Advanced techniques in molecular genetics and its implications on genetic testing and screening in the Arabian peninsula

Khaled K. Abu-Amero, PhD, FRCPath, Altaf A. Kondkar, MSc, PhD.

ABSTRACT

Molecular diagnosis of human disorders is referred to as the detection of the various pathogenic mutations in DNA and/or RNA samples in order to facilitate detection, diagnosis, sub-classification, prognosis, and monitoring response to therapy. The use of molecular biology techniques to expand scientific knowledge of the natural history of diseases, identify people who are at risk for acquiring specific diseases, and diagnose human diseases at the nucleic acid level. Molecular diagnostics combines laboratory medicine with the knowledge and technology of molecular genetics and has been enormously revolutionized over the last decades, benefiting from the discoveries in the field of molecular biology. This review will discuss in details the recent advances in molecular diagnostics and how the Arabian Peninsula can benefit from those techniques knowing for a fact the high percentages of consanguineous marriages and the tribal nature of marriages which resulted in high incidence of genetic diseases.

The intent of this review was to compile and provide the readers with up to date information regarding recent advances in molecular genetics techniques and how these techniques could be utilized for the purpose of genetic testing of various hereditary diseases present in the kingdom.

Newer techniques. We will concentrate here on new advanced techniques as basic molecular genetics techniques were covered in our previous review published almost a decade ago. The newer molecular genetics techniques and some of their applications are described below.

Array comparative genomic hybridization. Chromosomal abnormalities form the molecular basis of many human genetic disorders. Therefore, cytologists use microscopic method of karyotyping, combined with fluorescence in situ hybridization (FISH), which can detect aberrations as small as 3-5 megabases (Mb). Array comparative genomic hybridization (CGH) was originally designed to detect DNA copy number variations (CNVs) across the entire genome in chromosomal disorders and cancer research. Instead of using metaphase chromosomes, this method utilizes solid support or glass slides, “chips” arrayed with small segments of DNA as the targets for analysis. These chips are fabricated by high-speed robotics, which in an ordered fashion immobilizes small DNA fragments (referred to as probes), on to these chips. Probes vary in size from oligonucleotides (25-85 base pairs) to genomic clones such as bacterial artificial chromosomes (80,000-200,000 base pairs). Array CGH allow simultaneous enumeration of all 23 chromosome pairs while evaluating multiple loci along the length of each chromosome with a diagnosis rate of close to 100%. Unlike FISH, CGH arrays do not require cell fixation. Aneuploidy screening using microarray circumvents the need for karyotyping and can be automated. Image acquisition is carried out using high intensity laser scanning rather than fluorescence.
microscopy and quality control, normalization, and array analysis can be performed by in-built software tools or third party softwares. Since the analysis is fully automated, the entire procedure is performed is less than 24 hours. Newer commercial arrays can provides highest density coverage of over 2.6 million markers that includes ~750,000 SNP markers and ~2 million non-polymorphic (CNV) markers covering all the annotated genes, cancer genes, X chromosome genes, cytogenetic marker and haplo insufficiency genes to allow the detection of losses of heterozygosity, uniparental disomy, and regions identical-by-descent, in addition to low-level mosaicism and sample heterogeneity. Furthermore, these arrays provide a superior resolution for detection of smaller structural variations. However, there are certain minor limitations to array CGH in that it cannot detect polyploidy, balanced translocations or inversions, point mutations, small insertion-deletions, triplet repeat expansions, or gains and losses in regions of the genome not covered by the array. The successful clinical application of array CGH for genetic diagnosis of patients with unexplained developmental delay/intellectual disability, autism spectrum disorders, with or without multiple congenital anomalies has spurred interest in applying array CGH technology to pre-implantation genetic testing and aneuploidy screening procedures. In summary, array CGH has revolutionized the field of molecular cytogenetics and has driven cytogenetics from the microscope to the computer empowering high-throughput analysis of hundreds or thousands of discrete regions across the entire genome to identify chromosomal abnormalities. Array CGH is proving to be a powerful tool for the identification of novel chromosome syndromes, thus allowing accurate prognosis and phenotype-genotype correlations. Current guidelines and recommendations strongly support the use of array CGH for chromosomal analysis as a first-tier cytogenetic diagnostic test as an alternative to karyotyping for patients with intellectual disability, neurobehavioral phenotypes, and congenital malformations.

