



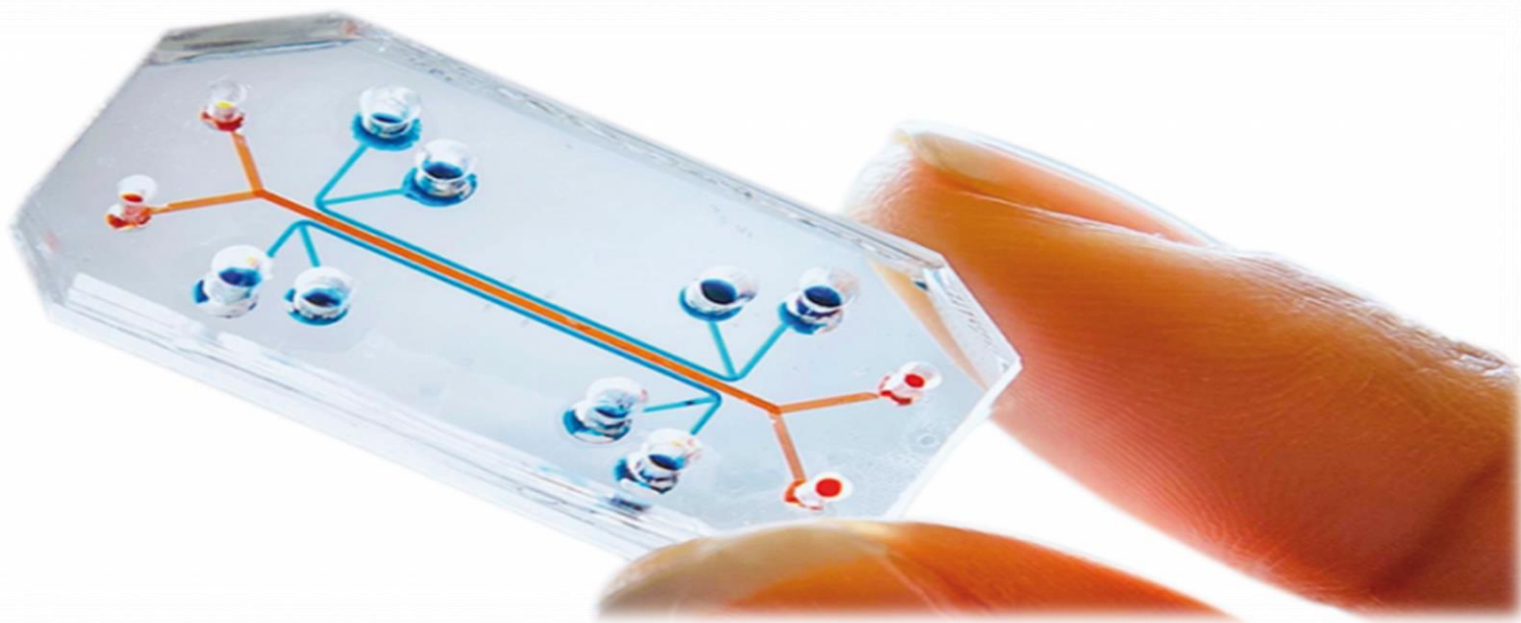
Biotechnology & Engineering

POSTGRADUATE COURSE

“Green Technology Foresight about challenges from biotechnology and ICT”

Digit-BioTech

2019-1-BG01-KA203-062371



LO3 BIOSENSORS & BIOCHIPS FOR SUSTAINABLE FUTURE

Basic level

AUTHOR:
TSVETINA MIHAILOVA & KLIMENT PETROV



Contents

Biosensors & Biochips: An overview	3
Blueprints for Biosensors: Design & Operation.....	3
1. Electrochemical biosensors	4
2. Enzyme-based biosensors	5
3. Immunological biosensor	7
4. DNA biosensors	9
5. The future of clinical biosensors	11
Blueprints for Biochips: Design & Operation.....	12
1. DNA microarray.....	14
2. Microfluidic chip.....	16
3. Protein microarray.....	19
4. Applications of biochips.....	21
References	22



Biosensors & Biochips: An overview

Biochips appeared as an innovative microtechnology platform for analysis of bio-molecule in the 1980s. A variety of technologies, such as life sciences, information technology, microelectronics and micromechanics, are involved in the underlying technologies. Biochips are considered important potential instruments in modern life science research, medical diagnosis, drug discovery, food safety monitoring and agriculture as high-performance, miniaturized, automated and cost-effective features. Biochips are expected to allow the speed and scope of the analytical process to be dramatically increased and to provide enormous economic value. Further more, in recent years, numerous governments and industrial companies in the world have invested heavily in this sector. Biochip technology is just in its early evolution for now. This is an area which is ever-changing. The future for biochip research and development is bright. There is furious competition in this area.

In biomedical and life science research, the miniaturization of chemical and biomedical laboratory processes on microchips is a rapidly expanding field. Lab-on-chip technologies will bring a lot of benefits over their macro-sized counterparts. In particular, higher surface-to-volume ratios result in reduced chemical requirements, reduced waste, better control, rapid processing and significant ability for parallel processing and process incorporation. Lab-on-chip technologies has the potential to have a significant socio-economic impact. In laboratory medicine, completely integrated micro devices for chemical synthesis and disease diagnosis offer a scientific breakthrough and a paradigm change in chemical processing. The human genome project has made a huge contribution to this technology.

In the field of analytics, biosensors allow major innovations that are both facilitating and facilitated by developments in synthetic biology. The potential of biosensors to identify a wide range of molecules easily and precisely makes them highly important to a variety of industrial, medical, ecological, and science applications. Biosensor design strategies are as numerous as their applications, with major groups of biosensors including nucleic acids, proteins, and transcription factors. Based on the expected use, and the parameters required for optimum performance, each of these types of biosensor has advantages and limitations. Especially, considerations such as ligand specificity, sensitivity, dynamic range, functional range, output mode, activation time, ease of use, and ease of engineering must be considered when choosing the design of the biosensor.

BLUEPRINTS FOR BIOSENSORS: DESIGN & OPERATION

Biosensors are sensors that transform bio-recognition processes through a physico-chemical transducer into observable signals, with electronic and optical techniques as two main transducers. The creation of biosensors addresses today's rapidly rising need for clinical diagnostics. A combination of advantages is brought on by the use of biosensors. Biosensors, first, are highly sensitive. This is because biomolecules have a high affinity for their targets, for example, antibodies catch antigens with a



dissociation constant at the nanomolar scale, and DNA - DNA interactions are so much stronger than antigen-antibody. Second, biological recognition is typically very selective. The enzyme and substrate are much like a lock and a key, for example. Such high selectivity frequently leads to biosensors which are selective. Third, the production of inexpensive, integrated and ready-to-use biosensor devices has become relatively easy to develop due to the development of the modern electronic industry. The ability to detect pathogens or perform genetic analysis in hospitals is certainly improved by these biological sensors; more importantly, they are especially useful for small clinics and even point-of-care analysis.

For biosensors with clinical applications, a range of new techniques have been developed. Biosensors are, in general, analytical devices constructed of an element of biological recognition and an optical/electronic transducer. The biological element is responsible for the capture of solution analytes and the transducer transforms the binding event to a measurable signal variation. By the nature of recognition, enzyme-based biosensors, immunological biosensors, and DNA biosensors, could categorize the type of biosensors. In addition, electronic biosensors (electrical or electrochemical), optical biosensors (fluorescent, surface plasmon resonance, or Raman), and piezoelectric biosensors (quartz crystal microbalance) are available depending on the type of transducer.

1. Electrochemical biosensors

For biological sensing, where electrodes function as either electron donors or electron acceptors, electrochemical techniques are particularly useful. Extensive electrochemical experiments have shown that the Marcus electron-transfer theory also complies with heterogeneous electron transfer between electrodes and surface-confined redox molecules, similar to donor-acceptor pairs in homogeneous solutions. This means that small distance changes in redox surface-confined molecules can cause wide variations in the heterogeneous electron-transfer rates that are supposed to translate into detectable changes of electrochemical signals. Hellinga and co-workers, for example, suggested an electrochemical sensing strategy that exploits protein ligand-mediated hinge-bending motions. A gold electrode was first coated with a self-assembled monolayer (SAM), which gives a versatile platform for site-specific protein immobilization. The maltose-binding protein (MBP) was then bound to the surface of the gold electrode with a particular orientation as the redox reporter group of ruthenium (Ru(II)) is fixed at a certain range above the electrode. As the ligand maltose binds to the active site, the Ru(II) reporter moves away from the electrode by causing a hinge-bending motion of MBP (Figure 1).

