ISSN: 0976-3031

Available Online at http://www.recentscientific.com

CODEN: IJRSFP (USA)

International Journal of Recent Scientific Research
Vol. 9, Issue, 3(I), pp. 25280-25287, March, 2018

International Journal of Recent Scientific Research

DOI: 10.24327/IJRSR

Research Article

NEXT GENERATION SEQUENCING IN FORENSIC SCIENCE

Sabir Ali[1]., Manish Yadav[1]., Ajai Singh*[1]., Salma Siddiqui[1]., Vaishnavi Pandey[1]., Anoop Kumar Verma[2] and Archana Raikwar[3]

[1]Department of Orthopaedic Surgery, King George's Medical University, Lucknow, Uttar Pradesh, India- 226 018
[2]Department of Forensic Medicine, King George's Medical University, Lucknow, Uttar Pradesh, India- 226 003
[3]Model Rural Health Research Unit, Ghatampur, Kanpur Nagar, Uttar Pradesh, India – 209206

DOI: http://dx.doi.org/10.24327/ijrsr.2018.0903.1831

ARTICLE INFO

Article History:
Received 15th December, 2017
Received in revised form 25th January, 2018
Accepted 23rd February, 2018
Published online 28th March, 2018

Key Words:
Next Generation Sequencing, Sequencing, Forensic Science.

ABSTRACT

Sequencing has made huge advancements and improvement in the field of genetics in the last few decades. One of the most important achievements of Next Generation Sequencing (NGS) is to produce millions of sequence reads in a short period of time and to produce large sequences of DNA in fragments of any size. Libraries can be produced from whole genomes or any DNA or RNA region of interest without the need to know its sequence beforehand. This allows for looking for variations and facilitating genetic identification. In this review, our objective was to do a profound analysis of existing NGS technologies and their application in forensics as well as a discussion about the advantages and drawbacks of these technologies in genetic identification. In addition to the degraded state of the sample, a major challenge is the limited amount of available sample in the forensic field. If the amount of DNA input required for preparing NGS libraries continues to decrease, nearly any sample could be sequenced; therefore, the maximum information from any biological remains could be obtained. Furthermore, microbiome typification could be a fascinating application to study for crime scene characterization. NGS technologies are going to be crucial for DNA human typing in cases like mass disasters or other happenings where forensic specimens and samples are compromised and degraded. With the use of NGS, it will be possible to achieve the simultaneous analysis of the standard autosomal DNA (STRs and SNPs), mitochondrial DNA, and X and Y chromosomal markers.

INTRODUCTION

In the area of sequencing in recent years, the molecular biology techniques have rapidly advanced and keep on moving ahead at a staggering speed (Figure-1). Many of the improvements within these techniques led to Next Generation Sequencing (NGS) development. All these NGS techniques shared the common mechanism and that is the massively parallel sequencing strategy, where thousands of reads are generated at the same time (Graham, 2012). The core innovations of NGS platforms are massive parallel chemical reactions, ultrahigh-resolution optics and computational methods to analyses very short read. These revolutionary technological advances have drastically reduced the sequencing cost and shortened the turnaround time to merely a few days. Nowadays, science is facing a challenge in relation to the amount of information derived from these technologies, and the problem of millions of misleading results (Wong, 2013). However, these technologies have the potential to launch a revolution of DNA sequencing along with its applications (sequencing single RNA or even protein molecules directly) and the challenge of implementing data storage with the processing (van Dijk *et al*., 2014; Yang *et al*., 2014). Currently, NGS methodology is mainly applied in biomedical fields, which has made it possible to improve research in the clinical diagnosis of cancer, genetically heterogeneous disorders and common and even rare diseases. In forensic science, NGS has been applied to whole mitochondrial sequencing.

Corresponding author:* **Ajai Singh
Department of Orthopaedic Surgery, King George's Medical University, Lucknow, Uttar Pradesh, India- 226 018

