlight the MiSeq FGx™, it is the first equipment fully validated for forensic science which is able to address thousands of forensic markers in a single test, through targeted sequencing of forensically relevant STRs and SNPs loci. Illumina kits combined with MiSeq FGx™ allow forensic laboratories to obtain genetic profiles from degraded, mixed, and limited DNA samples. This equipment produces high-quality, efficient forensic DNA profiles which assure sharing these profiles with others laboratories; moreover, this system performs mitochondrial sequencing which is also very demanded in forensic community. Finally, there are many other instruments but with an infrequent use, such as Oxford Nanopore systems: PromethION, GridION X5, MinION, and SmidgION; Helicos technologies; and seqll (tSMS technology, true single-molecule sequencing) (Products [n.d.](#)).

New technologies of sequencing allow an accurate processing of huge number of samples, including a sensitivity and specificity detection of differences on a population level, the discovery of novel responsible variants, and the verification of benign single/multiple-nucleotide polymorphisms. New studies such as human microbiome analysis, small RNA, noncoding RNAs, methylome analysis, ChIP sequencing, or methylated DNA immunoprecipitation force to realize new molecular designs to produce libraries for small inserts, repeated sequences, complex elements, etc. Thanks to this great development in NGS technologies and strategies, it will also improve their applications into forensic sciences and human identification.

Another field of technological development with relevance in forensic field is the single-cell sequencing studies. For NGS library preparation, it needs nanogram amounts of DNA, but a cell just contains 6–7 pg of DNA (Sanchez-Cespedes et al.

1998), so the whole-genome amplification (WGA) of the original DNA chains is a decisive step for single-cell sequencing. Methodologies developed for single-cell WGA, such as combine displacement pre-amplification and PCR amplification (indicated as PicoPLEX WGA Kit; Rubicon Genomics, Ann Arbor, MI, USA), degenerate-oligonucleotide-primed polymerase chain reaction (DOP-PCR) (such as marketed as WGA4 kit; Sigma-Aldrich, St. Louis, MO, USA), and multiple displacement amplification (MDA) (marketed as REPLI-g Single Cell Kit; QIAGEN, Germantown, MD, USA), allow us to examine the complete genome of a single cell, producing a library from this small amount of DNA. It is something incredible when talking about a crime scene sample or in the case of ancient remains.

In addition to the WGA workflow with single-cell also entails cell separation and equipment characterization. In the near future, these systems may be very useful for cases of multiple violations or in disaster victims' identifications. Now we could find a varied number of single-cell platforms. One of the first to appear and currently very widespread is the C1 (Fluidigm Corporation, CA, USA). C1 prepares single-cell templates DNA or mRNA sequencing, epigenetics, or miRNA expression. Its principal applications are identify rare cell types, survey cells diversity, or characterize cellular functions (Fluidigm | Single-Cell Analysis n.d.). Chromium Controller (10× Genomics, Inc., CA, USA) is a high-throughput automated barcoding and library construction for sequencing applications, equipment that coupled to Illumina sequencers, and allows the analysis of all types of nucleic acids, highlighting the analysis of single cell CNVs; (Home—10× Genomics n.d.). BD Rhapsody™ System (Becton, Dickinson and Company, NJ, USA) profiles the gene expression of thousands of single cells; using predesigned or customized assays, it reduces the experimentation time and the sequencing cost (BD Rhapsody™ Single-Cell Analysis System—BD n.d.). Cyto-Mine® (Sphere Fluidics Limited; CA, UK) is able to process up to ten million of varied mammalian cells in not more than half a daytime. Its workflow lets discriminating the screening of single cells to discover less common candidates (Products | Single Cell Analysis System | Cyto-Mine® n.d.). Finally, a solution that starts from Bio Rad digital PCR, ddSEQ™ Single-Cell Isolator, it is part of Illumina® Bio-Rad® Single-Cell Sequencing Solution, this technology co-encapsulates single cells and barcodes into sub-nanoliter droplets, where cell lysis and barcoding occurs (ddSEQ™ Single-Cell Isolator | Life Science Research | Bio-Rad n.d.).

23.3 Applications in Human Identification and Forensic Sciences

The use of NGS methodologies in “de novo” analysis of whole genomes (like the whole mitochondrial DNA sequence of *Neophema chrysogaster*) and in repeating the sequencing of human genomes (Human Genome Project with 1092 genomes of human samples from various populations, was developed by the use of a mixture of exome sequencing and low-coverage whole-genome) has been extensively proven. Hence, it is easy to visualize the possible uses of NGS in forensic sciences, avoiding

the principal technical limits of present technologies like capillary electrophoresis and the high rate of degraded or restricted samples of forensic cases (Irwin et al. 2011a, b; Miller et al. 2013).

Several current studies have integrated NGS data obtained from human hair generating relevant data in metagenomic analyses. A study based on forty-two extracts of DNA obtained from pericranium and pubic hairs produced approximately 80 thousand of sequences that, after post control analyses and filtering, almost 40,000 were accessible. This type of studies would be relevant for the identification of victims of sexual crimes and rapist when there are several evidences (Tridico et al. 2014). There are not many attentions in researching of Indels and CNVs, although repeats cover up nearly the half of the human genome, and STRs just the 15%. The reason is that pyrosequencing were just developed in platforms with enough read length to sequence the nucleus of STR loci applied in forensic genetics, and in these days, the majority of the forensic literature with NGS STRs data were developed using pyrosequencing methodologies. In not may reports, sequencing by synthesis and recently also semiconductor sequencing were also applied for the construction of libraries directly by PCR or using adapter ligation.

23.3.1 Short Tandem Repeats (STRs) and Single-Nucleotide Polymorphisms (SNPs)

DNA profile by using STR typing is the most common method for forensic analysis, and DNA sequencing has usually been used to indicate the origin and relationship between DNA samples in small quantity by the analysis of hypervariable section of mitochondrial DNA (mtDNA) (Berglund et al. 2011).

STR study has become a very common methodology for human forensic cases (Irwin et al. 2011a, b; Rockenbauer et al. 2014). STRs are normally studied by multiplexed PCR continued by capillary electrophoresis (CE)-based separation (Jobling and Gill 2004). Moreover, normal STR typing gives enough discrimination power for many of the current applications, and it has been proven that just using routine STRs can solve mistakes in relationship analysis (Li et al. 2012; Tsai et al. 2013). Furthermore, it is not capable to distinguish in the production stutter data produced from mixtures (Butler et al. 2004; McNevin et al. 2005), and dye artifacts are usually found (Zeng et al. 2015). The use of autosomal SNPs for human classification in forensic identification also has several problems. Bi-allelic SNPs are not as polymorphic as multi-allelic STRs. In DNA mixtures, analysis from several individuals SNPs are not very informative, but the analysis of a incremented number of SNPs as well as the use of multiple genotyping methodologies may neutralize this event (Kayser and de Knijff 2011). Some SNP panels have been used for individual identification; Fudan ID Panel covers 175 SNP markers and has been improved and validated by Ion Torrent™ PGM (Li et al. 2017).