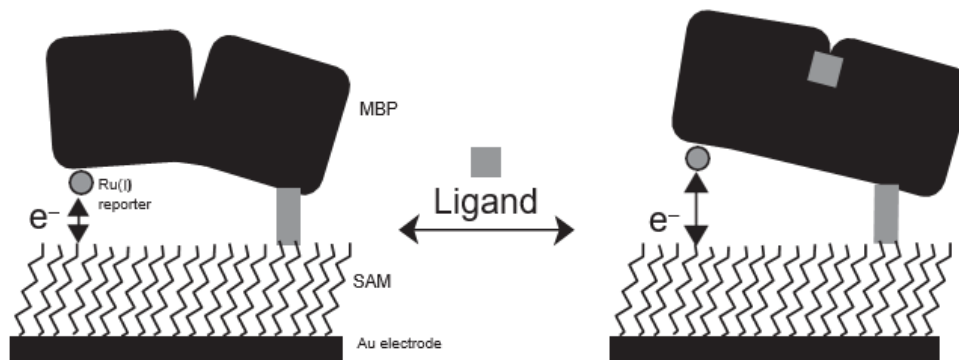


Figure 1. Construction and operation of electrochemical biosensor

This maltose-binding induced distance change causes concentration-dependent decreases in electrochemical signals, thus providing a way for maltose to be electronically sensed. The use of this highly generalized sensing approach to detect various analytes with a family of proteins or enzymes undergoing ligand-binding induced conformational changes has also been shown.

2. Enzyme-based biosensors

The first biosensors ever documented were glucose oxidase (GOD)-based biosensors, established by Clark and Lyons in 1962. Hyperglycemia, a chronically raised concentration of blood glucose, is common for *diabetes mellitus*. As a result, regular control of their blood glucose concentration is essential for diabetic patients. The benefit of electrochemistry coupled with enzyme catalysis is this biosensor, and its newer versions. An electrode immobilized with GOD was Clark's biosensor. The oxidized form of GOD interacts with glucose in the presence of glucose and produces gluconic acid and reduced GOD, with two electrons and two protons involved. As dissolved oxygen reacts with reduced GOD, this glucose oxidation also consumes oxygen in the solution, thereby producing hydrogen peroxide and oxidized GOD, and lowering oxygen pressure. As a result, by electrochemically sensing oxygen with a Clark oxygen electrode, the electrode can detect the glucose. This kind of sensor is considered a biosensor of the "first generation." The Yellow Springs Instrument Company (Ohio, USA) marketed this first-generation biosensor in the 1970s.

The second-generation biosensor substitutes the naturally-existing substrate, oxygen, with small artificial redox molecules that act as redox mediators and exchange electrons between electrodes and



2019-1-BG01-KA203-062371

enzymes. To improve the sensor efficiency, that is, sensitivity and signal-to-noise ratio, a variety of soluble redox molecules, such as ferrocene, thionine, methylene blue, methyl viologen, were used. These mediators were initially dissolved in a solution. They obtain electrons from the electrodes and then, or vice versa, these electrons are shuttled to the redox center of the enzymes. Immobilized mediators were suggested as a step forward, in order to enhance reagentless biosensors. For example, for hydrogen peroxide, Ruan et al in 1998 documented a reagentless solid-state sensor. Gold electrodes were first modified with L-cysteine, and then multilayers of horseradish peroxidase (HRP) were connected by glutaraldehyde to the amine group of cysteine, and thionine was further linked to the enzyme by the same chemical link. As a result, the gold electrode was immobilized by both the enzyme and the mediator, which could detect hydrogen peroxide sensitively in the test solution without further reagent addition. An essential advantage of this configuration of the biosensor is that the mediator is fixed on the electrode surface, thereby preventing the diffusion problem.

An alternative solution that included the use of redox polymers was stated by Heller and co-workers. First, they prepared a doped polymer with Os^{2+} complex. This type of polymer functions as a "molecular wire" and exchanges electrons between the enzyme and the electrode. The Os-polymer and glucose oxidase are then co-immobilized on the carbon electrode, which generates a sensitive response to the presence of glucose. They were able to make almost 100% immobilized enzyme molecules electroactive by using these redox polymers, which contributed to a very high sensitivity method of glucose detection.

The commercialization of the enzyme-based second-generation biosensor was quite successful. In 1987, MediSense was established and the pensized ExactechTM glucose sensors were released. This success has led to a health care revolution for diabetic patients. Instead of traveling to hospitals, they were able to control their concentration of blood glucose at home. The MediSense and later amperometric biosensor systems consist of GOD-coated disposable, screen-printed carbon electrodes and mediators (test strips). The sensor starts to function when a droplet of blood is applied to the test strip and records the amperometric response, which is transformed to a digit displayed on the LCD, indicating glucose concentration.

More recently, by designing a reconstructed GOD enzyme, Xiao et al (2003) published a new glucose biosensor generation. They first prepared apo-GOD free of the cofactor of flavin adenine dinucleotide (FAD), then functionalized and reconstructed a 1.4-nm gold nanoparticle with FAD into the apo-GOD. By using a dithiol monolayer, such a reconstructed enzyme was aligned with gold electrodes (Figure 2). They demonstrated that this artificial enzyme's electron transfer turnover is as high as 5000 s^{-1} , approximately 8-fold higher than the normal enzyme (700 s^{-1}). The gold nanoparticle in this system serves as an electron relay for the electrical wiring of the enzyme's redox center. In this area, the glucose biosensor established by Xiao et al represents a new path, free from any mediator and highly sensitive. More recently, by using single wall carbon nanotubes (SWNTs) instead of gold nanoparticles, Willner's group documented a modified version of this sensor and realized a similarly superior efficiency. Although there is still no commercialization of this technology, it is expected that state-of-the-art biosensors will be further improved.

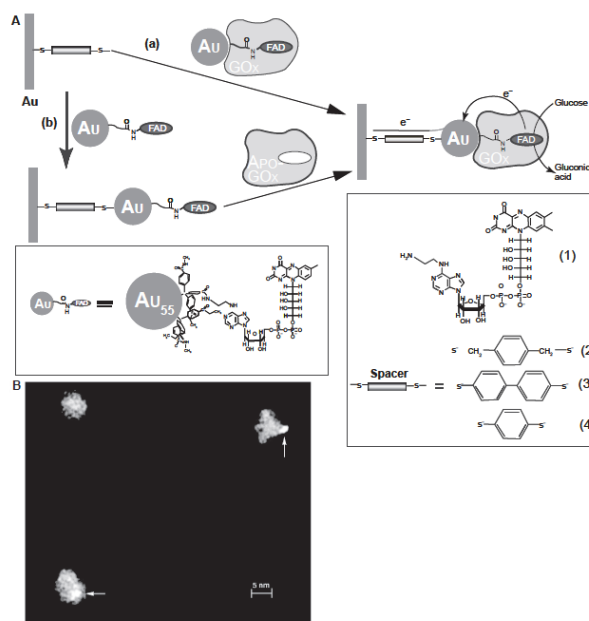


Figure 2. Construction and operation of enzyme - based biosensor

3. Immunological biosensor

To identify environmental or clinically important targets, immunological biosensors depend on a highly specific immunological system, i.e., antibodies and antigens. In reality, immunological biosensors are a modern variant of the enzyme-linked immunosorbent assay (ELISA), with lower costs, increased speed and convenience of service, and sensitivity that is comparable or even higher.

Among the most common ones are electrochemical immunological biosensors. Two types of immunological biosensors are available. First, the electrode is immobilized with a capture antibody, which catches a particular target antigen. Via a secondary antibody tagged with redox molecules or enzymes, signal transduction is achieved. Second, an electrode antigen is immobilized, which detects specific antibodies.

Ju and co-workers established a carcinoembryonic antigen (CEA) amperometric immunological biosensor. Thionine and HRP-labeled CEA antibodies were co-immobilized on a glassy carbon electrode crosslinked with glutaraldehyde. In the solution, which was coupled to the electrode reaction of thionine,

2019-1-BG01-KA203-062371

HRP catalytically reduced hydrogen peroxide, leading to a catalyzed signal. The redox center of HRP was partly blocked by catching CEA, leading to attenuation of amperometric signals.

Rusling and colleagues have recently taken advantage of SWNTs to enhance immunological biosensor performance (Figure 3).

