

The History and Evolution of Molecular Diagnostics

Molecular Diagnostics (MDx) is among the most highly complex disciplines in modern science, a fact which has made its development both long and arduous. The theoretical gestation of genetics spanned well over 2,000 years, from classical Greek thinkers like Hippocrates and Aristotle, all the way to the advent of cellular biology in the 19th Century and through to modern times.

Deoxyribonucleic Acid (DNA) and Ribonucleic Acid (RNA) were formally named in the first part of the 20th Century, with the genetic coding mechanism finally cracked in 1965. The Human Genome Project (HGP) was begun by the US government in 1987, during which time a multitude of scientists and clinicians began employing the knowledge that had thus far been gleaned for their own particular uses.

The Human Genome Project was completed in 2003, stimulating greater advances not only in DNA and RNA sequencing, but in other areas of Molecular Diagnostics. Simultaneous advances in laboratory information systems made it possible not only to store and copy the growing dataset, but to use cutting-edge digital algorithms to help extract the meaning within. And advances in proteomics and metabolomics helped to increase our understanding of the biosciences in near equal measure.

The latest MDx technologies include nanopore readers, one of which was recently used to read the Ebola virus genome in only 44 seconds. Such advances – combined with concurrent advances in Computational Biology and associated areas using software algorithms – are helping to dramatically reduce diagnostics times, allowing clinicians to treat their patients far more quickly and efficiently.

Introduction

Molecular Diagnostics is a branch of the biotechnology industry, and involves the detection of genetic patterns in DNA and RNA, as well as patterns in protein generation and usage. These detected patterns (both genomic and proteomic) are used by scientists for classification and for the broadening of knowledge in connected fields, and by clinicians for aiding with diagnosis, prognosis and therapeutic monitoring.

The industry is a natural outgrowth of both science and healthcare, and of medical technology and molecular biology specifically. Uses of Molecular Diagnostics are manifold, including clinical pathology, forensics testing, epigenetics, immunotherapy and immunosuppression, metagenomics, molecular endocrinology, molecular oncology, toxicology, personalized medicine and more.

Today the Molecular Diagnostics industry is an exploding field with a current market size of approximately \$8B.

Two Millennia of Inquiry

Before the rise of Molecular Diagnostics in the 1980s, clinicians were limited to a generalized approach to diagnosing health problems and for creating solutions for those problems. They simply didn't have any tools they could use to access the genetic information of their patients, or for putting that information to use. But what they did have was access to human history, which itself is one of the greatest tools for stimulating innovation.

The history of genetics stretches back over two millennia to Hippocrates, who is credited as being the first person to have speculated about inherited traits, which he also suspected involved some type of material transfer system in the reproductive process.

Fast-forward to the mid-1800s with Charles Darwin at the helm of evolutionary biology, promoting the idea that evolution was "descent with modification". Then in 1866 Gregor Johann Mendel, an Augustinian friar whose pioneering work on the pea plant clearly showed documentable inheritance patterns. And although his work – which made use of breeding experiments – didn't gain much traction until it was re-discovered in 1900 by another team (Hugo de Vries, Carl Correns and Erich von Tschermak), by 1925 the model known as 'Mendelian inheritance' was widely accepted. August Weismann's contribution in 1883 of his Germ Plasm theory was also instrumental in bringing about the coming sea change within scientific circles.

On the clinical side, Oxford physician Sir Archibald Edward Garrod collected historical information in 1902 about the family of one of his patients, and concluded that his patient's alkaptonuria was a recessive disorder passed down through inheritance. This was quite possibly the first time in history that genetics – the idea that human disease can be inherited – was discussed in a clinical setting.

DNA and RNA

In 1869 a Swiss physician named Friedrich Miescher analyzed the pus that was in the discarded bandages of one of his patients, and discovered that the discharge contained microscopic material. He called the substance 'nuclein' because his microscope allowed him to see that the substance resided within the nucleus of the cells. Then in 1878, Albrecht Kossel was the first scientist to isolate nucleic acid, and went on to isolate the five primary nucleobases later in his career.

At the beginning of the 20th Century the

terms DNA and RNA had not yet been invented – but their existence was no secret in microbiology circles. In those early days DNA was known as ‘thymus nucleic acid’, because the tissue that was being studied had been extracted from a thymus gland. In a similar way, RNA was known as ‘yeast nucleic acid’, with yeast as the source material.