In recent times, MPS is very concerned for the forensic genetic community. It offers the option to discover several hundred to thousand markers at the same time. Using large-scale sequencing offers multiple enhancements to forensic science.

Lately, NGS has been applied to study nuclear microsatellite markers, where random sequencing of a small fraction of the genome has been proven to produce high density of potential microsatellite loci in a low cost and rapid process. The DNA commission of the International Society for Forensic Genetics has just taken some considerations for using typing, analysis, and naming MPS STRs for forensic uses (Parson et al. 2016).

Furthermore, it will be possible to complement STR data with other informative forensic markers obtained by large-scale sequencing (Berglund et al. 2011). New sequencing technologies present a completely new paradigm for sequence data generation, contributing the chance to sequence up to millions of individual DNA fragment (like in DNA mixtures), and the elevated throughput sequencing when compared with Sanger sequencing with a much lower cost. Using new instruments to detect STR amplicons will make it much more discriminative for genetic identification, and with a degraded DNA input of just 31 pg complete profiles could be generated, also we could obtain partial informative profiles out of 5 pg of DNA (Zeng et al. 2015; Elena et al. 2016).

SNPs and Indels could be found in the repeated unit of STRs or in the flanking region, if nucleotide substitutions, deletions, or insertions are sequenced, but these changes are not detectable by CE analysis (Rockenbauer et al. 2014). NGS identifies the number of repetitions presented in the STRs, but also the sequence of the polymorphisms presented in them (Novroski et al. 2016). Moreover, new technologies allow to generate individual sequences of the different alleles present in a STR amplicon mixture (Van Neste et al. 2012). More identifiable alleles lead to more statistical power study and decrease the necessary number of STR loci that is necessary to be analyzed to solve a case with a high probability of coincidence. In criminal cases, it may also be easier to solve sample mixtures, when the same alleles appear in CE analyses and may be latter identified with more information by NGS. The possibility to discriminate between persons with identical allele lengths as far as the characterization of variations could be crucial in forensic or affiliation analysis. Statistical power will be increased using NGS technologies with a reduced number of SNPs or STRs (Rockenbauer et al. 2014).

The STR Sequencing Project, known as STRSeq, was started to simplify the characterization of sequence-based alleles at STR loci targeted in commercial kits (Gettings et al. 2017). This tool entails a curated database of sequence diversity at identification of STR loci, in addition to the nomenclature, essential elements, and variations described in accordance with the current guidelines (Parson et al. 2016).

So, the power of new sequencing technologies can only be used by also including complex STRs, analyzing both SNPs and STRs in parallel. NGS instrument provides a possibility for creating an all-in-one multiplex with important identification markers that include, e.g., STRs, SNPs, InDels, and mtDNA markers (Irwin et al. 2011a, b). In recent years, some biotech corporations are developing focused NGS-Panel into the forensic area using the same genomics areas applied till now, helping researchers to study their samples instead of customizing their own panels and primers. Currently, Thermo Fisher has designed some kits focused on forensic samples, Precision ID Panels, all of them based on Ion PGM System: Precision ID

GlobalFiler™ NGS STR Panels and GlobalFiler™ PCR Amplification Kit, NGS and STRs kits respectively, including the same 21 autosomal STRs and amelogenin sex markers. However, this kit adds STR sequence motifs to the usual STR profile, isometric heterozygotes, and known SNPs in flanking regions (Wang et al. 2017). In addition, three other panels have been developed to investigate about the biogeographical origin of the samples and discriminate individuals based on SNP analysis and mitochondrial DNA whole genome. The Illumina® ForenSeq™ DNA Signature kit includes 200 genetic loci (autosomal STRs, Y-STRs and X-STRs, ancestry informative SNPs, identity informative SNPs, and phenotypic informative SNPs). As the size of the targeted amplicons range from 64 to 231 bp for SNPs and 61 to 430 bp for STRs, this kit has been postulated as a good choice in MPS analysis of degraded human remains (Almohammed et al. 2017). Promega Corporation has developed the PowerSeq™ Systems Prototype Auto/Y analyzing 22 autosomal STR markers, amelogenin, and 23 Y-STR markers designed specifically for NGS on the MiSeq FGx® System (Montano et al. 2018; Silva et al. 2018). Many evaluation studies have been done in order to assess sensitivity, reproducibility, mixtures, and concordance as well as the analysis of casework and ancient DNA samples (Fattorini et al. 2017; Jäger et al. 2017; Just et al. 2017; Xavier and Parson 2017).

Thanks to the genome association studies, GWAs, it has been possible to identify SNPs for the characterization of the phenotype, indicating their ethnic origin and physical characteristics. These data are very useful in criminal cases because it can help recognize an individual, which will later be confirmed compared with a reference sample. Moreover, the capability to predict facial morphology and estimate individual-specific appearance through DNA is a very important goal to identify unknown persons (Irwin et al. 2011a, b; Rockenbauer et al. 2014). In order to achieve this great precision, it is most likely that it cannot be done with a few markers; on the contrary, a large set of genetic markers will be necessary. So, NGS and high-density arrays are useful for genotyping large numbers of SNPs, which can successfully deal with DNA degraded and in low quantity (Kayser and de Knijff 2011). Currently, a good application of NGS is discovering forensic SNP markers (Seo et al. 2013). TruSeq™ CHIP protocol has been modified and developed to detect 160 human identification, phenotypic SNPs, and ancestry; this new method is called TruSeq™ Forensic Amplicon (Warshauer et al. 2015a). This protocol is less labor-intensive than other techniques. In addition, low input DNA (1 ng) is required for library preparation (Warshauer et al. 2015a, b). Additionally, a genome-wide SNP array, containing 906,600 autosomal SNPs, Y-SNPs, and mitochondrial DNA has been developed by Affimetrix (Bridges et al. 2011).

The ability to produce a huge volume of data (in Novaseq more than 30 billion short reads per run) is the key advance of NGS technologies, obtaining these genetics data economically (Metzker 2009). Developing a software for the analysis of forensic NGS data will be one of the main challenges in genetics identification; it is needed because it is not possible to analyze the enormous amount of data manually. The application tool must be consistent and validated before the software can be applied in real case work. *Forenseq Universal Analysis Software* has been developed by Illumina, specifically designed to support forensic genomics

applications. Recently, there has been developed a new web application, *toaSTR*. It is a user-friendly tool for STR allele calling in massively parallel sequencing data independent of the system and the forensic kit used (Ganschow et al. 2017). The program differentiates automatically isoalleles from stutter and artefacts on a sequence bases; it facilitates an automatic allele calling with minimal need for review (Ganschow et al. 2017).