**Single nucleotide polymorphism arrays.** Another similar microarray based technology referred to as single nucleotide polymorphism (SNP) array that utilizes 10 million genome-wide SNP markers has great potential in clinical diagnostics. In addition to the detection of copy number variations, the genotype information obtained with SNP array analysis enables the detection of stretches of homozygosity and thereby the possible identification of recessive disease genes, mosaic aneuploidy, or uniparental disomy. Single nucleotide polymorphisms are most common gene variations (1 in 300 bp) found in the human genome that can serve as markers for genome-wide association study. A typical low density coverage SNP array will have 100,000 SNPs and a high resolution array may have more than 1 million SNPs to assess with appropriate software to analyze the generated genetic information. Single nucleotide polymorphism arrays like other methods also rely on whole genome amplification (WGA). Newer methods of performing WGA like MDA and GenomePlex can be used for SNP array. After amplification, the DNA is labeled with fluorescent dyes and hybridized to the array. The signal intensities are then interpreted as SNP calls (allele 1 and allele 2) using powerful computer software to allow for diagnosis of inheritance. While oligonucleotide arrays with high-density exonic coverage remain the gold standard for the detection of CNVs, Single nucleotide polymorphism arrays allow for detection of consanguinity and most cases of uniparental disomy and provide a higher sensitivity to detect low-level mosaic aneuploidies. Suniparental disomy arrays can test genetic disorders as well as chromosomal abnormalities at the same time and are finding its way in laboratories performing prenatal diagnosis and preimplantation genetic testing.

**Next generation sequencing.** Sanger sequencing is currently the most widely used, best established, and best accepted sequencing technology that is considered gold standard to detect small sequence variations. However, the throughput is limited to analysis of a single gene in a single experiment. The limitation has been overcome by next generation sequencing (NGS) that offers massive parallel sequencing of the entire genome in a time- and cost-effective manner. Various NGS platforms such as Roche 454, Illumina GA IIx, and Life Technologies SOLiD, and Ion Torrent are now commercially available with different detection chemistries. All the platforms require genome or exome library preparation that involve fragmenting genomic DNA and attaching specific adapter sequences, followed with exome enrichment (for exome sequencing), and these clonal amplicons then serve as sequencing substrate to yield sufficient signal-to-noise reads.

Currently, signal outputs are based on luminescence, fluorescence, and changes in ion concentration. It is powerful technology with its reduced cost per base, high accuracy, and low turn-around time to generate the data has enabled clinicians and researchers to identify genetic susceptibility markers and inherited disease traits of rare single gene disorders. Identifying causal SNPs in the coding regions (exonic variants) and those present in other non-coding (unknown functional) regions of the genome are now becoming an integral part of
Molecular genetics in Saudi Arabia … Abu-Amero & Kondkar

medical clinical genomics, and NGS has revolutionized the practice of molecular diagnostic testing. Indeed, NGS has already been used to help diagnose highly genetically heterogeneous disorders, such as X-linked intellectual disability, congenital disorders of glycosylation and congenital muscular dystrophies; to detect carrier status for potentially lethal and genetically heterogeneous inherited cardiomyopathies and cardiac channelpathies; and to provide a non-invasive method to analyze fetal genome that may allow diagnosis of all inherited and de novo genetic diseases by sequencing of the maternal plasma DNA. Another important clinical utility of NGS is in the diagnosis of mitochondrial disorders. Conventional diagnosis of mitochondrial disorders is complicated by the dual genome involvement namely the mitochondrial genome and ~1500 genes encoded in the nuclear genome. Thus, it requires the use of several different approaches to diagnose the maternally inherited mitochondrial DNA disorder and the autosomal recessive, dominant, and X-linked nuclear genes in a stepwise manner making it extremely cumbersome and difficult. Next generation sequencing can provide a rapid alternative to analyze the entire mitochondrial genome, haplotyping mtDNA point mutations and deletions, and identify the causative gene(s) in a single step. Clinical diagnostics laboratories are actively engaged in implementing NGS-based tests. Next generation sequencing-based genetic testing’s can be utilized in 3 formats: a) Comprehensive gene panel testing; b) Whole genome sequencing (WGS) and c) Whole exome sequencing (WES).