Phoebus Levene identified the nucleotide unit of yeast nucleic acid in 1909, carefully noting the base, sugar and phosphate components. Then in 1929, while studying thymus nucleic acid, he identified deoxyribose sugar. Leven’s ‘tetranucleotide hypothesis’ was the first to suggest that thymus nucleic acid was composed of four nucleotide units linked by the phosphate groups – although he thought that the ordering of the nucleotide units was the same in every cell.

It was Nikolai Koltsov who, in 1927, proposed his idea that hereditary traits were carried by two-stranded – or “mirrored” – molecules, with each strand acting as a genetic template. The very next year, Frederick Griffith performed an experiment using two different but compatible bacterial species, which provided the first real indication that DNA – as it was later known – was the true carrier of hereditary information.

In 1944, Oswald Avery noticed that when one pneumococcus species was intermingled with another, the first innocuous species suddenly took on the deadly characteristics of the second. He called this the ‘transforming principle’. Avery’s finds had a huge impact on fellow researcher Erwin Chargaff, and in 1950 Chargaff published his ‘Chargaff’s Rules’ which, in part, stated that the composition of DNA is subtly different between different animal species.

The Hershey–Chase experiments, conducted by Alfred Hershey and Martha Chase in 1952, focused on the virus Enterobacteria phage T2. The team discovered that because the DNA of the virus entered the host bacterium without the entry of much of the virus’ protein, that DNA was the likely genetic transfer mechanism. Although the

experiment was inconclusive, Hershey went on to win the 1969 Nobel Prize – along with Max Delbrück and Salvador Luria – for their work on viral genetics.

The Double-Helix and the Code of Life

By the 1950s DNA and RNA were finally recognized as the containers of the genetic information in living systems. The double-helix model of DNA was provided by Francis Crick and James Watson in 1953, using X-ray diffraction imagery (with the famous ‘Photo 51’ taken by Raymond Gosling in the previous year).

In 1953, astronomer George Gamow created his ‘RNA Tie Club’, for which he hand-picked 20 scientists – a number which matched the number of amino acids in the human body – and had each of those scientists wear a tie which corresponded to a single amino acid. His tie club provided Gamow with a way to show society his part in the race that was on at the time: who would be the one to crack the genetic code.

In 1959, working separately, Marthe Gautier and Jérôme Lejeune each discovered that Down’s syndrome was caused by an additional gene in the DNA cluster – an extra copy of chromosome 21 (which is why the disorder was also called ‘trisomy 21’). This was absolutely groundbreaking because it was the first time that chromosomes themselves were being identified, using karyotype techniques such as Giemsa staining.

After the revolutionary discovery of the double-helix formation of DNA by Francis Crick and James Watson in 1953, they postulated that the strand pairing suggested the potential of being part of a copying mechanism. This suggestion won the pair a Nobel Prize in 1962, and stimulated further research by others.

In 1957 Arthur Kornberg discovered the first DNA polymerase, which opened up the possibility of synthetic copying of genetic material for the first time, and won him the Nobel Prize two years later in 1959. His work was incomplete, however, as much more needed to be learned about the replication process.

At around the same time, Marshall

Nirenberg began working on the enigma of DNA coding at the National Institute of Health (NIH). His beginning work was based on experiments in which he attempted to show the possibility of RNA-triggered protein synthesis. Using the cytoplasm of E.Coli bacteria, he tested each of the 20 amino acids on the samples to find out which amino acid would be incorporated into the final protein.

In 1961 Nirenberg’s team finally discovered that phenylalanine was the amino acid which could be forced to form the repeating protein chain with Uracil. It was this experiment which proved to the team that the genetic code could be broken.

Because Nirenberg was racing against Nobel laureate Severo Ochoa to be the first to crack the code of life, Nirenberg’s colleagues generously set down their own experiments to assist in the effort. Ochoa decided to give up the race in ’62, because he felt that since Nirenberg had the lead, his own time would be better spent on a different task.

After four years of constant toil, with dozens of scientists scrambling about his lab, Nirenberg finally cracked the genetic code in 1965.

A Chain Reaction

During that same period, H. Gobind Khorana was working on a way to completely synthesize a gene using oligonucleotides. He developed multiple techniques, some using oligonucleotides as structural elements, and others making use of oligonucleotides as templates and primers. Khorana was awarded a Nobel Prize in 1968.

In 1971 Khorana’s team came up with an idea they called ‘repair synthesis’. This effort, developed as an attempt to increase yield, is today seen as one of the initial steps toward Polymerase Chain Reaction (PCR) amplification, a process that would become extremely instrumental in the eventual rise of Molecular Diagnostics.