23.3.2 Lineage Markers

NGS has been used to discover systematically Y-STRs. 4500 Y-STRs have been genotyped, and the mutation rates have been estimated for 702 of them (Willems et al. 2016). Furthermore, sequence-based studies are beginning to reveal the full extent of structural variants (CNVs and inversions) and the great acceptance for gene loss on the Y chromosome compared with the autosomes (Massaia and Xue 2017) as far as sequence microvariance among Y-STRs (Warshauer et al. 2015a, b; Kwon et al. 2016; Iacovacci et al. 2017).

On the other hand, in Y-chromosome, the most interesting finding of NGS is the discovery of new Y-SNPs. These previously unknown Y-SNPs are relevant for Y-chromosomal phylogenies and haplogroup nomenclatures used for anthropological research and forensic identification. This finding has direct influence over phylogenies trees, appearing as new Y-chromosomal categories. Nevertheless, to include new Y-SNPs and updating and having a global consensus in forensic and anthropological sciences, a lot of work is still needed; most of the Y-chromosome recent project use Y phylogeny and Y-SNPs defined by Karafet et al. (2008). Due to the development of NGS analysis, new lineages have been described, and there is a need in the updating of the Y-chromosomal tree (Larmuseau et al. 2015). A new haplogroup has been described, A00, as the deepest-rooting known haplogroup in the Y-chromosome tree that diverged 275 thousand years ago (Mendez et al. 2013).

Many populations have been sequenced for several Y-SNPs in order to include all Y-haplogroups present and update the Y-Chromosome Consortium tree (Ochiai et al. 2016; Choi et al. 2017; Gao et al. 2017; Larmuseau et al. 2017). All the information derived from these studies has been accumulated, and the Y-chromosomal phylogenetic tree has been updated for criminal purposes (Van Geystelen et al. 2013).

mtDNA is present in higher number than nuclear DNA (nDNA), mitochondria vary in number and also mtDNA copies vary in mitochondria. So the probability of recovering useable DNA data is augmented in degraded samples that fail to yield helpful nDNA typing results (Parson et al. 2013). mtDNA is also informative of maternal biogeographic ancestry principally obtained by the HV1 and HV2 regions (Kaysner and de Knijff 2011).

The study of entire mtDNA control region offers a random match probability (RMP) of 1 in 120 and has provided valuable evidence in many cases. The extension to examine the complete mitochondrial genome is a logical consequence and required goal to maximize the information content of mtDNA analyses (Irwin

et al. 2011a). A study on Chinese Han population has determined that the RMP probability decreases 4.12% when analyzing the mitochondrial Genome (mtGenome) compared to control region analyses (Zhou et al. 2016).

New sequencing technologies have the potential to radically increase sample throughput, workflow efficiency, and detection resolution, so helping to get reliable and accurate entire mtGenome information (Yang et al. 2014; Zhou et al. 2016). Analyzing the whole mitochondrial genome in individuals with the same control region sequences will recognize different samples by the analysis of personal polymorphisms in the coding region not studied yet by traditional methods (Holland et al. 2011).

There is a substantial number of different variants on mtGenome, mainly due because there are dozens of mtDNA copies in each mitochondrion and up to hundreds of mitochondria in some tissues. Studying all mutations presented in mtDNA in tissues could be slightly arduous. If we compare NGS data with those analyzed by capillary electrophoresis, we could see the percentage of variables found using each technique, and how many of these mutations are homoplasmic or heteroplasmic. Several studies developed by NGS show us that this methodology allowed detecting a high percentage of mutations. These mutations can also be presented in homoplasmic or heteroplasmic states. Thanks to a study performed with 20 cases, it was possible to recognize more than 400 individual nucleotide substitutions that include four heteroplasmic variants confirmed by capillary electrophoresis with high equivalence (98%) (Zaragoza et al. 2010). While Sanger sequencing can detect heteroplasmy at a threshold of 10–20% but not resolve the variants, MPS is able to both detect and conclude heteroplasmy at levels of 1–2% (Rathbun et al. 2017; Gallimore et al. 2018); heteroplasmy can be detected accurately in 1 per 10,000 mtGenome copies with Illumina GAI (He et al. 2010). A recent study that analyzed six pairs of adult monozygotic twins has observed point heteroplasmies in five sets of twins, and a single nucleotide variant was detected in four sets. This results give an evidence for the hypothesis that variants of the mtGenome could be a biomarker to distinguish monozygotic twins from each other (Wang et al. 2015).

New sequencing methods are transforming data generation and have the potential to generate whole mitochondrial genomes profiles from even highly degraded specimens. It is quite simple and cost-effective. The study of mtDNA coding region data will be included in many routine forensic caseworks, and its study will not be dictated by the quantity or quality of the sample. This data allow accessibility to population databases that can be used to determine uncommonness mitochondrial haplotypes.

Recently, Precision ID mtDNA Whole Genome Panel kit has been tested to study the mtGenome of three skeletons dated to about early eighteenth to mid-nineteenth century. The three skeleton samples had the same whole mtDNA sequences with 38 mutations as compared to the rCRS, being assigned to haplogroup D4a1c. These results suggested that the three skeletons might belong to the same maternal line (Hashiyada et al. 2017). A recent study performed on chemically treated, degraded, high-quality, and nonhuman samples showed NGS methods to be exceedingly

sensitive, capable of generating entire mtGenome data from samples that failed to yield reliable sequences with standard PCR-based techniques (Marshall et al. 2017).

Many software packages have been developed for mtDNA profile generation. Both mitoSAVE (King et al. 2014) and GeneMarker HTS (Holland et al. 2017) generate haplotypes consistent with current forensic nomenclature guidelines and apply used-defined thresholds for profile reporting from high-quality samples. On the other hand, AQME automates mtDNA analysis from sequence data to forensic profiles and offers mtDNA haplogroup assignment. It has shown to produce accurate forensic profiles for high-quality, degraded, and chemically treated samples (Sturk-Andreaggi et al. 2017). In addition, MitoSuite supports quality check of alignment data, building consensus sequences, variant annotation, detection of heteroplasmic sites and haplogroup classification, contamination, and base substitution patterns (Ishiyama and Ueda 2017).

23.3.3 Microbiome

16S rRNA and metagenomic data “The Microbiome Project” enlisted around 300 people at two places in the USA, and proof in 15–18 body positions that characterizes skin, urogenital tract, gut, and oral cavity; studying them indicated that the microbial variety on the human body site was more similar to the similar human part on a different individual than to other body part on the same person (Clarke et al. 2017).