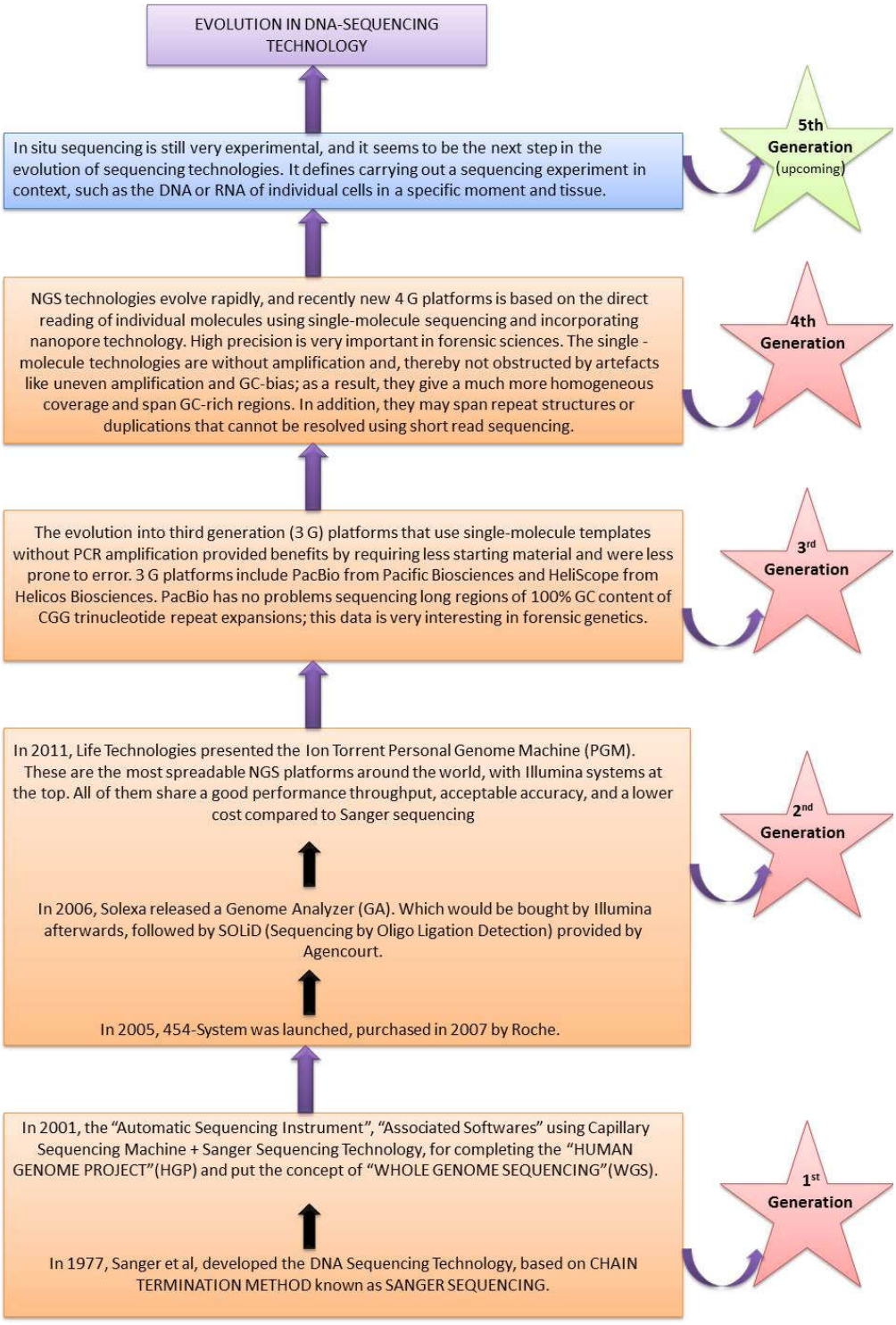

Figure 1 Evolutionary development of DNA Sequencing methodology [Sanger et al. (1977), Liu et al. (2012), Gut. (2013), Loomis et al. (2013), Buermans & den Dunnen. (2014) and Ku & Roukos. (2013)].

However, in the cases where a low quantity of DNA is one of the most important limiting factors, NGS has also obtained great advances, such as in Non-invasive Pre-natal Diagnosis and forensic casework. It has long been known that DNA of the foetus can be found in maternal blood, yet it is generally low in quantity and quality and it is not easy to discriminate foetal DNA from maternal blood. The enormous power of NGS is used for non-invasive pre-natal testing and to detect trisomies (Buermans & den Dunnen, 2014).

However, the progressions in Pre-natal Diagnosis are not comparable with forensic genetics. As there are not enough data about NGS and the quality of its results in the forensics area. Therefore multicentric studies are required about some ancient DNA reviews where the improvements of DNA extraction methods and short reads from NGS output have helped to increase the accessibility to this aDNA and its study (Hofreiter *et al*., 2015).These studies also consider enriching this aDNA, depending on each particular study, mentioning

custom project-specific SNP panels, as has been done for forensic aims. However, the limited amount of sample and its elevated grade of degradation is still the main obstacle for the NGS platforms (Figure-2) in the forensic field application.

Next Generation Sequencing In Forensic Science

It is not challenging to imagine the potential uses of NGS in forensic sciences while avoiding the main technical restrictions of current technologies such as- capillary electrophoresis and the high degraded or limited samples of forensic specimens (Irwin *et al*., 2011a; Miller *et al*., 2013), as the application of NGS technologies in 'de novo' characterization of entire genomes (such as the entire mitochondrial DNA sequence of Neophemachrysogaster) and in re-sequenciation of human genomes (Human genome project consisting in 1092 genomes of individuals from 14 populations, constructed using a combination of low-coverage whole-genome and exome sequencing) has been comprehensively established.