Although ‘repair synthesis’ was incapable of achieving the exponential yield of PCR, it did illustrate that a copy process was possible. At around

the same time, another member of Khorana's team – Kjell Kleppe – wrote a paper outlining an approach that was strikingly similar to PCR, although his proposed process was never fully realized.

Cetus Corporation – the company that was eventually to own the rights to the PCR process – was also founded in the same year (1971). They would eventually become the first Molecular Diagnostics company, with projects in the early '70s that included the development of diagnostic tests for genetic mutations.

One key that was missing, however, was a polymerase that could withstand high temperatures. In 1976, Taq polymerase was successfully extracted from *Thermus aquaticus*, a thermophilic bacterium. This polymerase could maintain stability at temperatures as high as 97.5 °C (with an optimum temperature range from 80 – 80 °C). This higher temperature range opened up the door to new avenues in sequencing.

Frederick Sanger put the Taq polymerase discovery to good use, and won the Nobel Prize in 1980 for his new method of DNA sequencing – known as the chain-termination method, or Sanger sequencing. This method greatly simplified the sequencing process, and was the most widely employed DNA sequencing method for nearly four decades after Sanger's discovery (he actually made the discovery in 1976).

In 1983 – back at Cetus Corporation – Kary Mullis used his knowledge of the Sanger method as the foundation for a new and revolutionary technique. He realized that a chain reaction could be triggered by the repeated use of DNA polymerase – a chain reaction that would provide amplification of any DNA segment.

PCR amplification was born – and with it, a whole new world of possibilities in the world of genetics.

In 1987, Applied Biosystems (AB) unveiled the AB370, the first automatic sequencing machine. Their revolutionary machine made use of capillary electrophoresis, which

increased the speed and accuracy of the sequencing process.

This was followed by Dupont with their Genesis 2000 which made use of a new fluorescent labeling technique.

The Human Genome

The first great leaps in modern genomics were finally accomplished – the knowledge of the coding mechanism, and the means of sample amplification. But without the deeper knowledge of genetics – the knowledge of what the code was actually 'saying' – little progress could be made in the area that all this would eventually be used for: health and medicine.

So the next race was on – the race to crack the human genome.

In 1987 the United States Department of Energy (DOE) established the first human genome project (HGP), in an effort to discover how to protect humans from the mutagenic effects of radiation. The next year the NIH was granted funding for a similar purpose, and the two agencies soon signed a cooperation agreement (Memorandum of Understanding). In 1989 the name of the project was changed to the National Center for Human Genome Research (NCHGR).

The project would make extensive use of technologies that were in development at that time, which included PCR and capillary sequencing, restriction fragment-length polymorphisms (RFLP) and pulsed-field gel electrophoresis. Also included was the use of artificial chromosomes developed using yeast and bacteria.

RFLP, developed in 1978 by David Botstein et al, was an adopted technique that had been used by Ray White on the *Drosophila* genome. Botstein's team felt that even with the vast differences between fruit flies and humans, attempting the technique with human genes was an exciting prospect. RFLP would soon become a mainstay in genetic marker technology.

Much of the work during the early days of the HGP was in developing methods of streamlining the process of DNA sequencing (extracting, marking

and categorizing genetic material from DNA samples). The scientists who were involved knew that once this streamlining could be accomplished, the 'elucidation of the genome' would come to fruition much more quickly.

By 2003 the NCHGR and their partners announced that their goal of mapping the human genome was complete, and although it seemed likely to some that this mapping would be continually refined as time marched onward, the existing mapping has been immensely useful to the many scientists and clinicians who have made use of this immensely important database.

The First Uses of Molecular Diagnostics

In 1991 the company Myriad was formed – one of the first of the Molecular Diagnostics companies. Their vision was to use genetic information to improve healthcare, and to do this they set out to develop the newest, cutting-edge diagnostic tools.

In 1996 Myriad announced the release of the first ever patient-level genetic testing system, BRACAnalysis. Detecting the presence of a BRCA1 or BRCA2 gene mutation, the test would provide clinicians with a fast and efficient diagnosis for hereditary breast and ovarian cancers. For their innovation Myriad was awarded the Laguna Niguel 1997 Best of Biotech Award.

In 1995 two different teams – Sooknanan and Malek, and Kacian and Fultz – were simultaneously developing alternatives to PCR, with nucleic acid sequence-based amplification (NASBA), transcription mediated amplification (TMA) and self-sustained sequence replication (3SR). These techniques relied on RNA polymerase rather than DNA polymerase, to render a reverse transcriptase that would produce DNA from the RNA templates.