The human microbiota consists of 10–100 trillion symbiotic microbial cells harbored by each person and the genes these cells port. There are as many bacteria in our body as human cells (Hampton-Marcell et al. 2017). Understanding these genomes, we could understand the microbiome impacts on human health, but it is also interesting as innovative tool of genetic identification (Ursell et al. 2012). Two facts make microbiome as a relevant point for forensic identification: (a) the high number of these bacterial cells in a human body; (b) the subspecies level of these bacteria appears to be unique to each person. All of this offers an excellent chance to find a new identifiable marker unique to each person (Hampton-Marcell et al. 2017). There are unique methodologies that fit and classify some interesting taxa. In forensic field, microbial analysis focuses on identifying precise bacterial strains related to terrorism, illness, pollution, microbial postmortem variations, and trace evidence signs. The current use of NGS of taxonomically and/or phylogenetically useful genomic regions (i.e., 16S rRNA, 18S rRNA, and ITS) has optimized the excellent potential use of microbiome to estimate the postmortem intervals, to identify clandestine graves, and to unify people spaces or objects analyzing their skin microbes. But postmortem intervals are one of the most improved forensic investigations using metabolomics, by giving novel biomarkers linked to chronological variations after death, mainly to the high diversity and strength to unstable conditions to microbes (Ursell et al. 2012).

Currently, there are different focuses on NGS, mass spectrometry, and computational methods for improving the ability to characterize the microbial diversity

(microbiome); increasing the new applications of microbiome analysis such as, oral microbiome. Such profile may be unique to each individual and thereby useful for investigating genetics purposes. Microbiome will offer relevant data about each person's latest actions and about the place they live (Wong et al. 2013). Skin microbiome is also very personal; it is recognized that two individual's hands can vary around a 80% by the types of microbe discovered there. They are also good trace evidence, unique by person but can be easily transferred to objects associated with a given person, such as a computer keyboard, computer mouse, and cell phone, among others. Moreover, the structure of personalized skin microbial communities is stable over time although variable by season; and highly specific of person's gender and lifestyle (Metcalf et al. 2017). Nowadays, there is a commercial kit that are based on skin microbiome profiling for forensic human identification, such as hidSkinPlex, comprising of 286 bacterial (and phage) family-, genus-, species-, and subspecies-level markers (Schmedes et al. 2018).

Necrobiome or thanatobiome (defined as the community of organisms implicated in the decomposition events of a human body) is also a novel strategy used in forensic caseworks, which improves the data of forensic entomology and the obtained data from the physical description of corpses (Hyde et al. 2013; Pechal et al. 2014). NGS analysis of the necrobiome could establish the stage of decomposition, its succession, as well as important bacteria taxa from these communities. Moreover, they are also useful tools as postmortem interval estimator (Guo et al. 2016) and details about paleoforensic allowing to deepen in the genomic development of DNA microbes, pathogens as well as microbioma of putrefaction (Gorgé et al. 2016). Recent data included *Clostridium* spp. prevailed at long postmortem intervals (up to 10 days) and that these Gram-positive, anaerobic extremophiles also prevailed at shorter postmortem intervals (4 h) (Javan et al. 2017).

Lastly, details about geographical microbiome locations could also be of interest in forensic field. It is known that soil samples from hundreds of sites in the area have a specific profile, so any remains of soil sample in a scene of crime could be informative (Hampton-Marcell et al. 2017).

There are many advances to date in NGS and microbiome analysis; however, many things should be implemented in the near future and much work to be done such as including microbiome in databases such as CODIS could prove an effective method to lowering crime rates and clearing cases (Hampton-Marcell et al. 2017).

23.3.4 Epigenetic Analysis

It was in 1942 when Waddington coined the term "epigenetics," which he defined as changes in phenotype without changes in genotype. Nowadays, we know that epigenetics mechanisms interact with gene expression pathways and programs that can canalize different cell-type identities, including details of histone modifications and DNA methylation (Allis and Jenuwein 2016).

Many recent researches of epigenetic are focused on diseases such as cancer or inflammation, because epigenomic events are relevant for cellular reprogramming

(Allis and Jenuwein 2016). For example, an analysis demonstrated that KMT2D epigenetically activates PI3K/Akt path and EMT by linking both LIFR and KLF4 serving as a supposed epigenetic-based target for treating prostate cancer (Lv et al. 2017). There are also details in non-cancer diseases such as the analysis of patterns of DNA methylation and posttranslational histone modifications in chronic inflammatory diseases (Fogel et al. 2017).

In contrast, epigenetic has been explored slowly in forensic field. New researches are focused in the “epigenomic fingerprint” using epigenomic prediction of lifestyle and environmental exposures to improve DNA characterization in criminal cases where problems in DNA profiles identification are, complementing its predictions (Vidaki and Kayser 2017).

Many of the studies in epigenetic in forensic field are focused on markers for the prediction of tissue-identification and monozygotic twin bias, by using DNA methylation methodologies (methylation-sensitive restriction endonuclease, bisulfite modification, methylation-CpG linking protein, and third-generation sequencing methods, among others) (Kader and Ghai 2015).

In relation to the use of epigenetic in the estimation of human age, there is a study which analyses a total number of 27 CpG sites at three genetic loci (SCGN, Secretagogin; KLF14, Kruppel-Like Factor 14; and DLX5, distal-less homeobox 5 gene) in relation to their methylation status with age (samples range from 5 to 73 years). Finally, it is discovered that specific CpGs in SCGN and KLF14 can be used as potential epigenetic markers to estimate age using saliva and blood specimens, data with relevant interest in forensic field. However, there are still many challenges in finding universal DNA methylation markers using any body fluid, because many studies reported that DNA methylation is tissue specific. By the moment, just certain markers such as ELOVL2, which was describe as relatively stable age predictors for some cell types (Alghanim et al. 2017).

There are also data reporting the use of epigenetic for the discrimination of identical twins, demonstrating that there are twin-differentially methylated sites that can be useful in the case of twin genetic identification, that nowadays is currently impossible (Vidaki et al. 2017).

Many of the studies have been developed with great quantities of DNA, but promising reports also exist with few quantities of initial DNA (100 pg), which is relevant for forensic sciences (Yang et al. 2014). The application of epigenetic is in the initial steps in the forensic field, but promise results are described according to this data.

23.3.5 MicroRNA Analysis

In 1993 Lee was the first to describe microRNAs (miRNAs). They are defined as small RNA molecules encoded in the genome of plants and animals with highly conserved regions; their main role is to regulate the expression of genes (Lee et al. 1993). There are many data about their role in cancer development and evolution. Dr. Croce’s group was the first to highlight miRNAs role in B-cell chronic

lymphocytic leukemia cells; describing for the first time miR-15a and miR-16-1 in this cancer. Currently it is known that the role of miRNAs in cancer is wide, and they can act in different ways depending on their target genes. miRNA could function as either oncogene or tumor suppressor under certain circumstances. Recent researches are focused in identifying miRNA profile in exosomes of cancer (Peng and Croce 2016).

miRNAs have also been recently introduced in forensic field. One of the principal significant things is their little size (8–22 nucleotides in extent) which highlights them as the most promising molecules for avoiding degradation and tissue-specificity or very tissue-different expression. They are fitting for forensic detection, body fluid classification, and postmortem interval (PMI) deduction analysis (Yang et al. 2014).