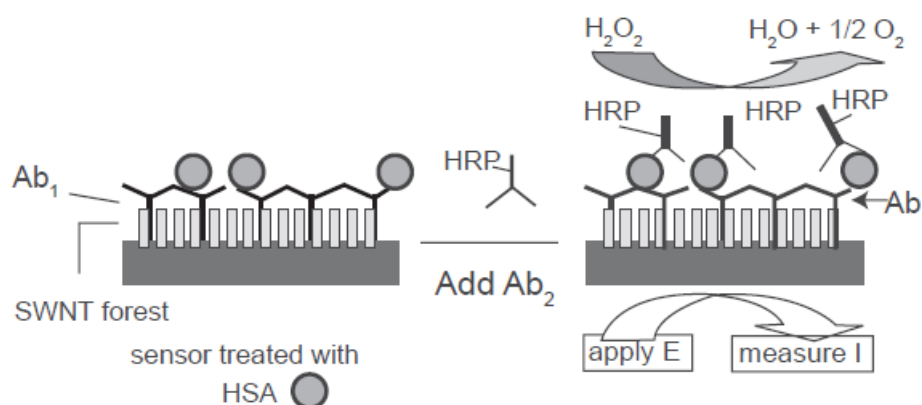


Figure 3. Construction and operation of immunological biosensor

By using metal mediated self-assembly, they designed vertically aligned arrays of SWNTs (SWNT forest) on pyrolytic graphite electrodes. Anti-HSA, using EDC/NHS, was then covalently bound to the carboxylated ends of the SWNT forest. The electrode was further incubated with a secondary HRP-labelled anti-HSA antibody after catching the HSA target. The HSA target in the test solution can be identified based on the catalytic signal of HRP for hydrogen peroxide. The detection sensitivity, which was around 1 nM, was dramatically improved by the use of SWNT forests. This was probably due to the improved reactivity of electron transfer of HRP encapsulated in SWNT forests.

One of the most relevant clinical tools has been the immunological assays. However, existing assay methods, such as ELISA, require large and costly instruments as well as well-trained specialists. For the development of cheap, miniaturized, and compact devices, electrochemical methods are well adapted. As a result, the development of electrochemical immunological biosensors to meet field and point-of-care analysis is highly desirable. It is important to mention that the use of disposable screen-printed electrodes could be crucial towards this goal, comparable to glucose biosensors. In order to carry out high-throughput (HTS) assays, it is also important to establish antibody microarrays based on electrochemistry.



2019-1-BG01-KA203-062371

4. DNA biosensors

There has been tremendous scientific and technological interest in the identification of DNA hybridization events. The rapidly growing interest in chip-based clinical diagnosis has particularly demonstrated this significance. Therefore, a variety of techniques, including optical, acoustic and electronic approaches, have been developed over the years. In past decades, fluorescent detection has dominated state-of-the-art genosensors among them. Electrochemical methods, however, which have proven effective in simple chemical species, especially metal ions, have attracted increasingly growing interest in biologically related species detection applications.

The benefits of electronic detection include: 1) electrochemical detection is typically inexpensive thus enabling highly sensitive and rapid screening; 2) several electroactive labels, e.g. metallocenes, are stable and environmentally insensitive, unlike fluorophores that often have "photo-bleaching" problems; 3) appropriate molecular design and synthesis that generate a variety of derivatives, each with a specific redox potential, have made it possible to label 'multi-color'; 4) The rapidly established silicon industry has paved the way for the mass production of integrated circuits, making electronic detection particularly appropriate and compatible with microarray-based technologies; 5) The exponential growth of interfacial science and technology has unravelled mysteries in the precise control of surface properties that are one of the key barricades in bioelectronic applications.

At moderate applied voltages, DNA itself is electrochemically silent, while significant interferences are predicted at high voltages that cause DNA bases to be oxidized/reduced (. Millan was the first to suggest sequence-selective DNA target detections based on electroactive hybridization indicators that provide electronic signals and double and single-stranded DNA discrimination. "Sandwich" type detections were suggested in an effort to reduce the high background derived from the minor binding of hybridization indicators to ssDNA. A DNA strand possessing an electroactive label has been introduced to act as the signaling molecule in addition to an immobilized DNA probe. Similarly, with nanoparticle probes, Park and colleagues in 2002 have developed an array-based electrical DNA detection that demonstrates high sensitivity and selectivity. A technology based on the relatively high oxidation activity of guanine and its cooperation by exogenous redox catalysts was developed by Thorp one year later. The discrimination of ds/ss is obtained by the fact that guanine has relatively low electron transfer reactivity in duplexes, due to the steric effect. In the detection of PCR products, this method is highly sensitive, but relatively poor in discriminating hybridization events. Moreover, this method is only possible on ITO surfaces so far, because the high oxidation potential still excludes the use of gold.

Despite progress, the development of an all-in-one (i.e. reagentless) sensor that specifically signals target capture is still highly important (i.e., obviating further treatment with either signal molecules or hybridization indicators). A viable means to this end is provided by DNA or RNA aptamers. Aptamers are well-structured DNA or RNA that have high affinity and selectivity for particular targets as well as natural enzymes, thus showing superior robustness to fragile enzymes. They have been a very promising tool for therapy and diagnosis. Until then, for virtually any given target, the well-developed *in vitro* selection was able to produce aptamers. In view of these benefits, oligonucleotide aptamers are

2019-1-BG01-KA203-062371

predicted to be the biosensing components of the next generation. An easy, organized hairpin-like DNA with an electroactive label (electronic DNA hairpin) was used by Fan and colleagues as the building block to define hybridization events (Figure 4).

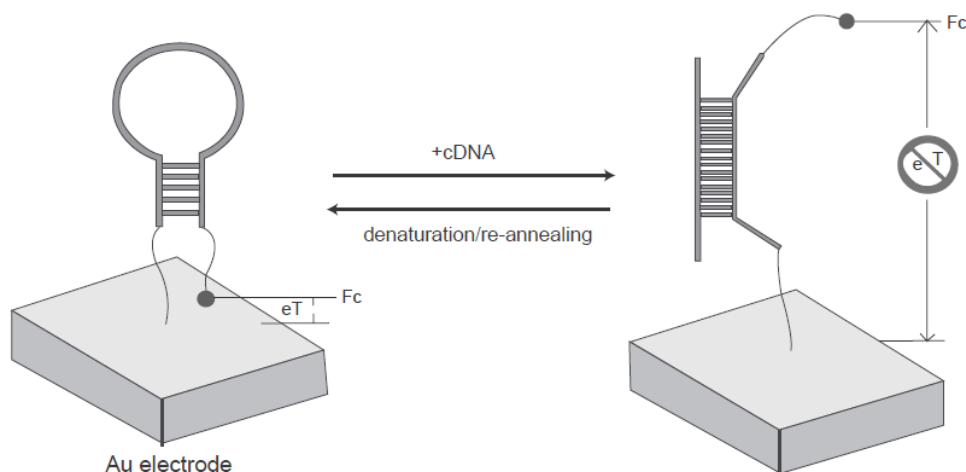


Figure 4. Construction and operation of DNA biosensor

Hairpin-like DNA was an incredibly fascinating aptamer that forms the basis of homogeneous hybridization recognition of fluorescent “molecular beacons”. The DNA sequence has been designed so that in the absence of targets, this “beacon” is in the near state while it will be “turned on” when it reaches its particular gene target. The presence of the design of the stem-loop in the structure offers an on/off switch as well as a stringency to differentiate against single DNA hybridization mismatch. A thiolated terminus gives a sticky end to the gold surface of this electronic DNA hairpin, while a ferrocene tag transduces electronic signals at the other end. The initial hairpin localizes the ferrocene proximal to the electrode surface, thus allowing interfacial electron transfer. After hybridization, the formation of the linear duplex structure disrupts the hairpin and forces apart the ferrocene and the electrode. This significant distance change (up to a few nm) effectively blocked the interfacial electron transfer and leads to the diminution of corresponding electrochemical current signals. This strategy offers the opportunity to identify 10 pM DNA targets. More importantly, such a design takes advantage of integrating within a single surface-confined hairpin structure the capturing part (probe sequence) and the signaling part (electroactive species). In contrast to most previously proposed solid-state DNA sensors, this design is therefore effectively reagentless, i.e. no exogenous reagent is required during the recognition process



2019-1-BG01-KA203-062371

apart from DNA targets. This provides the basis for the development of a portable, continuous DNA analyzer that may be useful for medical and military applications.

A medium for long-range electron transfer (ET) via its base stacking was suggested as the DNA double helix. Although this issue has been discussed for a long time, Barton and colleagues have proven electrochemically that well-oriented gold electrode DNA films enable long-range electron transfer and that such ET is highly sensitive to base stacking perturbations like mismatches. They found that electroactive intercalators like methylene blue (MB) could be effectively reduced by a fully matched DNA duplex-modified electrode. The existence of only a single mismatch, however, converts the wire-like ET medium into an insulator, totally disrupting the ET between the MB and the electrode. Via cyclic voltammetric or coulometric assays, which form the basis of a rapid DNA mutation screening sensor, such a difference can be easily read. Barton and colleagues also demonstrated that electrocatalysis could improve the sensitivity of this strategy. In solution, the addition of ferricyanide constantly pulls electrons from electrochemically reduced MB, amplifying electron flow through the double helix of DNA. This helps in detecting ~10⁸ molecules of DNA with a 30- μ m electrode. They have also developed DNA-based sensors to detect DNA-binding proteins in parallel with DNA detection. Some DNA-binding proteins or enzymes are believed to interact with the base pair stacking of DNA, transforming the double helix of DNA from effective ET wires to insulators. They established a sensitive way to electrically assay a variety of DNA-binding proteins based on the comparable sensing strategy. Crucially, these sensors discriminate successfully against proteins that bind to DNA, but they do not disrupt base stacking. This certainly confirms that the signal cut-off on protein binding is due to the alteration of the ET medium relevant to base stacking.