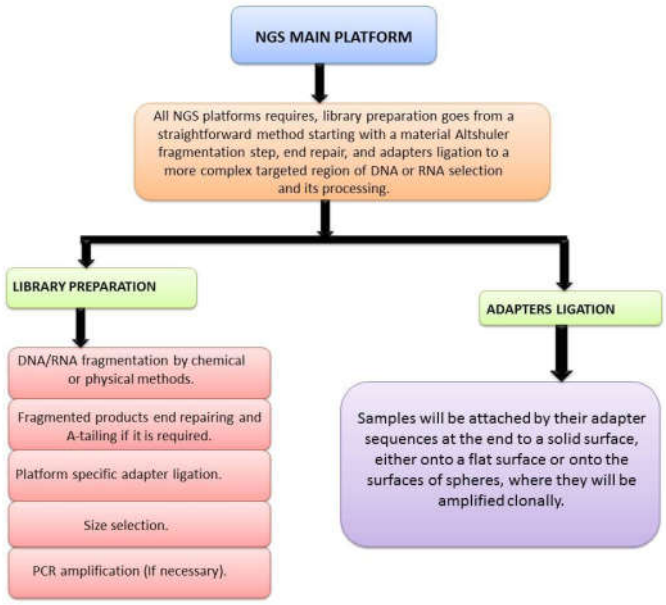

Figure 2 Basic methodology involved in NGS Platform [Sanger *et al*. (1977), Liu *et al*. (2012), Gut. (2013), Loomis *et al*. (2013) and Ku & Roukos. (2013)]

The recent advancement includes NGS data produced from human hair, obtaining relevant results in metagenomic analyses. Forty-two DNA extracts obtained from the human scalp and pubic hairs generated a total of 79 766 reads, yielding 39 814 reads post-control and abundance filtering. As the pyrosequencers were the only platforms with sufficient read length to sequence the core STR loci used in forensic genetics and to date most of the forensic literature with NGS STR data have been produced with pyrosequencing technology. Therefore this could provide relevant data for victims of sexual assault and offenders when other associative evidence is absent (Tridico *et al*., 2014).Indels and CNVs have not attracted much attention, even though repeats cover almost half, and STRs only 15%, of the human genome. In a few reports, sequencing by synthesis and recently semiconductor sequencing have been used. The libraries were constructed directly by PCR or by adapter ligation. The advantages and disadvantages of NGS is summarizes in **Table -1.**

Short Tandem Repeats and Single Nucleotide Polymorphisms

Nowadays short tandem repeats (STRs) typing DNA profiling is the standard method for forensic analysis and until now DNA sequencing has been employed to the hypervariable region of mitochondrial DNA (mtDNA) to establish the origin and relationship between DNA samples available in low quantities (Berglund *et al*., 2011).STR analysis has become a well-established technology for human forensic casework (Irwin *et al*., 2011b; Rockenbauer *et al*., 2014). Several STR database systems have been created, including the Combined DNA Index System (CODIS) utilized in the US. This system currently uses a standard set of 13 STR loci, which are highly polymorphic, genetically unlinked and reside in non-coding regions (Budowle *et al*.,2001). These STR alleles are routinely analysed by multiplexed PCR followed by capillary electrophoresis (CE)-based separation(Jobling & Gill, 2004)., The process to generate the DNA profiles are typically conducted in 8–10 hours by using the traditional laboratory methods, Over recent years, significant efforts have been made to enable the rapid generation of STR profiles to include rapid PCR, direct PCR, microfluidic chip-based amplification and detection and integrated microfluidic devices for rapid sample to profile processing, as alternative approaches. Generating DNA profiles in an automated fashion with the use of integrated devices has future applications at booking stations and in other field locations such as airports, border crossing or in kinship determination and disaster victim investigation (Romsos & Vallone, 2015). Although standard STR typing provides sufficient discrimination power for most applications, it has been proved that the use of only the routine STRs can lead to errors in kinship analysis (Li *et al*., 2012; Tsai *et al*., 2013). Moreover, the output data stutter products from alleles from a secondary participant in mixtures cannot be discriminated (Butler *et al*.,2004; McNevin *et al*., 2005) and dye artefacts are commonly found (Zeng *et al*., 2015). The current challenges are to use autosomal Single Nucleotide Polymorphisms (SNPs) for human individual identification in forensic identification. Bi-allelic SNPs are less polymorphic than multi-allelic STRs. SNPs are less informative in the analysis of mixtures of DNA from multiple individuals, although for this effect (Kayser & de Knijff, 2011), the use of a higher number of SNPs relative to STRs united with multiple genotyping technologies may recompense. The application of large-scale sequencing provides multi-dimensional improvements to forensic analyses. Out of these, the major one is to complement STRs with more informative markers obtained from large-scale sequencing (Berglund *et al*., 2011). while compared with Sanger capillary sequencing at a considerably lower cost per nucleotide, the NGS technologies confront entirely new approaches for DNA sequence data generation, offering the possibility to sequence up to millions of individual DNA strands (like in DNA mixtures) and higher throughput sequencing.