There are data exposing a set of miRNAs as validated ones for human identification in several body fluids (blood, saliva, semen, menstrual blood, and vaginal secretions) and skin such as miR10b, miR203, miR374, miR451, and miR943 as well as proper candidate housekeeping genes like miR26b, miR92, miR144, and miR484 for normalizing miRNA gene expression data (Sirker et al. 2017). Recent publications have also indicated that miR-451a and miR-142-3p were observed in venous profiles and miR-205 absent; in menstrual blood samples, miR-451a, miR-141-3p, and miR-205; miR-891a, miR-10b, miR-142-3p, and miR-205 were observed in all semen samples; in saliva, miR-142-3p and miR-205 were observed (Mayes et al. 2018).

There are also efforts in identifying miRNA organ-specific markers such as hsa-miR-219a-5p, hsa-miR-122-5p, hsa-miR-205-5p, hsa-miR-208b-3p, and hsa-miR-206 for the identification of brain, liver, skin, heart, and skeletal muscle, respectively. Further, hsa-miR-9-5p and hsa-miR-124-3p as well as hsa-miR-499a-5p, hsa-miR-1-3p, and hsa-miR-133a-3p were found to be promising markers for the identification of brain and muscle in general, respectively (Sauer et al. 2017). As the usefulness of miRNAs in the forensic field to determine the origin of the evidence is more and more proven, there has been developed some strategies to multiplex the analysis of miRNA for body fluid identification (Mayes et al. 2018). However, many challenges are still open for its wide application in human identification and database updating.

References

- Alghanim H, Antunes J, Silva DSBS, Alho CS, Balamurugan K, McCord B (2017) Detection and evaluation of DNA methylation markers found at SCGN and KLF14 loci to estimate human age. *Forensic Sci Int Genet* 31:81–88. <https://doi.org/10.1016/j.fsigen.2017.07.011>
- Allis CD, Jenuwein T (2016) The molecular hallmarks of epigenetic control. *Nat Rev Genet* 17:487. <https://doi.org/10.1038/nrg.2016.59>
- Almohammed E, Zgonjanin D, Iyengar A, Ballard D, Devesse L, Sibte H (2017) A study of degraded skeletal samples using ForenSeq DNA Signature™ Kit. *Forensic Sci Int Genet Suppl Ser* 6:e410–e412. <https://doi.org/10.1016/j.fsigss.2017.09.158>

- Alvarez-Cubero MJ, Saiz M, Martínez-García B, Sayalero SM, Entrala C, Lorente JA, Martínez-Gonzalez LJ (2017) Next generation sequencing: an application in forensic sciences? *Ann Hum Biol* 44(7):581–592. <https://doi.org/10.1080/03014460.2017.1375155>
- BD Rhapsody™ Single-Cell Analysis System—BD (n.d.). <https://www.bd.com/en-us/offerings/capabilities/genomics/bd-rhapsody-single-cell-analysis-system>. Accessed 7 June 2018
- Berglund EC, Kiialainen A, Syvänen AC (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig Genet* 2(1):23
- Breviglieri G, D'Aversa E, Finotti A, Borgatti M (2019) Non-invasive prenatal testing using fetal DNA. *Mol Diagn Ther* 23:291–299. <https://doi.org/10.1007/s40291-019-00385-2>
- Bridges M, Heron EA, O'Dushlaine C, Segurado R, Morris D, Corvin A et al (2011) Genetic classification of populations using supervised learning. *PLoS One* 6(5):e14802. Retrieved from <http://dx.doi.org/10.1371/journal.pone.0014802>
- Buermans HPJ, den Dunnen JT (2014) Next generation sequencing technology: advances and applications. *Biochim Biophys Acta* 1842(10):1932–1941. <https://doi.org/10.1016/j.bbadis.2014.06.015>
- Butler JM, Buel E, Crivellente F, McCord BR (2004) Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis* 25(10–11):1397–1412. <https://doi.org/10.1002/elps.200305822>
- Ku C-S, Roukos DH (2013) From next-generation sequencing to nanopore sequencing technology: paving the way to personalized genomic medicine. *Expert Rev Med Dev* 10(1):1–6. <https://doi.org/10.1586/erd.12.63>
- Choi SS, Park KH, Nam DE, Kang TH, Chung KW (2017) Y-chromosome haplogrouping for Asians using Y-SNP target sequencing. *Forensic Sci Int Genet Suppl Ser* 6:e235–e237. <https://doi.org/10.1016/J.FSIGSS.2017.09.100>
- Clarke TH, Gomez A, Singh H, Nelson KE, Brinkac LM (2017) Integrating the microbiome as a resource in the forensics toolkit. *Forensic Sci Int Genet* 30:141–147. <https://doi.org/10.1016/j.fsigen.2017.06.008>
- Collins FS, Lander ES, Rogers J, Waterston RH, Conso I (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431:931–945
- ddSEQ™ Single-Cell Isolator | Life Science Research | Bio-Rad (n.d.). <http://www.bio-rad.com/es-es/product/ddseq-single-cell-isolator?ID=OKNWBSE8Z>. Accessed 7 June 2018
- van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C (2014) Ten years of next-generation sequencing technology. *Trends Genet* 30(9):418–426. <https://doi.org/10.1016/j.tig.2014.07.001>
- Elena S, Alessandro A, Ignazio C, Sharon W, Luigi R, Andrea B (2016) Revealing the challenges of low template DNA analysis with the prototype Ion AmpliSeq™ Identity panel v2.3 on the PGM™ Sequencer. *Forensic Sci Int Genet* 22:25–36. <https://doi.org/10.1016/j.fsigen.2015.07.011>
- Fattorini P, Previderé C, Carboni I, Marrubini G, Sorçaburu-Cigliero S, Grignani P et al (2017) Performance of the ForenSeq™ DNA Signature Prep kit on highly degraded samples. *Electrophoresis* 38(8):1163–1174. <https://doi.org/10.1002/elps.201600290>
- Fluidigm | Single-Cell Analysis (n.d.). <https://www.fluidigm.com/applications/single-cell-analysis>. Accessed 7 June 2018
- Fogel O, Richard-Miceli C, Tost J (2017) Epigenetic changes in chronic inflammatory diseases. *Adv Protein Chem Struct Biol* 106:139–189. <https://doi.org/10.1016/bs.apcsb.2016.09.003>
- Gallimore JM, McElhoe JA, Holland MM (2018) Assessing heteroplasmic variant drift in the mtDNA control region of human hairs using an MPS approach. *Forensic Sci Int Genet* 32:7–17. <https://doi.org/10.1016/J.FSIGEN.2017.09.013>
- Ganschow S, Wiegand P, Tiemann C (2017) toaSTR: a web-based forensic tool for the analysis of short tandem repeats in massively parallel sequencing data. *Forensic Sci Int Genet Suppl Ser* 6:e119–e121. <https://doi.org/10.1016/j.fsigss.2017.09.034>
- Gao T, Yun L, Zhou D, Lang M, Wang Z, Qian X et al (2017) Next-generation sequencing of 74 Y-SNPs to construct a concise consensus phylogeny tree for Chinese population. *Forensic Sci Int Genet Suppl Ser* 6:e96–e98. <https://doi.org/10.1016/j.fsigss.2017.09.043>