5. The future of clinical biosensors

Despite the rapid improvement in the development of biosensors, clinical applications of biosensors are still uncommon, with an exception being the glucose monitor. This is in direct contrast to the critical need for point-of-care testing in small clinics. We assume the specifications below are relevant. First, high sensitivity: Improvement of sensitivity is an ever-lasting priority in the development of biosensors. It is clear that the sensitivity criterion ranges from case to case. For example, because glucose levels are high in the blood, one does not need very high sensitivity for glucose detection. This is basically part of the reason why glucose monitors have been successful. However, in many situations, in order to meet the requirements of molecular diagnostics and pathogen detection, it is very important to establish highly sensitive biosensors with optimum single-molecule detection. Second, high selectivity: In the application of biosensors, this may be a significant barricade. Most of the biosensors mentioned in the literature function very well in laboratories, but in real test samples, serious problems can be addressed. As a result, in order to prevent non-specific surface adsorption, it is important to establish novel approaches to surface modification. Third, multiplexing is crucial for saving assay time, which is particularly important for laboratory or clinic assays. It is therefore important to establish arrays



2019-1-BG01-KA203-062371

of high-density electrodes as well as electrochemical instruments that can conduct a large number of assays simultaneously. Forth, in order to increase portability, it is essential to create miniaturized biosensors, thus satisfying the field and point-of-care test requirements. Fifth, it is appropriate to integrate and highly automate an ideal biosensor. A solution to this goal is offered by current lab-on-a-chip technologies (microfluidics). We can expect all these features to be integrated into successful biosensors in the future, and can easily detect minute targets within a short period of time.

BLUEPRINTS FOR BIOCHIPS: DESIGN & OPERATION

The term “biochip” has taken on several meanings. Any device or component introducing biological materials, either extracted from biological species or synthesized in a laboratory on a solid substrate can be considered to be a biochip in the most generic sense. In practical terms, however, both miniaturizations, usually in microarray format, and the possibility of low-cost mass production are often involved in biochips. The electronic nose or artificial nose chip, the electronic tongue, the Polymerase chain reaction chip, the DNA microarray chip (gene chip), the protein chip and the biochemical lab-on-a-chip are some examples which meet these qualifications. In the gene chip and the protein chip, the most dynamic biochip research has been done.

Much attention has been paid, in particular, to biochips integrating conventional biotechnology with semiconductor processing, micro-electro-mechanical systems (MEMS), optoelectronics and digital signal and image acquisition and processing.

Thousands of genes and their derivatives (i.e., RNA and proteins) in a given living organism are generally assumed to act in a complicated and coordinated manner that creates the mystery of life. Traditional methods in molecular biology, however, typically operate on a "one gene in one experiment" basis, which assumes that the throughput is very limited and it is difficult to achieve the "whole picture" of gene function. A new technology, called the DNA microarray, has gained considerable attention among biologists in the last decades. This technology aims to measure the whole genome on a single chip so that, meanwhile, researchers can get a better picture of the interactions between thousands of genes.

A gene or DNA chip corresponds to a two-dimensional array of small reaction cells (100 x 100 μm each) produced using high-speed robotics on a solid substrate. A silicon wafer, a thin sheet of glass, plastic, or a nylon membrane could be the solid substrate. Trillions of polymeric molecules from a particular sequence of single strand DNA fragments are immobilized in each reaction cell (Figure 5).

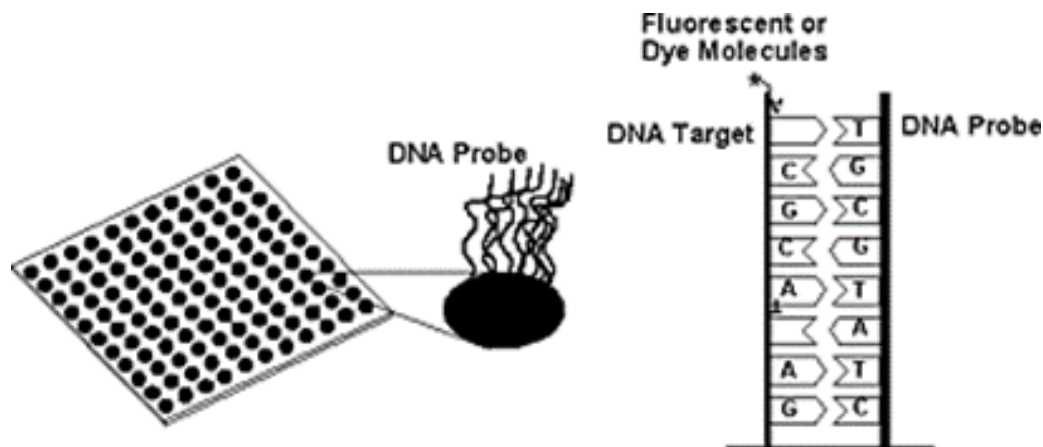


Figure 5. A schematic illustration of a gene chip

DNA fragments may be either short (about 20 to 25) base sequences (A, T, G, and C) or longer complementary DNA strands (cDNA). In each cell, the unique sequence of bases (e.g. CTATGC...) is preselected or configured depending on the expected use. The probes are also called recognized sequences of single-strand DNA fragments immobilized on the substrate. Double-strand DNA fragments are formed when unknown fragments of single-strand DNA samples, called the target, react (or hybridize) with the probes on the chip, where the target and the probe are complementary according to the base pairing rule (A paired with T, and G paired with C). The target samples are often labeled with tags, such as fluorescents, dyes, or radio-isotope molecules, to facilitate the diagnosis or analysis of the hybridized chip. Each is labeled with its own distinguishable tag when the targets contain more than one type of sample. This type of DNA microarray chip provides a platform where, based on the size of the array, the unknown target or targets can theoretically be defined with very high speed and high throughput by matching the components involved in the research and development of biochip technology with tens of thousands of different types of probes through hybridization in parallel, and the related technical disciplines are indicated in Figure 6. Fundamentally, biochip technology is interdisciplinary; it is important that scientists and engineers from different disciplines cooperate synergistically to push this novel technology from a lab interest to practical devices and systems.

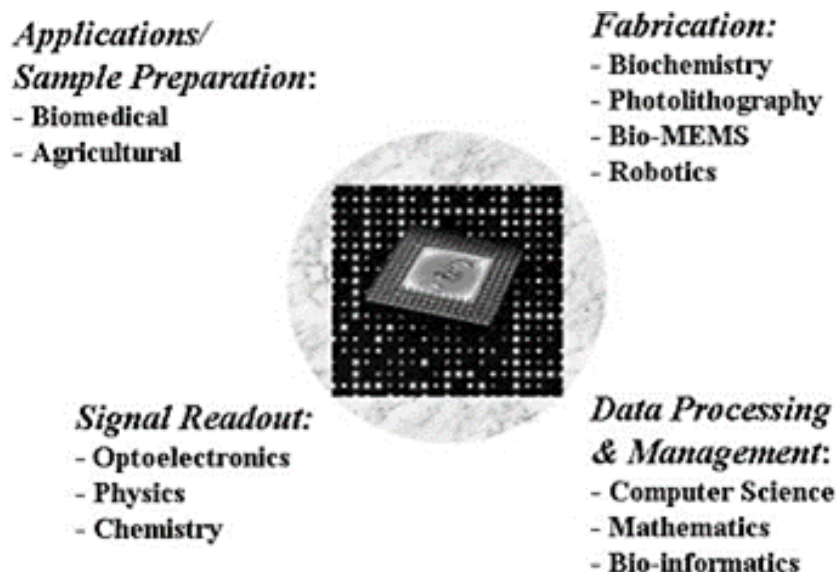


Figure 6. The components and the associated technical disciplines involved in the R&D of biochip technology

1. DNA microarray

With regard to the property of the arrayed DNA sequence of known identity, there are two variants of the DNA microarray technology:

Type I: probe cDNA (500~5,000 bases long) is immobilized using robot spotting to a solid surface such as glass and exposed either separately or in a mixture to a set of targets. This strategy, “traditionally” referred to as DNA microarray, is largely known to be developed at Stanford University.

Type II: *in situ* (on-chip) or by conventional synthesis accompanied by on-chip immobilization, an array of oligonucleotide (20~80-mer oligos) or peptide nucleic acid (PNA) probes is synthesized. The array is exposed to hybridized, labeled sample DNA and determines the identity/abundance of complementary sequences. This method, “historically” named DNA chips, was established at Affymetrix INC., which sells its photolithographically manufactured products under the Genechip trademark. Oligonucleotide-based chips are developed by several companies using alternative *in-situ* synthesis or deposition technologies.