Comprehensive gene panel testing. This approach is also referred to as NGS-based targeted gene sequencing to sequence all the causal genes for a particular group of diseases such as inherited cardiomyopathies and the congenital muscular dystrophy to name a few (Table 1). These disorders are difficult to analyze using the conventional PCR-based Sanger sequencing because of the involvement of mutation(s) in large number of different genes (such as more than 30 for Charcot-Marie-Tooth). In diseases such as familial adenomatous polyposis coli (APC) and Lynch syndrome caused by mutations in the APC and DNA mismatch repair genes, respectively, a targeted NGS-based strategy is sufficient for diagnostic purpose.

Whole exome sequencing. Exome sequencing refers to sequencing of the entire protein coding regions of the human genome. The protein coding regions of the genome are known as the exome and constitute 1% (30 million bases) of the human DNA and 85% of the disease causing mutations are located in these regions. Whole exome sequencing provides nucleotide by nucleotide sequence information of, the human exome of an individual to a depth coverage necessary to build a consensus sequence with high accuracy. This is then compared to standard population-based normal reference sequence and the result is interpreted by genetics and genomics board-certified laboratory directors and clinical geneticist. Unlike the panel testing, WES is unbiased and not limited to specific set of genes. With commercial availability of exome capture kits, coupled with NGS technology, exome sequencing has become an efficient way to establish a diagnosis in a wide range of diseases with genetic etiology.

In clinical settings, WES can be requested when a patient history and clinical findings strongly suggest that there is an underlying genetic etiology. In some cases, the patients may have negative results for multiple genetic tests known to be associated with the disease without identifying an etiology or for diseases characterized by marked genetic heterogeneity. In other cases, the clinicians may opt to perform WES early in diagnosis with or without a clear clinical diagnosis in an effort to expedite a possible diagnosis. Exome sequencing has dramatically altered patient disease diagnosis and management, saving considerable time and money, and reducing the unnecessary testing and clinically invasive procedures. For couples, planning

<table>
<thead>
<tr>
<th>Disease category</th>
<th>Gene panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congenital muscular dystrophy</td>
<td>PMM2, RYR1, SEPN1, COL6A1, COL6A2, COL6A3, ITGA7, ITGA9, RYR2, LAM A2, LARGE-IT1</td>
</tr>
<tr>
<td>Dilated cardiomyopathy</td>
<td>DES, LMNA, TAZ, G4.5, LAMP2, ZASP/CYPHER/LDB3, SCN5A</td>
</tr>
</tbody>
</table>

Table 1 - Examples of gene panel tests as per disease category by next generation sequencing.
their families and future the high clarity of genetic information can be very helpful. The success rate of causal gene identification is about 50% for even disease cases that have been thoroughly evaluated prior to referral for clinical exome sequencing.\(^{29}\)

This approach typically involves the use of family trio for testing; sequencing of affected individual, the mothers and parents to deduce inherited variants or de novo mutations that are associated with a disease. For X-linked conditions samples from additional family members are required. However, WES can also be performed on affected individual only and compared to reference database to identify the disease-related gene variants. Whole exome sequencing has been successfully used to identify the genetic cause of numerous Mendelian disorders, in patients with intellectual disability,\(^{30,31}\) and many other rare genetic diseases.\(^{32,33}\)

**Whole genome sequencing.** Whole genome sequencing in contrast to WES, involved sequencing the coding and the non-coding regions of the entire human DNA to identify the genetic causal defect.\(^{34}\) The full potential of WGS can be realized only when we gain a much better understanding of the functions of non-coding regions. Currently, however, comprehensive gene panel-based tests and WES are leading the way in clinical diagnostics. Clinically, WGS should primarily be considered for patients who have exhausted all currently available options. There are certain limitations to NGS-based testing. Triplet repeat expansions, large deletions and duplications are better detected by methods other than NGS. Detection of small in-dels may not be as accurate as detection of base substitutions. Also, genes closely related to pseudo-genes are not uniquely captured by this method. The challenge for laboratories embarked upon deploying NGS-based molecular diagnostic testing would not only be financial but to learn new chemistries and instrumentation as well as entering into the complex area of bioinformatics for NGS data analysis.