Using NGS to sequence STR amplicons will make it more discriminative as a genetic marker for individuals, and complete profiles could be generated with an input of just 62 pg (Zeng *et al*.,2015). It has been proved that there are nucleotide substitutions, deletions or insertions in the flanking region or in the repeat unit of STRs not detectable by CE-analysis (Rockenbauer *et al*., 2014).

Table 1 Summarized advantages and disadvantages of NGS. [Jobling & Gill, 2004]

MARKER			
Autosomal STRs	**Autosomal SNPs**	**Y Chromosomal STRs**	**Mitochondrial SNPs**
Advantages And Disadvantages In Forensics			
▪ High discrimination power ▪ Low mutation rates. ▪ Easy inclusion in databases.	▪ Less discrimination power than STRs, but very useful in compromised samples.	▪ Low discrimination power. ▪ Sharing within patrilline. ▪ Possible population structure problems.	▪ High copy number. ▪ Good survivalin old/damaged samples. ▪ Low discrimination power.
Advantages of Next Generation Sequencing			
▪ Detection in mixed samples and easy to calculate proportions (% reads).	▪ Millions of SNPs can be detected (lineage, bioancestry and phenotypic markers). ▪ Detectable in mixed samples with high reads. ▪ Specific kits for their study.	▪ Detection in mixed samples and easy to calculate proportions (% read).	▪ Easy analysis of complete mitochondrial genome. ▪ Many kits available.
Disadvantages of Next Generation Sequencing			
▪ Much more expensive. ▪ Difficult to use multi-allelic STRs due to the read length. ▪ Enormous volume of data, complex statistical treatment and storage.	▪ More expensive, but depending on the number of SNPs, price can be equal.	▪ Much more expensive. ▪ Difficult to use multi-allelic STRs due to the read length. ▪ High rates of repeat sequences complicates analysis. ▪ Enormous volume of data, complex statistical treatment and storage.	

NGS technologies identify both the number of repeats and the sequence polymorphisms present in the STR core unit. Furthermore, NGS can generate individual sequences of the alleles present in an STR amplicon mixture (VanNeste *et al*., 2012). More identifiable alleles mean more statistical power of the STR investigation and essentially reduce the necessary number of loci that need to be typed to solve a case to a certain level. It may also be easier to resolve DNA mixtures in crime case investigations, when alleles that appear identical in CE-analyses may be further characterized by NGS. About half of the STRs selected for the European Standard Set and the CODIS set are simple STR repeats, with only one repeat unit and little sequence variation, except for the variation in a number of repeats. The possibility to distinguish between individuals with identical allele lengths as far as the characterization of mutation events could be essential in forensic or kinship analysis. These data will increase the statistical power with a reduced number of SNPs or STRs (Rockenbauer *et al*., 2014).

Currently, Thermo Fisher has designed a human identification STR (HID-STR) 10-plex panel (amelogenin, CSF1PO,D16S539, D3S1358, D5S818, D7S820, D8S1179, TH01, TPOX, and vWA) where the primers have been designed specifically for the purpose of SGS and the data analysis is supported by Ion Torrent TM software. The experiment runs with success and full profiles were obtained for all biological samples from real crime and identification cases, in which only partial profiles were previously obtained with capillary electrophoresis assays. However, still, the full power of NGS can only be exploited by also including new complex STRs, as some problems with the presence of stutters are still present and should be solved with improved software (Fordyce *et al*., 2015).

Therefore, NGS provides an opportunity for creating an all-in-one multiplex with relevant forensic markers that include, for example, STRs, SNPs, InDels and mtDNA markers (Irwin *et al*., 2011a). In recent times, certain companies have developed panels focused on the forensic field using the same markers applied until now, helping researchers to process their samples instead of designing their own panels and primers. SNPs are identified by some genome-wide association studies have that can predict ethnic background and appearance traits, this approach is known as forensic DNA phenotyping. This information is very useful in criminal case resolutions, as it can help identify an individual, which will later be confirmed with a reference sample. On the other hand, the ability to estimate individual-specific appearance via DNA predicted facial morphology is an important goal towards identifying unknown persons (Irwin *et al*., 2011a; Rockenbauer *et al*., 2014).

If such predictions turn out to be possible, it is unlikely to be achievable with small sets of genetic markers. Therefore, new technologies, such as NGS, are needed for the parallel genotyping of large numbers of SNPs, which can successfully deal with small amounts of degraded DNA (Kayser & de Knijff, 2011). A modified protocol for the TruSeqTM ChIP protocol has recently been developed to detect 160 human identification, ancestry and phenotypic SNPs, called TruSeq TM Forensic Amplicon (Warshauer *et al*., 2015) and, massively parallel sequencing (MPS) has been described as a good method for the discovery of forensic SNP markers (Seo *et al*.,2013).This method is less labour intensive than other techniques. Furthermore, library preparation requires low input DNA (1 ng) (Warshauer *et al*., 2015). A genome-wide SNP array that covers 906 600 SNPs of autosomal X and Y chromosome and mitochondrial DNA (Bridges *et al*., 2011) was developed by Affymetrix. Recently, NGS has been used to analyze nuclear microsatellite markers, where it has been

shown that random sequencing of a small fraction of a genome can result in a high density of potential microsatellite loci with low cost and rapid processing times. The major advancement offered by NGS is the ability to produce an enormous volume of data cheaply (Metzker, 2009).