Depending on the type of molecule immobilized, biochip is made primarily in two formats. CDNA arrays are also referred to as biochips containing PCR products of 200 base pairs to 2KB size immobilized along the length of the molecule by covalently cross linking to the surface of the array. Alternatively, oligonucleotide probes may either be synthesized *in situ* on the array, or covalent links to

2019-1-BG01-KA203-062371

the termini may be fixed by pre-synthesized oligos. Genechip engineering includes several different parts, such as manufacturing, sample preparation and hybridization of the target sequence, hybridization results detection, design of the oligonucleotide probe, and hybridization image analysis, and various applications, as seen in Figure 7.

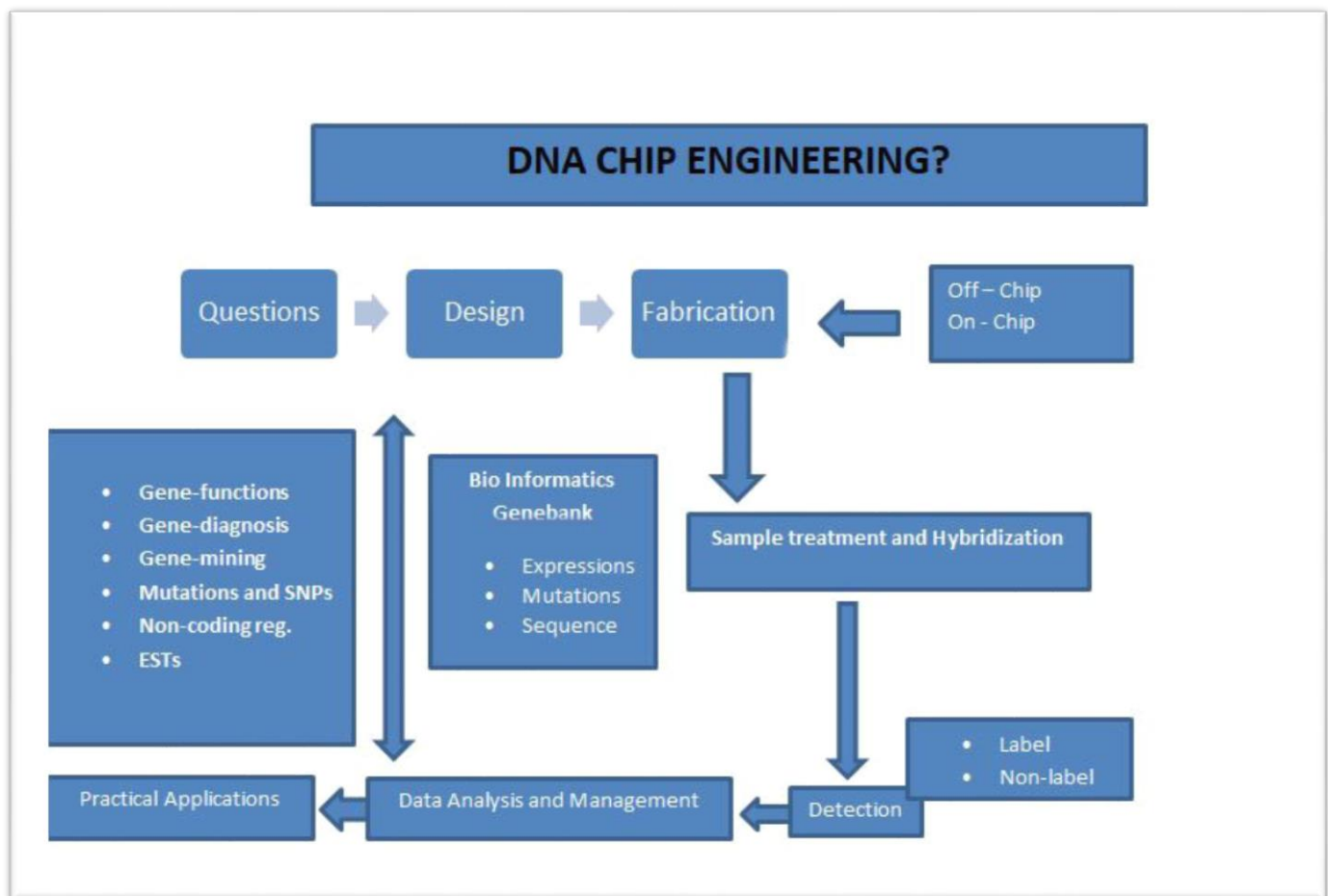


Figure 7. Several important aspects related to Genechip technology

First, according to a particular target, many relevant gene sequences will be selected from the DNA database (single nucleic polymorphism for a specific genes, differential expression for a given group of genes or mutation identification). By determining the sequence and length of each probe and its exact position on the chip, a series of unique oligonucleotide probes will be designed based on the



2019-1-BG01-KA203-062371

selected sequences. With the spotting method or on-chip synthesis, the synthesis of the DNA microarray can be carried out. The target genes are usually necessary for fluorescence to be amplified and labeled. Selection of appropriate PCR primes, optimization of amplification and hybridization conditions will be needed in most cases. For the hybridization outcomes on a genechip, there are several different detection strategies. A traditional approach is fluorescent detection. For the treatment of such a large amount obtained from a genechip, data analysis based on the fluorescent images and database configuration is required. The use of PCR products corresponding to the genes as probe molecules is a common platform for preparing microarrays. Biological sources of cDNA libraries offer an effective template for PCR amplification of the probes. This platform is therefore referred to as cDNA Arrays.

The use of oligonucleotide probes instead of PCR products has several benefits. First, they are typically of similar length and can be created in such a way that they have similar properties of hybridization. Second, they may be configured to hybridize against the specific gene region; the PCR product also allows cross-hybridization between homologous genes disturbing members of the same gene family's gene expression profiles. Furthermore, the need for tedious PCR amplification of the probe molecules is eliminated by oligonucleotide arrays and decreases the risk of error due to clone handling and contamination during transfer.

2. Microfluidic chip

In recent decades, microfluidic technology has been rapidly developed and provides multitudinous applications in the life sciences. Thanks to the distinct benefits provided by system miniaturization, the microfluidics revolution arose, including high analytical efficiency, increased sensitivity, improved analytical performance, fast multiplexing parallelization, the ability to manage and process reduced reagent volumes and dramatically reduced instrumental footprints.

Microfluidics could simultaneously offer analytical efficiency and high throughput capability as a miniaturization technology, without the lack of accuracy and automation. Microfluidics technology is not only a powerful tool for the fast screening and study of drug development during the drug application process, but also for its miniaturized devices to lower costs and reagent consumption.

Significant progress has been made in developing drug screening components and systems based on microfluidics over the past few years. For drug screening, different types of microfluidic chips are used to improve screening efficiency and decrease costs. Several common types of chip technology are described in the following paragraphs.

2019-1-BG01-KA203-062371

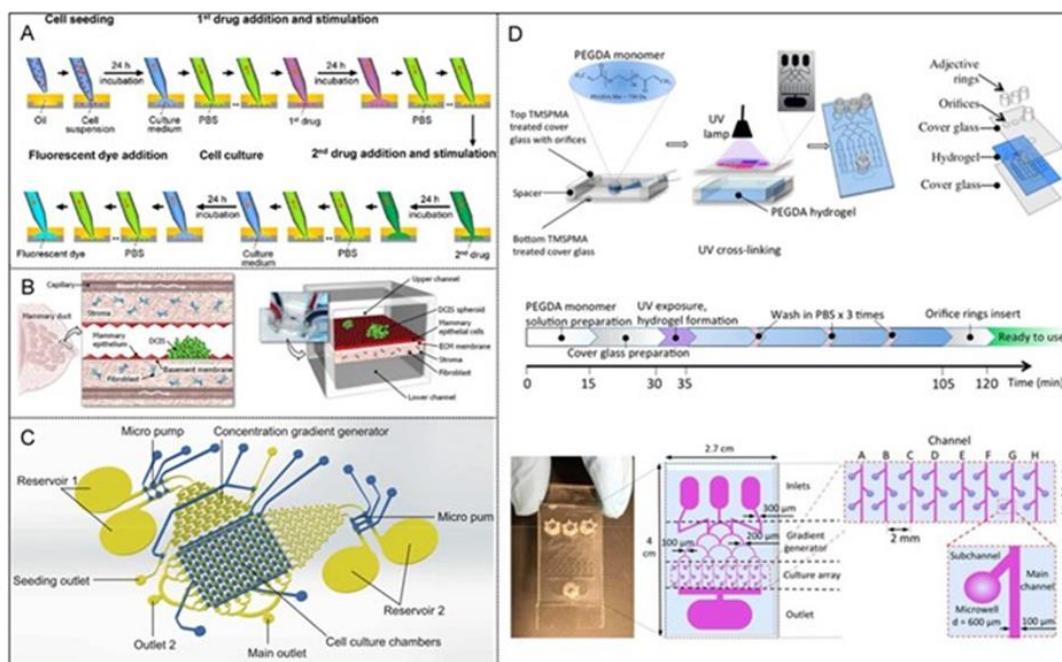


Figure 8. Application of microfluidic chip in drug screening

Droplet microfluidics – In order to perform experiments in continuous or segmented flow, droplet microfluidic technology utilizes liquid droplets compartmentalized by an immiscible fluid as nanoliter to picoliter separate reaction vessels. Such methods show significant advantages, such as reduced sample usage, improved reaction speed and increased efficiency and reproducibility, handling incredibly small volumes with robust composition control.