**Multiple ligation-dependent probe amplification.** Gene deletions and duplications constitute 5% of all disease-causing mutations. It is now well accepted that number of human genetic disease are caused due to CNV as compared to the reference genome.\(^{35}\) Multiplex Ligation-dependent probe amplification (MLPA) is a recently developed method that can detect CNV of human genes and can be used as a genetic test for molecular diagnosis of diseases associated with major deletions of CNV, including small intragenic rearrangements in up to 40 DNA loci in a single reaction. The assay utilizes 40 different probes and a single pair of primers one of which is fluorescently labeled. Each probe consist of two half probes (5’ and 3’ half probes) that upon binding to perfectly matched target DNA sequence are ligated and amplified. Because only ligated probes are amplified the number of probe ligation products corresponds to the number of target sequence in the test sample which are then size separated by capillary electrophoresis.\(^{37}\) The area under the peak is measured after normalization to normal controls indicating the relative amount of target DNA sequence in the added test sample.

Multiplex Ligation-dependent Probe Amplification genetic testing can be used in several types of inherited neuromuscular disorders such as Duchenne muscular dystrophy (DMD), Becker muscular dystrophy (BMD), Spinal muscular atrophy (SMA), Charcot-Marie-Tooth (CMT) disease, and Hereditary neuropathy with liability of pressure palsies (HNPP) for diagnostic purpose in affected patients, to female carriers, and to for the evaluation the risk of recurrence (Table 2). Over 300 probe sets are commercially available from MRC Holland for large range of common and rare inherited disorders.\(^{36}\) The major advantage of MLPA over other conventional techniques is that it can be multiplexed to detect CNV across the specific target gene without compromising the rate of detection. To site as an example, all the 79 exons of the DMD gene can be analyzed in 2 reaction tubes with a detection rate of >99% providing high throughput and cost reduction.\(^{38}\) The American College of Medical Genetics and Genomics has recently recommended a routine carrier screening for SMA, a severe autosomal recessive neurodegenerative disorder with a frequency of 1 in 25-50, and MLPA can be considered for such a screening as a cost-effective approach.\(^{39}\) Multiplex ligation-dependent Probe Amplification analysis has now become a part of the molecular diagnostic testing approach to detect intragenic copy number changes (dosage alterations) as well as microdeletions involving the target gene and the flanking genes e.g. for diagnosis of neuroborromatosis type 1 and 2 (NF1 and NF2 genes) or Legius syndrome (SPRED1 gene) DNA-based sequencing (Tier-1) are performed along with copy number analysis by MLPA (Tier-2).

**Application of advanced genetic techniques in preimplantation genetic testing.** Traditionally, prenatal genetic testing is widely practiced to detect of wide range of genetic abnormalities via amniocentesis and chorionic villus sampling (CVS) during the pregnancy period. Preimplantation genetic (PG) testing offers an alternative to parents at high risk of transmitting genetic disease to their of spring or to women at high risk of chromosomal aneuploidy without the risks and burdens associated with prenatal diagnosis. Besides, the
gastrointestinal motility disorders, including rare or difficult to diagnose diseases, such as collagen vascular and connective tissue disorders (SLE, lupus, scleroderma) and retinal degenerations (RP, retinitis pigmentosa), macular degeneration, and von Hippel-Lindau syndrome. PGD has also been used to avoid the birth of children with single-gene disorders, such as Duchenne muscular dystrophy, cystic fibrosis, hemophilia, and von Willebrand disease. It has also been used to prevent the transmission of chromosomal abnormalities, such as translocations and inversions, and to test for genetic variations that may increase the risk of developing certain diseases, such as familial cancer syndromes (Li-Fraumeni, BRCA1, BRCA2). PGD also allows for the selection of embryos that are not at risk for genetic disorders, such as chromosomal abnormalities, and can be used to screen for non-syndromic deafness and other genetic conditions. In addition, PGD can be used to test for genetic variations that may increase the risk of developing certain diseases, such as familial cancer syndromes (Li-Fraumeni, BRCA1, BRCA2) and other genetic conditions, such as non-syndromic deafness. The use of PGD is becoming more widespread due to advances in molecular genetics and the ability to detect genetic variations with high accuracy.