Almost immediately, these innovations made it possible to measure HIV-1 RNA in blood plasma, a test which became the first widespread use of molecular diagnostic technology. Having opened the door to viral load testing, kits soon became available for the hepatitis C

virus (HCV), cytomegalovirus, Epstein-Barr virus and BK virus.

It was during these same years that the diagnosis of herpes simplex virus (HSV)-associated encephalitis was aided by the testing of HSV DNA in cerebral spinal fluid, as an alternative to brain biopsy. Shortly thereafter a test for the detection of Chlamydia trachomatis was approved by the US Food and Drug Administration – their first-ever approval of a molecular diagnostics testing kit.

Another innovation during the late 90s was the simultaneous detection and amplification of target nucleic acid in real time. This process not only eliminated the risk of carryover contamination, it also reduced the turnaround time to just a few hours.

In 1995, Craig Venter employed what was known as the ‘whole genome shotgun sequencing’ technique to complete the Haemophilus influenza (bacterium) genome. Although the shotgun sequencing technique was said by some to be unreliable – which is why it was rejected for use by the HGP – it nonetheless became quite popular in the genetics community.

Back in 1993 the NCHGR had established their Division of Intramural Research (DIR). Three years later, eight separate NIH divisions morphed into a singular institute, the Center for Inherited Disease Research (CIDR), for the study of genetic causes in relation to disease. In 1997 the NCHGR was renamed the National Human Genome Research Institute (NHGRI).

In 1996 the Health Insurance Portability and Accountability Act (HIPAA) came into effect, which – among other things – made provisions for standardizing electronic health care (including standards for electronic medical records (EMR) and other electronic transactions), and established national identifiers for various entities in the medical establishment.

In 1999 the American Board of Medical Genetics and the American Board of Pathology announced molecular genetic pathology as a new joint subspecialty. In the same year, Sydney D. Finkelstein

et al devised a system of embedding genetic tissue in cold-temperature plastic resin; this resin was the water-miscible methyl methacrylate polymer known as Immunobed.

Also in 1999 a team led by Khuong Truong developed the technique known as Fluorescence In Situ Hybridization (FISH) for the detection of lung cancer. The technique used two-color fluorescence on the long and short arms of chromosome 3 to detect an imbalance.

In the late 90s the pharmaceutical industry recognized the need for simultaneous advances in diagnostic tools on the clinical side of the equation – advances which were already being made by various individuals and teams. In 1999 the Association for Molecular Pathology (AMP) had co-founded The Journal of Medical Diagnostics, which provided an informational outlet for specialists working in the field, and The Journal of Molecular Diagnostics started the same year.

The Millennium Turns

Although 2003 marked the official completion of the human genome project, a bulk of the work had already been accomplished by 2000. This provided the catalyst for an explosive environment of innovation within the scientific community, fueled perhaps in part by the psychological thrust of the turn of the millennium. Target-based drug discovery became a huge part of the R&D budgets of pharmaceutical companies, as biochemists put the estimated number of genes that could be pharmacologically targeted at between 3000 and 5000 (of the roughly 30,000 total genes in human DNA). But scientists eventually began to see that this was likely a gross overestimate, with only 500 targets discovered in the ensuing five years of highly active research.

The turn of the millennium also saw an expansion of point-of-care testing devices and systems, which for the first time allowed physicians and nurses to use such technology directly at the location of their patients. This expanse of point-of-care medical testing systems

has been progressing ever since, and will undoubtedly continue far into the future.

In 2001 Dr. Bert Vogelstein of John Hopkins University was presented the Award of Excellence by the AMP for his major contributions to the understanding of genetics in connection with colorectal cancer. In his acceptance speech, Vogelstein discussed genomic micro- and macro-instability, and studies involving gene expression changes in tumors. He also discussed his team’s use of ‘digital PCR’ which, although the term had been invented in the early 90s, had seen a resurgence due to other connected elements (software enhancements, etc.).

In the same year, Dr. Rudy Leibl of Columbia University discussed the role of genetics in the expression of proteins and their effects on human body weight.

Bioinformatics

It was also around the turn of the millennium that advances in information technology – specifically bioinformatics – allowed huge productivity boosts for both scientists and clinicians. In 2000, Peter Cooper at the National Center for Biotechnology Information began doing workshops which taught specialists how to use the web-based, sequence-manipulative genetics tools that his organization had developed. Another web-based tool, the GeneTests genetic test database – run by Bonnie Pagon at the University of Washington, allowed instant access to tests – both scientific and clinical – from hundreds of labs worldwide.