- Gettings KB, Borsuk LA, Ballard D, Bodner M, Budowle B, Devesse L et al (2017) STRSeq: a catalog of sequence diversity at human identification Short Tandem Repeat loci. *Forensic Sci Int Genet* 31:111–117. <https://doi.org/10.1016/J.FSIGEN.2017.08.017>
- Gorgé O, Bennett EA, Massilani D, Daligault J, Pruvost M, Geigl E-M, Grange T (2016) Analysis of ancient DNA in microbial ecology. *Methods Mol Biol* 1399:289–315. https://doi.org/10.1007/978-1-4939-3369-3_17
- Guo J, Fu X, Liao H, Hu Z, Long L, Yan W et al (2016) Potential use of bacterial community succession for estimating post-mortem interval as revealed by high-throughput sequencing. *Sci Rep* 6(1):24197. <https://doi.org/10.1038/srep24197>
- Gut IG (2013) New sequencing technologies. *Clin Transl Oncol* 15(11):879–881. <https://doi.org/10.1007/s12094-013-1073-6>
- Hampton-Marcell JT, Lopez JV, Gilbert JA (2017) The human microbiome: an emerging tool in forensics. *Microb Biotechnol* 10(2):228–230. <https://doi.org/10.1111/1751-7915.12699>
- Han J-P, Yang F, Xu C, Wei Y-L, Zhao X-C, Hu L et al (2014) A new strategy for sperm isolation and STR typing from multi-donor sperm mixtures. *Forensic Sci Int Genet* 13:239–246. <https://doi.org/10.1016/j.fsigen.2014.08.012>
- Hashiyada M, Adachi N, Akane A, Matsumoto T, Yoshimura S (2017) The detailed examination of the human ancient mitochondrial DNA using the Ion PGM™ System. *Forensic Sci Int Genet Suppl Ser* 6:e156–e158. <https://doi.org/10.1016/J.FSIGSS.2017.09.061>
- He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE et al (2010) Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* 464:610–614. <https://doi.org/10.1038/nature08802>
- Hofreiter M, Paijmans JLA, Goodchild H, Speller CF, Barlow A, Fortes GG et al (2015) The future of ancient DNA: technical advances and conceptual shifts. *BioEssays* 37(3):284–293. <https://doi.org/10.1002/bies.201400160>
- Holland MM, McQuillan MR, O'Hanlon KA (2011) Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy. *Croat Med J* 52(3):299–313. <https://doi.org/10.3325/cmj.2011.52.299>
- Holland MM, Pack ED, McElhoe JA (2017) Evaluation of GeneMarker® HTS for improved alignment of mtDNA MPS data, haplotype determination, and heteroplasmy assessment. *Forensic Sci Int Genet* 28:90–98. <https://doi.org/10.1016/j.fsigen.2017.01.016>
- Home—10x Genomics (n.d.). <https://www.10xgenomics.com/>. Accessed 7 June 2018
- Hyde ER, Haarmann DP, Lynne AM, Bucheli SR, Petrosino JF (2013) The living dead: bacterial community structure of a cadaver at the onset and end of the bloat stage of decomposition. *PLoS One* 8(10):e77733. <https://doi.org/10.1371/journal.pone.0077733>
- Iacovacci G, D'Atanasio E, Marini O, Coppa A, Sellitto D, Trombetta B et al (2017) Forensic data and microvariant sequence characterization of 27 Y-STR loci analyzed in four Eastern African countries. *Forensic Sci Int Genet* 27:123–131. <https://doi.org/10.1016/J.FSIGEN.2016.12.015>
- International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431(7011):931–945. <https://doi.org/10.1038/nature03001>
- Irwin JA, Parson W, Coble MD, Just RS (2011a) mtGenome reference population databases and the future of forensic mtDNA analysis. *Forensic Sci Int Genet* 5(3):222–225. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-79953867044&partnerID=40&md5=0c14688affdaf9198aad1f6d2a783bc6>
- Irwin J, Just R, Scheible M, Loreille O (2011b) Assessing the potential of next generation sequencing technologies for missing persons identification efforts. *Forensic Sci Int Genet Suppl Ser* 3(1):e447–e448
- Ishiya K, Ueda S (2017) MitoSuite: a graphical tool for human mitochondrial genome profiling in massive parallel sequencing. *PeerJ* 5(5):e3406. <https://doi.org/10.7717/peerj.3406>
- Jäger AC, Alvarez ML, Davis CP, Guzmán E, Han Y, Way L et al (2017) Developmental validation of the MiSeq FGx forensic genomics system for targeted next generation sequencing in forensic DNA casework and database laboratories. *Forensic Sci Int Genet* 28:52–70. <https://doi.org/10.1016/j.fsigen.2017.01.011>