Based on the sequential process droplet array technique, droplet microfluidic methods may be used for drug combination screening (Figure 8A), to screen various dosing combinations and administration lengths, and to refine the optimum minimally consumed dosing regimen that is important for combined disease situations.

Organ-on-chip – The detailed regulation of microscale structure and flow enables the precise modeling of the microscale structures of organ tissue to be built. Organs-on-chips are biomimetic systems that are micro-engineered and represent essential functional units of living human organs. The functions of multiple organs and tissues, such as the liver, kidneys, lungs and intestines, have been replicated as in vitro models to date. These systems could be used as in vitro models that enable complex biological processes to be simulated and pharmacologically modulated.



2019-1-BG01-KA203-062371

The organ chip simulates the human body's basic processes, by using a number of cells to create a biomimetic chip with similar physiological functions on the chip's unique structure, which is comparable to the disease's actual external environment than the typical single-cell culture model.

In order to replicate the complex microarchitecture of cancerous tissue, a microsystem that allows co-culture of breast tumor spheroids with neighboring cells in a compartmentalized 3D microfluidic system has been established to help create the anti-breast cancer drug screening platform (Figure 8B). For the study of cancer cell migration and anti-cancer drug screening, co-culture of multiple tissues with a microfluidic system could be used.

Other microfluidic chips for drug screening – There are several other technologies applied to microfluidic chips for drug screening, in addition to the above-mentioned existing methods, which extend the ideas of researchers. The most significant and integrated aspect of drug development in most pharmaceutical and several biotechnology industries around the world has been an outstanding HTDS system.

With the use of open-access microfluidic tissue array systems and cell microarray chips, various concentrations and combinations of drugs can be screened. Various configurations and arrays of small culture chambers may be generated by the different design of the chip. Concentration gradient generators can provide an effective liquid concentration gradient, and these devices have been implemented and combined with microfluidic technologies by several research groups to optimize HTDS systems. The diffusive microfluidic mixers in Figure 8C could also recognize a fully automatic HTDS.

Cell array platforms are built using polydimethylsiloxane (PDMS) material in certain microfluidic channels for drug screening. The structure of the chips, though, is difficult and has some disadvantages, such as expensive silicon molds and biomolecular absorption. As a natural extracellular matrix (ECM), poly (ethylene glycol) diacrylate (PEGDA) hydrogel has identical mechanical properties and water content. PEGDA microfluidic hydrogels have been commonly used for cell encapsulation and are permeable to compounds such as water, biomolecules and chemicals. In order to research the combinatorial treatment effect of two drugs, microfluidic devices made of these types of materials and combined with 3D brain cell culture technologies were used (Figure 8D).

The allure of the microfluidic chip is that to perform various functions, it can have multiple designed structures and can be combined to extend its application range with different devices and testing equipment. However, considerable design, manufacturing and optimization efforts are needed. Each model has unique features of its own. It is used not only for drug screening, but for drug testing as well.



3. Protein microarray

In the research of proteins, protein microarrays are useful tools in an unbiased, high-throughput manner, as they allow up to thousands of individually purified proteins to be characterized in parallel. This technology's adaptability has made it possible to be used in a wide range of applications, including the study of proteome-wide molecular interactions, the analysis of post-translational alterations, the discovery of new drug targets, and the evaluation of pathogen-host interactions. In addition, the technology has already been shown to be effective in profiling the specificity of antibodies, as well as in identifying new biomarkers for autoimmune diseases and cancers in particular.

Proteins are complex biomolecules with a large spectrum of structures and functions, and as such, studying them in a high-throughput manner is a challenge. Three primary types of protein microarrays exist: functional, analytical, and reverse phase. Functional protein microarrays are assembled in a high-throughput manner with proteins purified/synthesized, allowing hundreds and even thousands of different proteins to be examined in parallel with their biochemical properties. In order to detect or measure complex biological samples, analytical protein microarrays use affinity reagents immobilized on the array. Finally, complex biological samples immobilized on the array are used by reverse phase protein microarrays which use affinity reagents for identification.

Functional protein microarrays are more capable of identifying weak interactions, more flexible for low-abundance proteins, and more capable of analyzing crude samples such as serum, compared to other methods, such as mass spectrometry. In terms of differences in proteome coverage, protein lengths, and production pipelines, several different types of functional protein microarrays have been produced to date.

Development of the Functional Protein Microarray – The whole set of proteins that can be expressed by a genome is the proteome. Typically, the creation of a purified proteome microarray requires the assembly of a genome-wide set of open reading frames (ORFs) cloned into an expression vector, encoded protein expression in cells, high-throughput individual protein purification, and protein immobilization on a microarray.

Application of Yeast Proteome Microarrays – For the profiling of proteome-wide molecular interactions, functional protein microarrays, especially purified proteome microarrays, are useful and enable detailed, unbiased screening. Researchers have used functional protein microarrays in fundamental research to study protein-protein interactions, protein-lipid interactions, protein-DNA interactions, protein-cell/lysates, small molecule binding, protein-RNA interactions, and PTMs, such as acetylation, SUMOylation, glycosylation, ubiquitylation, phosphorylation, and methylation (Figure 9A-9G). We summarize representative studies in Table I based on the research applications seen in Figure 9.

2019-1-BG01-KA203-062371

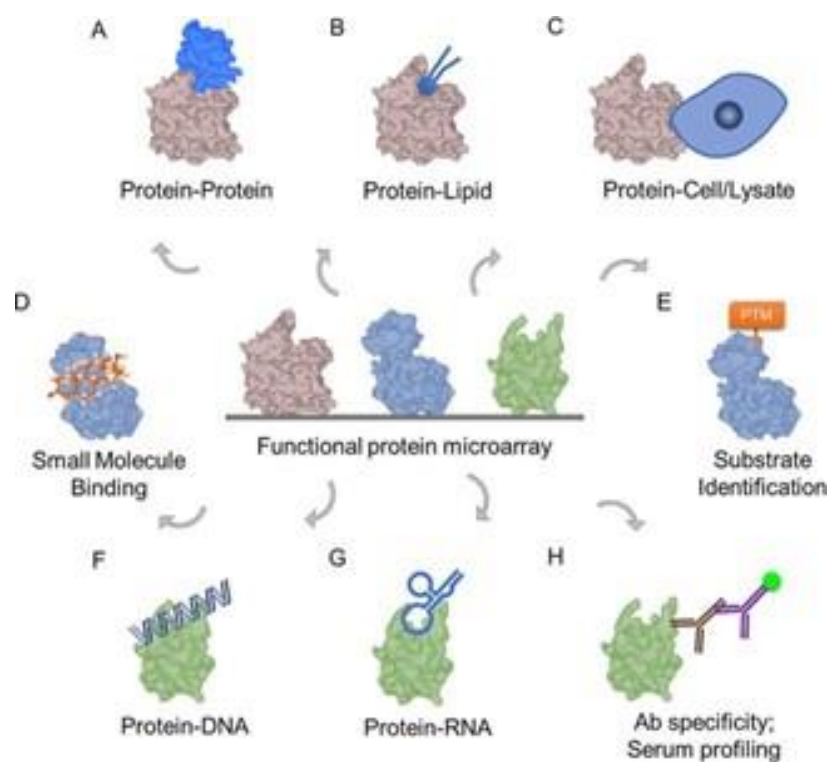


Figure 9. Application of Functional Protein Microarray

4. Applications of biochips

The most emerging applications of biochips are included in Table 1.

Table 1. Applications of biochips

Biochip in food	Biochip detecting genetically modified organisms (gmo's) in food.
Biochip in diagnostics	DNA biochip which revolutionise the way the medical profession performs tests on blood.
Biochip in Tuberculosis epidemic	Biochip technology expected to help combat the new variety of drugresistant strains of the disease.
Biochip in cancer	Biosensor chip technology provides quick and easy access to critical information regarding DNA damage from cancer-producing compounds and aid in early detection of colon cancer.
Biochip in drug development	DNA chips which find genetic differences between people who respond to a drug and those who do not, starting in Phase II, or mid-stage, clinical trials.