In biochem labs around the globe, Laboratory Information Management System (LIMS) and other specialized software were now making it possible to work with immense datasets with much greater ease, and allowed staggering computational analyses which were theretofore impractical or impossible. On the clinical side, physicians and lab technicians were finally able to take advantage of digital means for working with patient data using Electronic Medical Record (EMR) systems as well as Laboratory Information System (LIS). These IT advances reduced processing

times by eliminating the need for juggling paper-based files, and by speeding up the flow of information between specialists in disparate geographies using secure internet connections.

It was at the turn of the millennia that Psyche Systems released The LabWeb System, a comprehensive Laboratory Information System using web-based features like hyperlinked pages. The system was highly advanced at the time, and could incorporate many different multimedia types to enhance the user experience. Psyche Systems now has eight state-of-the-art software packages in various areas of molecular diagnostics, including Laboratory Information Systems, reporting systems and connectivity solutions.

Another stimulator on the electronic side of medicine was the establishment of Health Level Seven (HL7), which provided an international standardization with digitally transferred data (clinical and administrative) in the healthcare industry. This move implicitly required software companies to make these standards a reality within the software code that so much of modern healthcare relies on.

Quantitative PCR, or qPCR, was already in development in the 90s as an update to traditional PCR. With the production of complementary RT-PCR using an initial reverse transcription (RT) step, amplification of any type of RNA could now be accomplished. qPCR quickly became the new standard in the biological assay toolbox.

In 2001 Jules Meijerink et al discovered a common problem in clinical settings, in which patient samples and calibration samples showed unequal efficiencies during amplification (typically due to contamination with various compounds that were common to the qPCR process). They developed the efficiency compensation control (ECC) to compensate for the problem.

Next Generation Sequencing

In 2000, Jonathan Rothberg founded 454 Corporation as a subsidiary of Curagen. His ideas were based on work done by P. Mayer et al that was outlined

in a white paper of his from 1998. This work formed the foundation of the next generation sequencing (NGS) process (aka 'second generation process' and 'high-throughput sequencing').

NGS was a groundbreaking innovation that combined large-scale parallel sequencing, high throughput and economy of operation. The platforms based on this approach utilized miniaturized and parallelized platforms to process up to 600 million short reads of DNA per 10-hour run.

The new technique allowed – for the first time – an entire genome to be sequenced at once. This proved useful not only with normal genome sequencing, but also with genome resequencing, transcriptome profiling (RNA-Seq), DNA-protein interactions (ChIP-sequencing), and epigenome characterization.

Employing new advancements of NGS, 454 Life Sciences (part of the 454 Corporation) released their new Genome Sequencer 20 (GS20) machine in 2005. This was the very first complete NGS sequencer available on the open market, which caused an explosion of productivity around the world, in a process which became known as 454 pyrosequencing. Roche Diagnostics acquired 454 Life Sciences in 2007.

NGS machines found use in myriad fields, including metagenomics, molecular endocrinology, anatomic pathology, cytogenetics testing, immunosuppression and immunotherapy, precision medicine, and evolutionary biology.

Other Sequencing Technologies

Many newer high-throughput sequencing (HTS) methods were developed during this period.

In 2002 Ion semiconductor sequencing (aka Ion Torrent sequencing) made use of the detection of hydrogen ions that are released during the polymerization of DNA. This was one of the 'sequencing by synthesis' methods, in which dNTP is flooded into a microwell containing the DNA sample. If the dNTP is complementary to the leading template nucleotide, hydrogen ions are released,

which triggers an ion sensor.

In 2005 Single-molecule real-time (SMRT) sequencing was developed by Pacific Biosciences. SMRT made use of zero-mode waveguides (ZMWs) and phospholinked nucleotides. SMRT is the most widely used sequencing platform today, even though it was developed over a decade ago.

In 2007 another 'sequencing by synthesis' method was created by Illumina/Solexa, which used modified dNTPs containing a terminator which blocks further polymerization. This meant that only a single base could be added by a polymerase enzyme. This type of reaction was carried out simultaneously across millions of template molecules on a single surface, making it highly parallelized.

Combinatorial probe-anchor synthesis (cPAS) sequencing used a combination of sequencing by hybridization (SBH) and sequencing by ligation (SBL) techniques on pools of probes, each probe consisting of an anchor sequence and nine bases, where the interrogation position was then dye-labeled, washed and imaged.

Sequencing by hybridization used microarray technology – also known as DNA chips, ChIP sequencing, and bio-chips – which contain up to billions of synthetic oligonucleotides on a single surface.