The frequent challenges include it is not possible to analyze NGS data manually or even a manual genotype calling and one of the major challenges will be to develop a forensic NGS tool for analysis of the sequence data. The software application is completely reliable and validated before they can be applied in real casework. Illumina and Thermo Fisher Scientific have developed software solutions for analysis of their kits and even the software solution should also be able to identify two STR alleles of the same size but with different sequences, including it should be able to name the alleles according to the nomenclature suggested by the ISFG (Børsting & Morling, 2015). However, they are not sufficiently sophisticated for forensic genetics.

Lineage Markers in Forensic Science

The mitochondrial DNA (mtDNA) is present in greater amounts in the cell than nuclear DNA (nDNA) and thus, the likelihood of recovering useable DNA data is increased in forensic samples that fail to yield useful and DNA typing results (Parson *et al*., 2013). The mtDNA also provides information on maternal bio-geographic ancestry, which is mainly obtained from the hyper variable region (Kayser & de Knijff, 2011). Historically the so-called hyper variable segment 1 (HVS-1) of the mitochondrial control region, the only larger non-coding region in the mt Genome, was targeted providing a random match probability (RMP) of roughly one in 30individuals. With the entire mtDNA control region, the RMP augmented to one in 120 individuals and has provided useful evidence in many cases. The extension to analyze the entirement Genome is a logical consequence and desirable goal to maximize the information content of mtDNA analyses (Irwin *et al*., 2011b). Sequencing the entire mitochondrial genome from samples with the same control region sequences will differentiate samples by the analysis of private polymorphisms in the coding region not studied by Sanger sequencing (Holland *et al*., 2011). With mtDNA, NGS allows the entire mtDNA genome sequences to be obtained when only a small portion of the genome is sequenced and at a fraction of the cost of traditional approaches, requiring only small amounts of samples (Miller *et al*., 2013). In the case of mtDNA, Sanger capillary sequencing of the whole mtDNA was, until recently, the only method to detect variants. However, this technology employs an expensive and time-consuming technique (Davis *et al*., 2015; Parson & Bandelt, 2007). Another important limitation in this type of analysis is the availability of sufficient evidentiary material. The updated advances in microfluidics, digital imaging systems, and bioinformatics have improved these limitations of NGS technologies. Furthermore, the massive parallel sequencing enabled by NGS is revolutionizing genetic data generation and has the potential to generate entire mtDNA genome profiles from even highly degraded specimens. It is relatively straight-forward and cost-effective. The application of mtDNA coding region data in routine forensic casework will be dictated less by the quantity of the sample and/or the effort required to produce the data than by the availability of large high-quality entire mt Genome population databases that can be used to determine the rarity of mt Genome haplotypes.

The mtDNA and NGS have a significant amount of data in, because of the presence of hundreds to several thousand copies of this DNA per cell (present) in different tissues and this aspect produces a high percentage of mutations that can also be present in homoplasmic or heteroplasmic states. Although in different punctual mutations that could be useful for designing new haplo groups or personal identification can be found. Massive parallel sequencing DNA can be performed using various platforms, including454, SOLiD, Affymetrix re-sequencing chip, Illumina and Ion Torrent. Different platforms may also affect the depth of coverage and the ability to multiplex (Wong *et al*., 2013). Recent studies have proved that NGS results, such as the ones obtained by 454 platforms, are comparable to the detection of Sanger sequencing single nucleotide substitutions. High concordance (98%) was obtained for the detection of over 400 single nucleotide substitutions, including four heteroplasmic variants previously detected by Sanger sequencing (Zaragoza *et al*., 2010). Currently, there are various studies of mtDNA using NGS technologies. Out of which some have been performed with the Illumina GAIIx platform, as it facilitates with the generation of short reads, and its lower error rate compared to other technologies, as well as the production of a wide coverage read by a large number of reads generated (Zaragoza *et al*., 2010). In order to evaluate the performance and reliability of recent NGS, a comparative study on different biological sources and 64 mtGenomes from diverse phylogenetic backgrounds was performed via Ion Torrent Personal Genome Machine and Sanger-type sequencing as reference methods (Zaragoza *et al*., 2010).