References

- Adams DA, Brus L, Chidsey CED, et al. 2003. Charge transfer on the nanoscale: current status. *J Phys Chem B*, 107: 6668-97.
- Ashley GW, Henise J, Reid R, et al. 2013. Hydrogel drug delivery system with predictable and tunable drug release and degradation rates. *Proc. Natl. Acad. Sci. USA*, 110: 2318–2323.
- Bard AJ, Faulkner LR. 2001. *Electrochemical Methods*. New York: John W Willey & Sons.
- Benson DE, Conrad DW, de Lorimier RM, et al. 2001. Design of bioelectronic interfaces by exploiting hinge-bending motions in proteins. *Science*, 293:1641–4.
- Boon EM, Ceres DM, Drummond TG, et al. 2000. Mutation detection by electrocatalysis at DNA-modified electrodes. *Nature Biotech*, 18:1096–100.
- Boon EM, Livingston AL, Chmiel NH, et al. 2003. DNA-mediated charge transport for DNA repair. *Proc Natl Acad Sci U S A*, 100:12543–7.
- Boon EM, Salas JE, Barton JK. 2002. An electrical probe of protein-DNA interactions on DNA-modified surfaces. *Nat Biotechnol*, 20:282-6.
- Brazill SA, Kim PH, Kuhr WG. 2001. Capillary gel electrophoresis with sinusoidal voltammetric detection: A strategy to allow four-“color” DNA sequencing. *Anal Chem*, 73:4882–90.
- Burgstaller P, Girod A, Blind M. 2002. Aptamers as tools for target prioritization and lead identification. *Drug Discov Today*, 7:1221–8.
- Choi Y, Hyun E, Seo J, et al. 2015. A microengineered pathophysiological model of early-stage breast cancer. *Lab Chip*, 15: 3350–3357.
- Cooper MA, Dultsev FN, Minson T, et al. 2001. Direct and sensitive detection of a human virus by rupture event scanning. *Nature Biotechnol*, 19:833–7.
- Dai Z, Yan F, Yu H, et al. 2004. Novel amperometric immunosensor for rapid separation-free immunoassay of carcinoembryonic antigen. *J Immuno Methods*, 287:13–20.
- Das, H.K. 2005. “Functional Genomics using Microarrays Technology.” Text book of Biotechnology, pp .1276-1288, Wiley Dreamtech Publisher.
- Drummond TG, Hill MG, Barton JK. 2003. Electrochemical DNA sensors. *Nat Biotechnol*, 21:1192–9.



2019-1-BG01-KA203-062371

Fan C, Plaxco KW, Heeger AJ. 2003. Electrochemical interrogation of conformational changes as a reagentless method for the sequence-specific detection of picomolar DNA. *Proc Natl Acad Sci U S A*, 100:9134–7.

Fan C, Plaxco KW, Heeger AJ. 2005. Biosensors based on binding-modulated donor-acceptor distances. *Trends Biotechnol*, 23:186–92.

Fan Y, Nguyen DT, Akay Y, et al. 2016. Engineering a brain cancer chip for high-throughput drug screening. *Sci. Rep.*, 6: 25062.

Fritz J, Cooper EB, Gaudet S, et al. 2002. Electronic detection of DNA by its intrinsic molecular charge. *Proc Natl Acad Sci U S A*, 99:14142–6.

Gao Z, Binyamin G, Kim H-H, et al. 2002. Electrodeposition of redox polymers and co-electrodeposition of enzymes by coordinative crosslinking. *Angew Chem Int Ed*, 41:810–13.

Gaylord BS, Heeger AJ, Bazan GC. 2002. DNA detection using watersoluble conjugated polymers and peptide nucleic acid probes. *Proc Nat Acad Sci U S A*, 99:10954.

Griffiths AD, Tawfik DS. 2000. Man-made enzymes – from design to in vitro compartmentalisation. *Curr Opin Biotech*, 11:338–53.

Heeger AJ. 2000. Nobel Lecture: Semiconducting and Metallic polymers: The fourth generation of polymeric materials [online]. URL: [http:// wwwnobelse](http://www.nobel.se).

Hook F, Ray A, Norden B, et al. 2001. Characterization of PNA and DNA immobilization and subsequent hybridization with DNA using acoustic- shear-wave attenuation measurements. *Langmuir*, 17:8305–12.

Hsiung LC, Chiang CL, Wang CH, et al. 2011. Dielectrophoresis-based cellular microarray chip for anticancer drug screening in perfusion microenvironments. *Lab Chip*, 11: 2333–2342.

Li Z, Su W, Zhu Y, et al. 2017. Drug absorption related nephrotoxicity assessment on an intestine-kidney chip. *Biomicrofluidics*, 11: 034114.

Lin D, Li P, Lin J, et al. 2017. Orthogonal screening of anticancer drugs using an open-access microfluidic tissue array system. *Anal. Chem.*, 89: 11976–11984.

Liu J, Zhang Y, Jiang M, et al. 2017. Electrochemical microfluidic chip based on molecular imprinting technique applied for therapeutic drug monitoring. *Biosens. Bioelectron.*, 91: 714–720.

Moore CD, Ajala O.Z, Zhu H. 2016. Applications in high-content functional protein microarrays. *Curr. Opin. Chem. Biol.*, 30: 21–27.

- Palecek E, Jelen F. 2002. Electrochemistry of nucleic acids and development of DNA sensors. *Crit Rev Anal Chem*, 32:261–70.
- Palecek E. 2004. Surface-attached molecular beacons light the way for DNA sequencing. *Trends Biotechnol*, 22:55–8.
- Park SJ, Taton TA and Mirkin CA. 2002. Array-based electrical detection of DNA with nanoparticle probes. *Science*, 295:1503–6.
- Patolsky F, Lichtenstein A, Willner I. 2001. Detection of single-base DNA mutations by enzyme-amplified electronic transduction. *Nature Biotech*, 19:253–7.
- Patolsky F, Weizmann Y, Wilner I. 2004. Long-range electrical contacting of redox enzymes by SWCNT connectors. *Angew Chem Int Ed*, 43:2113–17.
- S. Mi, Z. Du, Y. Xu, et al. 2016. Microfluidic co-culture system for cancer migratory analysis and anti-metastatic drugs screening. *Sci. Rep.*, 6: 35544.
- Schuster GB. 2000. Long-range charge transfer in DNA: transient structural distortions control the distance dependence. *Acc Chem Res*, 33:253-60. Sullivan CKO. 2002. Aptasensors—the future of biosensing? *Anal Bioanal Chem*, 372:44–8.
- Stone HA, Stroock AD, Ajdari A. 2004. Engineering flows in small devices: microfluidics toward a lab-on-a-chip. *Annu. Rev. Fluid Mech.*, 36: 381–411.
- Sugiura S, Hattori K, Kanamori T. 2010. Microfluidic serial dilution cell-based assay for analyzing drug dose response over a wide concentration range. *Anal. Chem.*, 82: 8278–8282.
- Taton TA, Mirkin CA, Letsinger RL. 2000. Scanometric DNA array detection with nanoparticle probes. *Science*, 289:1757–60.
- Thorp HH. 2003. Reagentless detection of DNA sequences on chemically modified electrodes. *Trends Biotechnol*, 21:522–4.
- Umek RM, Lin SW, Vielmetter J, et al. 2001. Electronic detection of nucleic acids—A versatile platform for molecular diagnostics. *J Mol Diag*, 3:74–84.
- Van Hove AH, Antonienko E, Burke K, et al. 2015. Temporally tunable, enzymatically responsive delivery of proangiogenic peptides from poly (ethylene glycol) hydrogels. *Adv. Healthc. Mater.*, 4: 2002–2011.
- Whitesides GM, Grzybowski B. 2002. Self-assembly at all scales. *Science*, 295:2418–21.
- Willner I. 2002. Biomaterials for sensors, fuel cells, and circuitry. *Science*, 298:2407.



2019-1-BG01-KA203-062371

Wosnick JH, Swager TM. 2000. Molecular photonic and electronic circuitry for ultra-sensitive chemical sensors. *Curr Opin Chem Biol*, 4:715–20.

Xiao Y, Patolsky F, Katz E, et al. 2003. Plugging into enzymes: nanowiring of redox enzymes by a gold nanoparticle. *Science*, 299:1877–81.

Xu H, Wu H, Huang F, et al. 2005. Magnetically assisted DNA assays: High selectivity using conjugated polymers for amplified fluorescent transduction. *Nucleic Acids Res*, 33:e83.