Sequencing by ligation (SOLiD sequencing) determines the underlying sequence of nucleotides in a given DNA sequence by making use of the mismatch sensitivity of DNA ligase.

Polony sequencing, developed by George M. Church at Harvard, used emulsion PCR, an in vitro paired-tag library, ligation-based sequencing chemistry and an automated microscope. This process was used to sequence a full *E. coli* genome in 2005, with an accuracy greater than 99.9999%.

DNA nanoball sequencing, developed by Complete Genomics, employed rolling circle replication for the amplification of genetic material into nanoballs. The sequence was then determined using unchained sequencing by ligation.

Heliscope single molecule sequencing was introduced by Helicos Biosciences. This solution added poly-A tail adapters to DNA fragments, and followed this step with extension-based sequencing and cyclic washes with fluorescently labeled nucleotides. Heliscope performed the reading.

Concurrent Advances in Molecular Diagnostics

Total growth in the Molecular Diagnostics industry from 2000 to 2005 was between 10% and 20%, and most of this growth was directly stimulated by advances in both sequencing technologies and in bioinformatics software technologies. These simultaneous advances made it possible to automate the sequencing process, freeing up time for scientists to focus on data analysis. And the new software tools allowed analysis to be done with greater speed and efficiency, allowing scientists to query databases in more specific ways.

For the first time in history, hospitals could perform genetic testing on a STAT basis, a fact which greatly improved the environment of healthcare. This was an important step in the rise of molecular medicine.

Back in 2001, cooperation between the NIH, the American College of Medical Genetics (ACMG) and the American College of Obstetricians and Gynecologists (ACOG) gave rise to a screening test for cystic fibrosis (CF), which was strongly advised for pregnant women. Assay platforms for the test included Innogenetics CFTR33 LiPA (Alpharetta, GA), Roche CF Gold, and the CF v3 OLA assay (Abbott-Celera, Abbott Park, IL), together with their associated reagent kits (ASRs).

As assay automation came to the fore during this period, sample processing was still an issue. Thermal Systems responded to the problem with the Kingfisher, a smaller nucleic acid extraction instrument. The Tecan provided an innovative platform with its use of robotics for liquid handling. Roche, MagnaPure and Qiagen were soon to follow with their new automated extractors.

The aforementioned ABI 310, which was a single channel CE system, and the ABI 3100, ABI's multi-channel system, made it possible to quickly test for fragile X syndrome and for HIV genotyping.

Sequence-based diagnostics (SBD) were routinely used to assess HIV drug resistance. CE systems became the leader on this front within the clinical environment, making it possible for clinicians to work with the HIV genome, in addition to facilitating the identification of yeast and fungal infections using 16S and 26S rRNA identification.

By 2005 hardware and software platforms had become available for applications as diverse as forensic testing (including forensic toxicology), paternity, immigration, genetic and infectious disease, molecular endocrinology, molecular oncology, evolutionary biology and anthropology, prenatal diagnostics, and the genetic modification of food crops.

In 2007, Applied Biosystems (now Life Technologies) released their SOLiD platform which made use of next-generation sequencing technology to deliver up to 4 billion bases of sequence data per run. This made the SOLiD platform the highest-throughput system of its day, and with a raw base accuracy greater than 99.94%.

Although multiplex-PCR was conceived in 1988 for the detection of genetic deletions, its use increased significantly in the 2000s in the wake of better and faster assay processing. In 2008, multiplex-PCR found new uses, for instance with the analysis of SNPs (single-nucleotide polymorphisms) and microsatellites (tracts of repetitive DNA).

Before the 2000s, Mass spectrometry (MS) was only employed for protein detection and characterization. Since the turn of the millennium, MS has been found to be highly useful for microbial and viral detection.

Point of Care Diagnostics

Point of Care (POC) diagnostics was a term invented in the 1980's, and refers to any test that can be performed near

a patient rather than in a laboratory environment. But it wasn't until the 2000s that POC technology would begin to include genetic testing for molecular diagnostics.

In 2003, following the biological threat of Anthrax at the United States Postal Service, the USPS awarded a \$175M contract to Northrop Grumman Corp., in cooperation with the company Cepheid. The result of the contract was the Biohazard Detection System (BDS), which began national rollout at USPS centers in 2004.

This system, called GeneXpert, could perform real-time PCR on a sample within 30 minutes, by users who had no need for biochem training. This system marked the birth of POC for molecular diagnostics.