- Javan GT, Finley SJ, Smith T, Miller J, Wilkinson JE (2017) Cadaver thanatomicrobiome signatures: the ubiquitous nature of Clostridium species in human decomposition. *Front Microbiol* 8. <https://doi.org/10.3389/fmicb.2017.02096>
- Jobling MA, Gill P (2004) Encoded evidence: DNA in forensic analysis. *Nat Rev Genet* 5 (10):739–751. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-5044225199&partnerID=40>
- Just RS, Moreno LI, Smerick JB, Irwin JA (2017) Performance and concordance of the ForenSeq™ system for autosomal and Y chromosome short tandem repeat sequencing of reference-type specimens. *Forensic Sci Int Genet* 28:1–9. <https://doi.org/10.1016/j.fsigen.2017.01.001>
- Kader F, Ghai M (2015) DNA methylation and application in forensic sciences. *Forensic Sci Int* 249:255. <https://doi.org/10.1016/j.forsciint.2015.01.037>
- Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18(5):830–838. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-43149087535&partnerID=40&md5=1b7dc85ee8aeea4dbbc309c07ac69a5d>
- Kayser M, de Knijff P (2011) Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 12(3):179–192
- King JL, Sajantila A, Budowle B (2014) MitoSAVE: mitochondrial sequence analysis of variants in excel. *Forensic Sci Int Genet* 12:122–125. <https://doi.org/10.1016/j.fsigen.2014.05.013>
- Kwon SY, Lee HY, Kim EH, Lee EY, Shin K-J (2016) Investigation into the sequence structure of 23 Y chromosomal STR loci using massively parallel sequencing. *Forensic Sci Int Genet* 25:132–141. <https://doi.org/10.1016/J.FSIGEN.2016.08.010>
- Larmuseau MHD, Van Geystelen A, Kayser M, van Oven M, Decorte R (2015) Towards a consensus Y-chromosomal phylogeny and Y-SNP set in forensics in the next-generation sequencing era. *Forensic Sci Int Genet* 15:39–42. <https://doi.org/10.1016/j.fsigen.2014.11.012>
- Larmuseau MHD, Otten GPPL, Decorte R, Van Damme P, Moisse M (2017) Defining Y-SNP variation among the Flemish population (Western Europe) by full genome sequencing. *Forensic Sci Int Genet* 31:e12–e16. <https://doi.org/10.1016/j.fsigen.2017.10.008>
- Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75(5):843–854. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8252621>
- Li L, Ge J, Zhang S, Guo J, Zhao S, Li C et al (2012) Maternity exclusion with a very high autosomal STRs kinship index. *Int J Legal Med* 126(4):645–648. <https://doi.org/10.1007/s00414-012-0668-8>
- Liu L, Li Y, Li S et al (2012) Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012:251364. <https://doi.org/10.1155/2012/251364>
- Li L, Wang Y, Yang S, Xia M, Yang Y, Wang J et al (2017) Genome-wide screening for highly discriminative SNPs for personal identification and their assessment in world populations. *Forensic Sci Int Genet* 28:118–127. <https://doi.org/10.1016/j.fsigen.2017.02.005>
- Loomis EW, Eid JS, Peluso P, Yin J, Hickey L, Rank D et al (2013) Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res* 23 (1):121–128. <https://doi.org/10.1101/gr.141705.112>
- Lv S, Ji L, Chen B, Liu S, Lei C, Liu X et al (2017) Histone methyltransferase KMT2D sustains prostate carcinogenesis and metastasis via epigenetically activating LIFR and KLF4. *Oncogene* 37:1–15. <https://doi.org/10.1038/s41388-017-0026-x>
- Marshall C, Sturk-Andreaggi K, Daniels-Higginbotham J, Oliver RS, Barritt-Ross S, McMahon TP (2017) Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples: implications for the use of a positive control in a next-generation sequencing procedure. *Forensic Sci Int Genet* 31:198–206. <https://doi.org/10.1016/J.FSIGEN.2017.09.001>
- Massaia A, Xue Y (2017) Human Y chromosome copy number variation in the next generation sequencing era and beyond. *Hum Genet* 136:591. <https://doi.org/10.1007/s00439-017-1788-5>

- Mayes C, Seashols-Williams S, Hughes-Stamm S (2018) A capillary electrophoresis method for identifying forensically relevant body fluids using miRNAs. *Legal Med* 30:1. <https://doi.org/10.1016/j.legalmed.2017.10.013>
- McNevin D, Wilson-Wilde L, Robertson J, Kyd J, Lennard C (2005) Short tandem repeat (STR) genotyping of keratinised hair: Part 2. An optimised genomic DNA extraction procedure reveals donor dependence of STR profiles. *Forensic Sci Int* 153(2–3):247–259. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-24044545493&partnerID=40>
- Mendez FL, Krahn T, Schrack B, Krahn AM, Veeramah KR, Woerner AE et al (2013) An African American paternal lineage adds an extremely ancient root to the human Y chromosome phylogenetic tree. *Am J Hum Genet* 92(3):454–459. <https://doi.org/10.1016/j.ajhg.2013.02.002>
- Metcalf JL, Xu ZZ, Bouslimani A, Dorrestein P, Carter DO, Knight R (2017) Microbiome tools for forensic science. *Trends Biotechnol* 35:814. <https://doi.org/10.1016/j.tibtech.2017.03.006>
- Metzker ML (2009) Sequencing technologies—the next generation. *Nat Rev Genet* 11(1):31–46
- Miller AD, Good RT, Coleman RA, Lancaster ML, Weeks AR (2013) Microsatellite loci and the complete mitochondrial DNA sequence characterized through next generation sequencing and de novo genome assembly for the critically endangered orange-bellied parrot, *Neophema chrysogaster*. *Mol Biol Rep* 40(1):35–42
- Montano EA, Bush JM, Garver AM, Larijani MM, Wiechman SM, Baker CH et al (2018) Optimization of the Promega PowerSeq™ Auto/Y system for efficient integration within a forensic DNA laboratory. *Forensic Sci Int Genet* 32:26–32. <https://doi.org/10.1016/J.FSigen.2017.10.002>
- Novroski NMM, King JL, Churchill JD, Seah LH, Budowle B (2016) Characterization of genetic sequence variation of 58 STR loci in four major population groups. *Forensic Sci Int Genet* 25:214–226. <https://doi.org/10.1016/J.FSigen.2016.09.007>
- Ochiai E, Minaguchi K, Nambiar P, Kakimoto Y, Satoh F, Nakatome M et al (2016) Evaluation of Y chromosomal SNP haplogrouping in the HID-Ion AmpliSeq™ Identity Panel. *Legal Med* 22:58–61. <https://doi.org/10.1016/J.LEGALMED.2016.08.001>
- Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L et al (2013) Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). *Forensic Sci Int Genet* 7(5):543–549
- Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P et al (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. *Forensic Sci Int Genet* 22:54–63. <https://doi.org/10.1016/j.fsigen.2016.01.009>
- Pechal JL, Crippen TL, Benbow ME, Tarone AM, Dowd S, Tomberlin JK (2014) The potential use of bacterial community succession in forensics as described by high throughput metagenomic sequencing. *Int J Legal Med* 128(1):193–205. <https://doi.org/10.1007/s00414-013-0872-1>
- Peng Y, Croce CM (2016) The role of MicroRNAs in human cancer. *Signal Transduct Target Ther* 1(1):15004. <https://doi.org/10.1038/sigtrans.2015.4>
- Products (n.d.). <https://nanoporetech.com/products>. Accessed 7 June 2018
- Products | Single Cell Analysis System | Cyto-Mine® (n.d.). <https://www.spherefluidics.com/products/integrated-systems/cyto-mine/>. Accessed 7 June 2018
- Rathbun MM, McElhoe JA, Parson W, Holland MM (2017) Considering DNA damage when interpreting mtDNA heteroplasmy in deep sequencing data. *Forensic Sci Int Genet* 26:1–11. <https://doi.org/10.1016/j.fsigen.2016.09.008>
- Rockenbauer E, Hansen S, Mikkelsen M, Børsting C, Morling N (2014) Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing. *Forensic Sci Int Genet* 8(1):68–72
- Sanger F et al (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci* 74:5463–5467
- Sanchez-Cespedes M, Cairns P, Jen J, Sidransky D (1998) Degenerate oligonucleotide-primed PCR (DOP-PCR): evaluation of its reliability for screening of genetic alterations in neoplasia. *BioTechniques* 25(6):1036–1038. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9863058>