The final NGS haplotypes already have a great impact on aligners, alignment parameters, and pre-alignment data filtering tools. Therefore, further development of alignment software would be desirable to facilitate the application of NGS in mtDNA forensic genetics. On the other hand, heteroplasmy analyzed through Sanger methods can be identified when the minor allele is present above 20%. However, heteroplasmy can be detected accurately in one per 10 000 mitochondrial genome copies with Illumina GAII (He *et al*., 2010). As a result, NGS technology has the potential to greatly assist in the analysis of whole mitochondrial sequences (Yang *et al*., 2014). Recently there are some NGS protocols applied; most of them to mtDNA, and some changes and steps to improve NGS data quality are being introduced. These steps include- Firstly, a preliminary DNA repair step and primer extension capture as methods of target enrichment with primers that allow a consistent read coverage of the entire mtGenome. Secondly, DNA repair and primer extension capture improve both the quantity and quality of NGS data in the forensic field. Mainly, the changes should be optimized for large quantities of the high quality of DNA to ancient DNA (Loreille *et al*., 2011). Finally, DNA libraries and barcoding improve the success of whole mtDNA sequencing from highly degraded samples in human identification (Templeton *et al*., 2013). Sanger capillary sequencing of mtDNA cannot be frequently used to sequence the regions of amplicons, downstream of lengths, and heteroplasmy homopolymers.

Previously in Y-chromosomes, many discoveries related to new Y-SNPs using NGS are unknown. These discoveries have a direct influence over Y-chromosomal phylogeny trees, revealing new categories that include not only a limited number of selected Y-SNPs assembled in a reduced Y-chromosome tree that captures the main Y-haplo-groups of worldwide significance, but also Y-SNPs and evolutionary lineages from all available (verified and non-verified) genetic data. However, further multicentric studies and researches are still required to establish the new Y-SNPs in forensic and anthropological sciences. Most of the current studies of Y-chromosomes use Y phylogeny and Y-SNPs defined by Karafet *et al*. (2008). With the development of NGS analysis, new evolutionary lineages have been explained and there is a need to update the Y-chromosomal tree (Larmuseau *et al*., 2015). Complete Genomics made the SNP calls of 35 whole genomes of males sequenced with a high sequencing coverage on the Complete Genomics Analysis (CGA) Platform (Drmanac *et al*., 2010). Later, the Personal Genome Project (PGP) and Singapore Sequencing Malay Project (SSMP) used the same platform. At the moment, 40 male genomes are available from the PGP and 46 Malays have been made publicly available from the SSMP (Wong *et al*., 2013). Ultimately, the 1000 Genomes Project wants to provide an extensive resource on human genetic variation by sequencing more than 1000 human genomes. In 2012, a set of 526 SNPs profiles were published as result of phase 1 of the Project (T. 1000 Genomes Project Consortium, 2012). All the information derived from these studies has been compiled and the Y-chromosomal phylogenetic tree has been updated for forensic purposes (Van Geystelen *et al*., 2013).

Oral Microbiome in Forensic Science

Oral microbiome analysis is another application for NGS. This analysis can reveal a microbiome profile in terms of species-level identification and relative abundance of micro-organisms. The microbiome can also reveal important information about an individual's recent activities and about geographic location (Wong *et al*., 2013). Accurately determining the time since death is an important aspect of forensic sciences and casework. New research indicates that this might be achieved by examining changes in the bacterial communities of the mouth that occur after death.

Necrobiome or Thanatomicrobiome in Forensic Science

Necrobiome/Thanatomicrobiome- necrobiome-the world of organisms involved in decomposition, including bacteria-could change the game for forensic scientists. "By knowing which microbes take over a dead body and how long it takes, forensic scientists might be able to use this technique to determine time of death or other aspects of a crime scene". The NGS is relatively a new tool used in forensic science, which complements the information of forensic entomology and the data collected from the physical appearance of cadavers (Hyde *et al*., 2013; Pechal *et al*., 2014). There are studies where a correlation between bacterial communities and evaluation of the carrion decomposition process has been shown to be a potentially useful post-mortem interval (PMI) estimator tool; in this way, the necrobiome could be used as a microbial clock indicator (Guo *et al*.,2016). A variant of this technique applied to Paleo forensic science would allow the deeper genomic evolution of DNA microbes, for example, both pathogens and microbiome of decomposition (Gorge *et al*., 2016).

Epigenetic Analysis in Forensic Science

Currently, it is depicted via a number of studies that epigenetic markers can also have various applications in forensic science. Most of the analyses have been proved with large amounts of DNA, but there are also good results with low amounts of starting DNA (100 pg), which is crucial for the forensic field (Yang *et al*., 2014). Illumina offers a broad portfolio of sequencing- and array-based epigenetic analysis tools that provide robust, simple-to-use, and cost-efficient solutions for studying epigenetic modifications and their impact on gene regulation. By working with leading epigenomics experts, Illumina ensures its solutions meet the field's rapidly evolving needs.

Microrna Analysis In Forensic Science

MicroRNAs (miRNAs) are non-protein coding molecules with important regulatory functions; many have tissue-specific expression patterns. In recent days, microRNA has been incorporated in the forensic field and their small size, makes them the best molecules to have resistance to degradation and tissue-specific or highly tissue-divergent expression; they are also suitable for forensic body fluid identification, species identification and post-mortem interval (PMI) inference analysis (Yang *et al*., 2014).