Other companies working on POC molecular diagnostics equipment in 2005 were IQuum, Nanosphere, Inc. and Nanogen Inc. in the U.S., Enigma Diagnostics, LGC and Lumora Ltd. in England, and IMM in Germany.

In 2006, Roche introduced LightCycler SeptiFast, a kit which could rapidly detect the cause of sepsis, by analyzing the genetics of microbial agents in the blood. Up to 25 different microbial agents could be discovered using the kit, which proved invaluable for hospitals around the world.

By 2008 companies such as TwistDX were using Recombinase Polymerase Amplification (RPA) isothermal technology to bring genetic technology out of the lab and directly to the patient. Although POC was already a term used for many years, devices were becoming smaller, more robust and more accurate as time progressed.

In 2009, funding for POC molecular diagnostics testing exploded, and by 2013 \$650M had been secured in the industry. But even with such funding, manufacturers at the time noted that although hospitals had readily adopted such test kits, adoption in clinics was dismal.

In 2011 David M. Pearce et al developed a new electrochemical method of detecting Chlamydia trachomatis, and

by 2015 his team's technique was put to use by Atlas Genetics in a POC kit called io System.

Newer POC systems include those by Vantix Diagnostics, the i-STAT Alinity system, Pandora's CDx system, and Luminex and their xTAG systems (8 different kits). And leading the pack is Oxford Nanopore Technologies with its MinION device, allowing DNA and RNA sequencing in real-time, using a USB device that weighs less than 100g.

Newer Technologies

It's an interesting fact that as MDx processes are being continually sped up by various means, including digital analytics and the use of robotics, older technologies have gotten a booster shot and brought back up to the fore – only in a slightly different context. One good example of this phenomenon is multiplex PCR, a process that makes use of standard PCR, but creates a parallelized system within which many PCR assays (and their associated processes) are occurring simultaneously, allowing much more complex analyses as a result.

On the other hand, brand new discoveries are being made more and more often, with the time shrinking between each new innovation, making modern tech-news a highly dense web of explosive modernization. This – the last section of this article – will focus on discoveries that have been made in the last few years, including the newest and greatest, bleeding-edge innovations that have helped in making MDx what it is today.

By 2012, Luminex had on offering their xTAG range of testing kits, including the xTAG Nucleic Acid Assay Technology, xTAG Cystic Fibrosis, xTAG Respiratory Viral Panel, xTAG Gastrointestinal Pathogen Panel, xTAG CYP2D6 Kit, xTAG CYP2C19 Kit.

Around the same time, the eSensor respiratory viral panel was introduced by GenMark Diagnostics. These new kits provided clinicians with more amazing tools to stick into their belts – so-to-speak – decreasing time for assays and analytics for their patients.

In 2015 the BGISEQ-500 sequencer was released by BGI, integrating automated sample preparation, sequencing and data analysis in an all-in-one machine that would fit on a desktop. The machine additionally made use of Radio Frequency Identification (RFID), as well as barcode scanning technologies, for further streamlining of genetic projects. In 2016 the machine was offered with a brand new thermoplastic called PEEK.

BGI's machine employs DNA Nanoballs (DNB) technology, which uses rolling-circle replication, as well as Probe-Anchor Synthesis (cPAS) technology.

In 2016 Illumina released their HiSeq 4000 Sequencing System, which uses patterned flow-cell technology which allows users to sequence "12 genomes, 100 transcriptomes, or 96 exomes in fewer than 3.5 days".

The newest sequencing technologies – called 3rd Generation Sequencing – can read over 10,000 base pairs, or map over 100,000 base pair molecules.

Nanopore Sequencing has made it possible to sequence single molecules of DNA and RNA without the need for amplification. The process also eliminates the need for chemical labeling. In this process, electrophoresis is used to transport a sample through a micro-orifice with a diameter of 1 nanometer. Sequencing is facilitated because of differences in electric current densities caused by different genetic samples, which affects the flow through the nanopore channel.

The aforementioned MinION, by Oxford Nanopore Technologies, puts nanopore sequencing to highly effective use. Their two larger systems – the benchtop GridION and the higher-throughput, higher-sample number PromethION – also employ nanopore technology.

Another cutting-edge area of research are those of Computational Biology, Statistical Genetics and Mathematical Modeling, each of which use statistical methods for extracting meaning from genetic datasets. Dr. Fengzhu Sun is one such specialist in these innovative fields. He and similar specialists are using computers and algorithmic processing to assist with MDx tasks as diverse as

genetic mapping, advanced sequence analysis, protein classification and long-time simulation of protein molecules, fold prediction, structure comparison algorithms, and comparative genomics.