Table 2 - Molecular genetics in Saudi Arabia

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene/Proportion of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMA</td>
<td>SMN1-SMN2 &gt;95%</td>
</tr>
<tr>
<td>DMD</td>
<td>DMD &gt;95%</td>
</tr>
<tr>
<td>BMD</td>
<td>BMD &gt;95%</td>
</tr>
<tr>
<td>CHM</td>
<td>CHM &gt;95%</td>
</tr>
<tr>
<td>HNPP</td>
<td>HNPP &gt;95%</td>
</tr>
<tr>
<td>CMT</td>
<td>CMT &gt;95%</td>
</tr>
<tr>
<td>DMD/BMD</td>
<td>DMD &gt;95%, BMD &gt;95%</td>
</tr>
<tr>
<td>SMA</td>
<td>SMN1-SMN2 &gt;95%</td>
</tr>
<tr>
<td>DMD</td>
<td>DMD &gt;95%</td>
</tr>
<tr>
<td>BMD</td>
<td>BMD &gt;95%</td>
</tr>
<tr>
<td>CHM</td>
<td>CHM &gt;95%</td>
</tr>
<tr>
<td>HNPP</td>
<td>HNPP &gt;95%</td>
</tr>
<tr>
<td>CMT</td>
<td>CMT &gt;95%</td>
</tr>
<tr>
<td>DMD/BMD</td>
<td>DMD &gt;95%, BMD &gt;95%</td>
</tr>
</tbody>
</table>

Despite the potential benefits of PGD, there are also concerns about the ethical, social, and economic implications of its use. These include the potential for eugenics, the risk of misdiagnosis, and the financial costs of testing and treatment. The rapid advances in the field of molecular genetics and the development of new technologies have allowed for the integration of these techniques into public health programs and the prevention of genetic diseases. However, the lack of access to these resources in many countries, particularly in developing nations, limits the ability to implement these technologies on a larger scale.

In conclusion, PGD provides a valuable tool for families who are at risk for genetic diseases, allowing them to avoid the transmission of these diseases to their children. However, the ethical, social, and economic implications of its use must be carefully considered in order to ensure its responsible implementation.
on outsourcing the genetic testing by sending samples to internationally known laboratories. Needless to say, that this practice has many disadvantages like the high cost endured, long turn-around times, breach of confidentiality and increased possibility of sample's losing or perishing during transition. Additionally, genetic testing in most overseas laboratories is based on the ethnicity of the individuals where the laboratory is located and therefore the mutation(s) regularly screened for in certain populations may not be suitable for testing in the Saudi patients. Therefore, there is an urgent need to create a national center for hereditary disorders (NCHD) which look after genetics testing for various hereditary disorders. The center should house under one roof services such as newborn screening, fully functional clinical genetics services, cytogenetic and molecular genetics laboratories. Additionally, genetic counseling could also be part of the services provided. The NCHD will be the reference hub for all hospital and health care providers for hereditary diseases. Obviously, creating such a center will be costly and will require advanced equipment, well trained staff and building which can house the various laboratories, clinics and of ces. The center will be able to look after genetic testing for all various hereditary diseases. Whether testing for monogenic common diseases, polygenic diseases. The center should also carry out and should run screening programs and research programs aiming at the identification of genetic risk factors for diseases common in the Saudi population. The center also should have a research and development unit with the task of translating research findings into appropriate clinical applications. The unit should also test new technologies and help integrate it into the routine clinical services. Different with the translation of research findings need to be understood and addressed if genetics and genomic research is to fulfill its promises toward improving diagnosis, treatment and prevention. I will finish my discussion with a statement from the WHO expert consultation\textsuperscript{46} concluded that “Genetic advances will only be acceptable if their application is carried out ethically, with due regard to autonomy, justice, education and beliefs and resources of each nation and community”.

We conclude that advanced genetic techniques could not be ignored as it of ered a more comprehensive testing for various hereditary disorders. There is an urgent need to create a national center for hereditary disorders which will take in its shoulders the task of integrating advanced genetic testing techniques into routine testing practices.

References

Molecular genetics in Saudi Arabia… Abu-Amero & Kondkar