CONCLUSION AND FUTURE PROSPECTS

Nowadays, there are many important databases such as CODIS and the UK National DNA Database (NDNAD). As the applications of genetic technologies have been applied to the forensic field. The major problems in introducing NGS to the field of forensics is the generation of a new type of data and the need for specific training for researchers in all these areas. There are specific kits such as the Amp FLSTRVR IdentifilerVR Direct PCR Amplification Kit that enable direct amplification of the 16 loci included in the IdentifilerVR Kit, eliminating the need to perform DNA extraction or purification steps for DNA database and other single source reference samples. The major advance offered by NGS is the ability to produce an enormous volume of data cheaply (Zaragoza *et al*., 2010). In forensics, especially population genetics; recent advances have produced new types of genetic markers that can overcome some of the main limitations of current and previous DNA profiling methods (Kayser & de Knijff, 2011). NGS is going to be crucial for DNA human typing in cases like mass disasters or other events where forensic specimens and samples are compromised and degraded. With the use of NGS, it will be possible to join the simultaneous analysis of the standard autosomal DNA (STRs and SNPs), mitochondrial DNA and X and Y chromosomal markers (Irwin *et al*., 2011a). Moreover, the benefits of NGS technologies have to be studied against its high cost and the strict standards required for forensic genetics (Kayser & de Knijff, 2011) as well as the multi-centric standardization exercises needed before implementing these new techniques in real forensic casework. Till date, there is no formal forensic assessment of the data quality generated by these NGS technologies, as is commonly exercised with standard sequencing techniques (Bandelt & Salas, 2012).

Furthermore, for the NGS results and their forensic applications, it is important to focus on the need for specific bioinformatics analysis tools. MyFLq (My-Forensic-Loci-queries) is a specific bioinformatics framework developed to process data without clustering, extracting maximal information with automatically determined regions of interest and it can be run natively on Illumina Base Space or independently on a forensic laboratory's server (Van Neste *et al*.,2015). Moreover, Illumina MiSeq has been proven to be ready to analyze STR profiles (Van Neste *et al*., 2012). The main advantage of NGS in forensics includes the wide range of information that can be obtained in a short period of time. However, its high costs, as well as the extensive training required in these technologies and the management of the heavily pre-processed data are still obstacles to overcome. The software applications are not far enough advanced and more effort from the companies, the forensic researchers and scientists are necessary. Software development, growth, and validation of software solutions will be critical aspects in the introduction of NGS in forensic sciences. Another main problem is that SGSSTR typing has a lack of uniformity in workflow and data analysis amongst the forensic genetic laboratories, making an inter-laboratory comparison of STR typing results difficult (Fordyce *et al*., 2015).