Other computer experts are working in areas such as query optimization and text mining. Query optimization allows MDx specialists more flexibility and power when they interface with datasets using databases such as SQL. And text mining brings in the immense power of Google, allowing the entire internet to be leveraged for the acquisition of subjective associations using unsupervised machine learning and natural language analysis.

One of the most explosive and important fields in MDx is biomarker discovery and validation, and many experts suggest that this is the key to progress. Biomarker discovery and validation will continue for many, many years, with each successive biomarker finding important use in hospitals and clinics around the world, and each one adding to the total body of knowledge.

One such discovery is the use of Single-Nucleotide Polymorphisms (SNP) as biomarkers. Recent genome-wide association studies were done in China on various populations, and it was discovered that SNP loci in five different genes were associated with susceptibility to esophageal squamous cell carcinoma (ESCC). Much more work is being done by many different teams around the world using SNP's as biomarkers.

In the last 5 years, nearly 1000 papers were published on the subject of biomarker discovery through the use of metabolomics, and in that same time many such biomarkers have been validated. This vindicates metabolomics as a highly useful tool in biomarker research, with its use of mass spectrometry and data-mining software.

A recent advancement in bioinformatics is The BioCompute Object (BCO) Project, which began as a collaborative effort between the FDA and George Washington University. It was instigated as a response to the challenge of sharing scientific workflows and

documentation by different teams, not only for the sake of peer reviews but perhaps more importantly to allow information sharing itself to be the catalyst for the evolution of science. The BCO Project now includes over 20 universities, biotechnology companies, pharmaceutical companies, and public/private partnerships.

Other recent MDx advances include work on miRNA and long non-coding RNA, Genetic testing for high-grade osteosarcoma, molecular and point-of-care diagnostics for Ebola, Serum biomarkers for viral hepatitis and hepatocellular carcinoma, electrochemical immunoassay for tumor markers based on hydrogels, effective early detection of oral cancer using a simple and inexpensive point of care device in oral rinses, and DNA Biosensors.

Antibiotic resistance poses a serious threat to humanity, and as such it will be a theme that is likely to impact the MDx space for some time. It has been suggested that antibiotic resistance will take over cancer as the leading cause of death by 2050.

The rise of antibiotic resistance can be attributed to over-prescription of antibiotics, in a healthcare system full of doctors who – being at a technical loss because they cannot quickly discover the exact cause of infection or sickness – will prescribe antibiotics. This problem is exacerbated by the public belief in antibiotics as a ‘cure-all’ (because the average person doesn’t understand the complexities of biology and medicine).

MDx is now coming to the aid, not directly but indirectly, by dramatically reducing the time it takes to pinpoint the cause of infection, via the detection of the specific microbial agent. New systems are allowing a patient’s sample to be loaded into a cartridge and placed into a process that will test for hundreds of different pathogens at once, as well as testing for antibiotic resistance markers. Using such systems, current state-of-the-art technology brings analytical results down to approximately five hours, which is a great boon to physicians who can then send the

needed prescription to their patient’s local pharmacy on the same day.

And on the proteomics front, data-independent acquisition (DIA) is the new technology on the block. Although the wrinkles have not been completely ironed out, DIA looks to be quite promising. Whereas data-dependent acquisition (DDA) has been the paradigm for a decade and a half, in combination with liquid chromatography-tandem mass spectrometry (LC-MS/MS) and multiple reaction monitoring (MRM), DIA samples every peptide in a protein digest, comprehensively and repeatedly, producing a complex set of mass spectra at the output of the process. The spectral output is interpreted using external spectral libraries derived from prior DIA experiments or from auxiliary DDA data, producing identification and quantification nearly equal to that achieved by MRM/PRM. Cutting-edge bioinformatics techniques are being introduced to address the remaining issues with DIA, and the potential results are quite promising.

Conclusion

The future is always bright wherever it features the minds of intellectually-driven men and women, who are ever in pursuit of that ideal of perfection which has been the dream of mankind since time immemorial. In the realm of MDx, physically diminutive POC devices with ever-increasing power and scope will always be the goal for clinicians – and for the armies of scientists and technicians who build those clinicians’ tools. This ideal is perfectly exemplified by the fictitious and oft-alluded-to tricorder, as seen in myriad episodes of Star Trek. And although many specialists see this allusion as trite, such a handheld device is slowly but surely in the process of being manifested by the legions of scientists and great thinkers of our great modern age.