- Sauer E, Extra A, Cachée P, Courts C (2017) Identification of organ tissue types and skin from forensic samples by microRNA expression analysis. *Forensic Sci Int Genet* 28:99–110. <https://doi.org/10.1016/j.fsigen.2017.02.002>
- Schmedes SE, Woerner AE, Novroski NMM, Wendt FR, King JL, Stephens KM, Budowle B (2018) Targeted sequencing of clade-specific markers from skin microbiomes for forensic human identification. *Forensic Sci Int Genet* 32:50–61. <https://doi.org/10.1016/j.fsigen.2017.10.004>
- Seo SB, King JL, Warshauer DH, Davis CP, Ge J, Budowle B (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. *Int J Legal Med* 127(6):1079–1086. <https://doi.org/10.1007/s00414-013-0879-7>
- Sequel System—PacBio (n.d.). <https://www.pacb.com/products-and-services/pacbio-systems/sequel>. Accessed 7 June 2018
- Sequencing Platforms | Compare NGS Platforms (Benchtop, Production-Scale) (n.d.). <https://emea.illumina.com/systems/sequencing-platforms.html>. Accessed 7 June 2018
- Silva DSBS, Sawitzki FR, Scheible MKR, Bailey SF, Alho CS, Faith SA (2018) Genetic analysis of Southern Brazil subjects using the PowerSeq™ AUTO/Y system for short tandem repeat sequencing. *Forensic Sci Int Genet* 33:129–135. <https://doi.org/10.1016/j.fsigen.2017.12.008>
- Sirker M, Fimmers R, Schneider PM, Gomes I (2017) Evaluating the forensic application of 19 target microRNAs as biomarkers in body fluid and tissue identification. *Forensic Sci Int Genet* 27:41–49. <https://doi.org/10.1016/j.fsigen.2016.11.012>
- Sturk-Andreaggi K, Peck MA, Boysen C, Dekker P, McMahon TP, Marshall CK (2017) AQME: a forensic mitochondrial DNA analysis tool for next-generation sequencing data. *Forensic Sci Int Genet* 31:189–197. <https://doi.org/10.1016/j.fsigen.2017.09.010>
- Tridico SR, Murray DC, Addison J, Kirkbride KP, Bunce M (2014) Metagenomic analyses of bacteria on human hairs: a qualitative assessment for applications in forensic science. *Investig Genet* 5(1):16. <https://doi.org/10.1186/s13323-014-0016-5>
- Tsai CW, Yang CH, Chou SL, Cheng SG, Pai CY (2013) Non-CODIS DNA markers could be more effective than CODIS-based STRs in problematic biological relationship cases. *Rom J Leg Med* 21(3):245–248. <https://doi.org/10.4323/rjlm.2013.245>
- Ursell LK, Metcalf JL, Parfrey LW, Knight R (2012) Defining the human microbiome. *Nutr Rev* 70 (Suppl 1):S38. <https://doi.org/10.1111/j.1753-4887.2012.00493.x>
- Van Geystelen A, Decorte R, Larmuseau MHD (2013) AMY-tree: an algorithm to use whole genome SNP calling for Y chromosomal phylogenetic applications. *BMC Genomics* 14(1):101. <https://doi.org/10.1186/1471-2164-14-101>
- Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D (2012) Forensic STR analysis using massive parallel sequencing. *Forensic Sci Int Genet* 6(6):810–818
- Vidaki A, Kayser M (2017) From forensic epigenetics to forensic epigenomics: broadening DNA investigative intelligence. *Genome Biol* 18(1):238. <https://doi.org/10.1186/s13059-017-1373-1>
- Vidaki A, Díez López C, Carnero-Montoro E, Ralf A, Ward K, Spector T et al (2017) Epigenetic discrimination of identical twins from blood under the forensic scenario. *Forensic Sci Int Genet* 31:67–80. <https://doi.org/10.1016/j.fsigen.2017.07.014>
- Wang Z, Zhang S, Bian Y, Li C (2015) Differentiating between monozygotic twins in forensics through next generation mtGenome sequencing. *Forensic Sci Int Genet Suppl Ser 5*:e58–e59. <https://doi.org/10.1016/J.FSIGSS.2015.09.023>
- Wang Z, Zhou D, Wang H, Jia Z, Liu J, Qian X et al (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler™ NGS STR Panel and the Ion PGM™ System. *Forensic Sci Int Genet* 31:126–134. <https://doi.org/10.1016/J.FSIGEN.2017.09.004>
- Warshauer DH, Davis C, Holt C, Han Y, Walichiewicz P, Richardson T et al (2015a) Massively parallel sequencing of forensically relevant single nucleotide polymorphisms using TruSeq™ forensic amplicon. *Int J Legal Med* 129(1):31–36. <https://doi.org/10.1007/s00414-014-1108-8>
- Warshauer DH, Churchill JD, Novroski N, King JL, Budowle B (2015b) Novel Y-chromosome short tandem repeat variants detected through the use of massively parallel sequencing. *Genomics Proteomics Bioinformatics* 13(4):250–257. <https://doi.org/10.1016/J.GPB.2015.08.001>

- Williams GA (2012) Identification and resolution of capability gaps in forensic science. Doctoral thesis, University of Huddersfield. (Deposit: <http://eprints.hud.ac.uk/id/eprint/17500/>)
- Willems T, Gymrek M, Poznik GD, Tyler-Smith C, Erlich Y (2016) Population-scale sequencing data enable precise estimates of Y-STR mutation rates. *Am J Hum Genet* 98(5):919–933. <https://doi.org/10.1016/j.ajhg.2016.04.001>
- Wong L-P, Ong RT-H, Poh W-T, Liu X, Chen P, Li R et al (2013) Deep whole-genome sequencing of 100 Southeast Asian Malays. *Am J Hum Genet* 92(1):52–66. <https://doi.org/10.1016/j.ajhg.2012.12.005>
- Xavier C, Parson W (2017) Evaluation of the Illumina ForenSeq™ DNA Signature Prep Kit—MPS forensic application for the MiSeq FGx™ benchtop sequencer. *Forensic Sci Int Genet* 28:188–194. <https://doi.org/10.1016/j.fsigen.2017.02.018>
- Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. *Genomics Proteomics Bioinformatics* 12(5):190–197. <https://doi.org/10.1016/j.gpb.2014.09.001>
- Zaragoza MV, Fass J, Diegoli M, Lin D, Arbustini E (2010) Mitochondrial DNA variant discovery and evaluation in human cardiomyopathies through next generation sequencing. *PLoS One* 5(8). <https://doi.org/10.1371/journal.pone.0012295>
- Zeng X, King JL, Stoljarova M, Warshauer DH, LaRue BL, Sajantila A et al (2015) High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. *Forensic Sci Int Genet* 16:38–47. <https://doi.org/10.1016/j.fsigen.2014.11.022>
- Zhou Y, Guo F, Yu J, Liu F, Zhao J, Shen H et al (2016) Strategies for complete mitochondrial genome sequencing on Ion Torrent PGM™ platform in forensic sciences. *Forensic Sci Int Genet* 22:11–21. <https://doi.org/10.1016/J.FSIGEN.2016.01.004>