References

1. Bandelt HJ, Salas A (2012). Current next generation sequencing technology may not meet forensic standards. *Forensic Sci Int Genet* 6:143-145.
2. Berglund EC, Kiialainen A, Syv€anen AC (2011). Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig Genet* 2:23.
3. Børsting C, Morling N (2015). Next generation sequencing and its applications in forensic genetics. *Forensic Sci Int Genet* 18:78–89.
4. Bridges M, Heron EA, O'Dushlaine C, Segurado R, Morris D, Corvin A, Gill M, Pinto C. (ISC) TISC (2011). Genetic classification of populations using supervised learning. *PLoS One* 6:e14802.
5. Budowle B, Masibay A, Anderson SJ, Barna C, Biega L, Brenneke S, Brown BL, *et al*. (2001). STR primer concordance study. *Forensic Sci Int* 124:47–54.
6. Buermans HPJ, den Dunnen JT (2014). Next generation sequencing technology: advances and applications. *Biochim Biophys Acta* 1842: 1932–1941.
7. 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* 25:1397–1412.
8. Davis C, Peters D, Warshauer D, King J, Budowle B (2015). Sequencing the hypervariable regions of human mitochondrial DNA using massively parallel sequencing: enhanced data acquisition for DNA samples encountered in forensic testing. *Leg Med* 17:123–127.
9. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, *et al*. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* 327:78–81.
10. Fordyce SL, Mogensen HS, Børsting C, Lagac_e RE, Chang C-W, Rajagopalan N, Morling N (2015). Second-generation sequencing of forensic STRs using the Ion TorrentTM HID STR 10-plex and the Ion PGMTM. *Forensic Sci Int Genet* 14:132–140.
11. Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56–65.
12. Gorg_e 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.
13. Graham AW (2012). Identification and resolution of capability gaps in forensic science. [University of Huddersfield]: *University of Huddersfield.*
14. Guo J, Fu X, Liao H, Hu Z, Long L, Yan W, Ding Y, *et al*. (2016). Potential use of bacterial community succession for estimating post-mortem interval as revealed by high-throughput sequencing. *Sci Rep* 6:24197.
15. Gut IG (2013). New sequencing technologies. *Clin Transl Oncol* 15:879–881.
16. He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE, Diaz LA, *et al*. (2010). Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* 464:610–614.
17. Hofreiter M, Paijmans JLA, Goodchild H, Speller CF, Barlow A, Fortes GG, Thomas JA, *et al*. (2015). The future of ancient DNA: technical advances and conceptual shifts. *BioEssays* 37:284–293.
18. Holland MM, McQuillan MR, O'Hanlon K. (2011). Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy. *Croat Med J* 52:299–313.
19. Hyde ER, Haarmann DP, Lynne AM, Bucheli SR, Petrosino JF (2013a). The living dead: bacterial community structure of a cadaver at the onset and end of the bloat stage of decomposition. Moreau CS, editor. *PLoS One* 8:e77733.
20. Irwin J, Just R, Scheible M, Loreille O (2011a). Assessing the potential of next generation sequencing technologies for missing persons identification efforts. *Forensic Sci Int Genet Suppl Ser* 3:e447–e448.
21. Irwin JA, Parson W, Coble MD, Just RS. (2011b). mtGenome reference population databases and the future of forensic mtDNA analysis. *Forensic Sci Int Genet* 5:222–225.
22. Jobling MA, Gill P (2004). Encoded evidence: DNA in forensic analysis. *Nat Rev Genet* 5:739–751.
23. 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* 18:830–838.
24. Kayser M, de Knijff P (2011). Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 12:179–192.
25. Ku CS, Roukos DH (2013). From next-generation sequencing to nanopore sequencing technology: paving the way to personalized genomic medicine. *Expert Rev Med Devices* 10:1–6.
26. 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.
27. Li L, Ge J, Zhang S, Guo J, Zhao S, Li C, Tang H, *et al*. (2012). Maternity exclusion with a very high autosomal STRs kinship index. *Int J Legal Med* 126:645–648.
28. Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, *et al*. (2012). Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012:251364.
29. Loomis EW, Eid JS, Peluso P, Yin J, Hickey L, Rank D, McCalmon S, *et al*. (2013). Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res* 23:121–128.
30. Loreille O, Koshinsky H, Fofanov VY, Irwin JA (2011). Application of next generation sequencing technologies to the identification of highly degraded unknown soldiers' remains. *Forensic Sci Int Genet Suppl Ser* 3:e540–e541.
31. 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* 153:247–259.
32. Metzker ML (2009). Sequencing technologies - the next generation. *Nat Rev Genet* 11:31–46.
33. 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* 40:35–42.
34. Parson W, Bandelt H-J (2007). Extended guidelines for mtDNA typing of population data in forensic science. *Forensic Sci Int Genet* 1:13–19.
35. Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Fendt L, *et al*. (2013). Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). *Forensic Sci Int Genet* 7:543–549.
36. 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* 128:193–205.
37. 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* 8:68–72.
38. Romsos EL, Vallone PM (2015). Rapid PCR of STR markers: applications to human identification. *Forensic Sci Int Genet* 18:90–99.
39. Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467.
40. 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* 127:1079–1086.
41. Templeton JEL, Brotherton PM, Llamas B, Soubrier J, Haak W, Cooper A, Austin JJ (2013). DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investig Genet* 4:26.
42. Thompson JF, Steinmann KE (2010). Single molecule sequencing with a HeliScope genetic analysis system. *Curr Protoc Mol Biol* 7:10.1–10.14.
43. 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* 5:16.
44. 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* 21:245–248.
45. Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C (2014). Ten years of nextgeneration sequencing technology. *Trends Genet* 30:418–426.
46. 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* 14:101.
47. Van Neste C, Gansemans Y, De Coninck D, Van Hoofstat D, Van Criekinge W, Deforce D, Van Nieuwerburgh F (2015). Forensic massively parallel sequencing data analysis tool: implementation of MyFLq as a standalone web- and illumina BaseSpace(VR)-application. *Forensic Sci Int Genet* 15:2–7.
48. Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D (2012). Forensic STR analysis using massive parallel sequencing. *Forensic Sci Int Genet* 6:810–818.
49. Warshauer D, Davis C, Holt C, Han Y, Walichiewicz P, Richardson T, Stephens K, *et al*. (2015). Massively parallel sequencing of forensically relevant single nucleotide polymorphisms using TruSeqTM forensic amplicon. *Int J Legal Med* 129:31–36.
50. Wong L-JC (2013). Next Generation Sequencing. Translation to Clinical Diagnostics. New York (NY): *Springer*.
51. Wong L-P, Ong RT-H, Poh W-T, Liu X, Chen P, Li R, Lam KK-Y, *et al*. (2013). Deep whole-genome sequencing of 100 Southeast Asian Malays. *Am J Hum Genet* 92:52–66.
52. Yang Y, Xie B, Yan J (2014). Application of next-generation sequencing technology in forensic science. *Genomics Proteomics Bioinformatics* 12:190–197.
53. 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:e12295.
54. Zeng X, King JL, Stoljarova M, Warshauer DH, LaRue BL, Sajantila A, Patel J, *et al*. (2015). High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. (Forensic Sci Int Genet) 16:38–47.

****