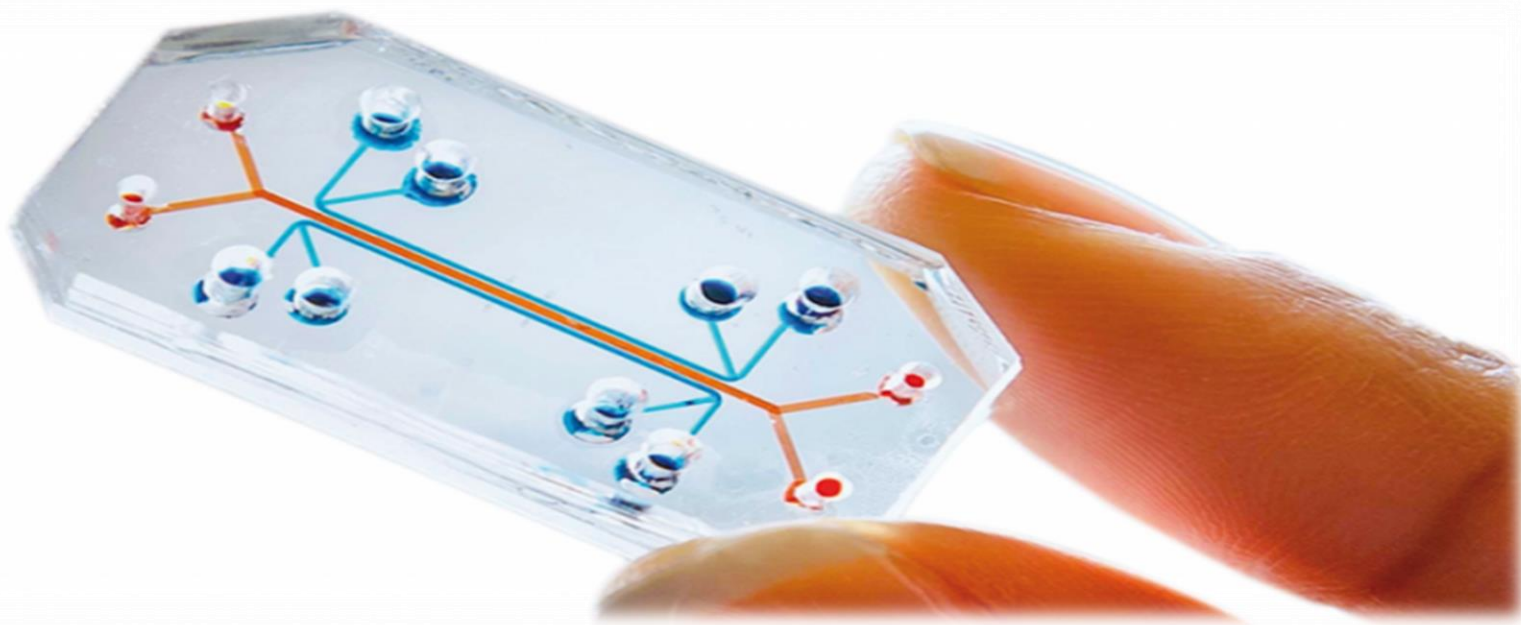
Yu CJ, Wan YJ, Yowanto H, et al. 2001. Electronic detection of single-base mismatches in DNA with ferrocene-modified probes. *J Am Chem Soc*, 123:11155–61.

Yu X, Kim SN, Papadimitrakopoulos F, et al. 2005. Protein immunosensor using single-wall carbon nanotube forests with electrochemical detection of enzyme labels. *Mol Biosyst*, 1:70–8.



Project website: www.digit-biotech.eu

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



LO3 BIOSENSORS & BIOCHIPS FOR SUSTAINABLE FUTURE

Advanced level

AUTHOR:

TSVETINA MIHAILOVA & KLIMENT PETROV



Contents

Biosensors & Biochips Technologies: Contribution to the Future Sustainable Life.....	3
Biosensors & biochips: advances in medical diagnostics	4
1. Glucose biosensors in diabetic management.....	6
2. Cardiovascular disease detection using biosensors	7
3. Biosensor for detection of cancer.....	7
4. Biochip in diagnostics	8
5. Biochip in Tuberculosis epidemic.....	8
6. Biochip in cancer.....	9
Biosensors & Biochips applied in food and agriculture.....	9
1. Nanomaterials in biosensing technology.....	12
2. Nutrient and quality detection	12
3. Detection of pathogens.....	13
4. Detection of toxins	14
Biosensors & Biochips for environmental monitoring.....	16
1. Pesticides.....	17
2. Pathogens	18
3. Potentially toxic elements	19
4. Toxins.....	20
5. Endocrine disrupting chemicals	20
6. Other environmental compounds	21
References	21



Biosensors & Biochips Technologies: Contribution to the Future Sustainable Life

The field of synthetic biology has exploded over the past decade, having a major influence on fields such as metabolic engineering, protein engineering, digital biology, and whole-genome engineering. In the framework of iterative "design-build-test" development cycles, a significant portion of synthetic biology innovation has taken place. In the field of synthetic biology, progress can be associated with innovations in each of the processes of "design", "build" and "test". For example, there has been a major push to standardize elements within synthetic biology, with significant attention being paid to modularity and "plug and play" components. This modularization, along with the accelerated progress in systems biology, has allowed the "design" stage to become less time-consuming and less reliant on advanced knowledge. In recent years, the cost of DNA sequencing and synthesis has also decreased dramatically, allowing large constructs to be synthesized cheaply. In the 'build' phase, this has facilitated a rapid improvement, helping researchers to investigate a larger percentage of the space of the biological solution. Finally, within the synthetic biology "test" phase, high-throughput screening has also become a focal point. The increased "design" and "build" potential has contributed to an increased demand for success in the assessment of the plethora of new designs. In turn, this was done by incorporating robots and high-throughput analytics into the laboratory setting, in which new models can be evaluated to a level that is not achievable for human researchers.

Biosensors represent a groundbreaking emerging technology for high-throughput screening that can be implemented. Most precisely, they are classified as an analytical tool consisting of biological components used to detect and generate a signal for the presence of a target ligand. Synthetic biology is at the forefront of biosensors, both as a tool for high-throughput screening, but also as the direct result of developments within the field of synthetic biology itself. In addition, because of the unparalleled specificity and sensitivity that biological parts provide relative to conventional analytical methods, biosensors have gained expanded interest as alternatives to traditional analytics.

The design and construction of biosensors is a multidisciplinary endeavor and can include expertise in areas such as protein engineering, molecular biology, affinity chemistry, molecular dynamics of nucleic acid, materials sciences, and nanotechnology. Biosensors interface with a target ligand at their most simple stage, undergo some type of modification, and output a signal. There is a great variety of potential configurations in all the parts of this process. Target ligands range from single atoms such as calcium, to entire proteins such as thrombin, all the way through. Processes as varied as enzymatic activity, fluorescence, electrical current generation, and transcriptional activity include output signals. The mechanisms that transduce ligand recognition into functional signals are just as diverse.



2019-1-BG01-KA203-062371

In the field of analytics, biosensors represent a significant step forward. In order to move analytics away from purely physics- or chemistry-based frameworks, the integration of biological components in sensory diagnostics has begun. This has allowed analytical functions that are not well adapted to conventional methods to conduct a vast diversity and specificity of biological components. The theoretical and demonstrated biosensor applications cover a significant range of human society and activity. Biosensor applications are grouped into three broad categories, depending on their measurement scale.

- ❖ Group Diagnostics: Environmental, Agricultural, and Industrial Applications
- ❖ Point-of-Use Diagnostics: Medical, and Security Applications
- ❖ Single-Cell Diagnostics: Metabolic Engineering, and Synthetic Biology Applications

BIOSENSORS & BIOCHIPS: ADVANCES IN MEDICAL DIAGNOSTICS

Biosensors consist of a biocatalyst that can recognize a biological element and a transducer that can turn the biocatalyst and the biological element combination occurrence into a measurable parameter.

The biocatalyst may be biomolecules such as enzymes, DNA, RNA, metabolites, cells, oligonucleotides, etc., and electrochemical, calorimetric, optical, acoustic, piezoelectric, etc. transducers. Biosensors using immobilized cells, enzymes and nucleic acids have come into the field in recent years in disease diagnostics. For engineering disease diagnostic biosensors, nanobiosensors utilizing the ultra-small size and unique properties have also been applied. The use of biosensors can quickly determine the health status, the onset and progression of the disease and, with the assistance of a multidisciplinary combination of chemistry, medical science and nanotechnology, can help to prepare treatment for many diseases. The devices are cost-effective, highly responsive, fast, user-friendly, and can be manufactured for human use in bulk. Numerous biosensors for the diagnosis of three major diseases, such as diabetes, cardiovascular disease and cancer, are the most developed ones.

Such biosensors, coined by Cammann, are analytical instruments that transform an electrical signal into a biological response. Biosensors can usually be highly precise and should be recyclable and irrespective of physical limitations such as pH, temperature. Practical approach to the design of a biosensor requires manufacturing, immobilization, transduction devices that offer multidisciplinary research engineering in both chemistry and biology.

Based on their working mechanism the diagnostic biosensors are divided into four major groups:

1. Enzyme-based biocatalytic biosensors.
2. Bioaffinity group, i.e. antibody, antigen and nucleic acid presence.
3. Microbes, i.e., microorganism-containing biosensors.
4. Nanosensors, i.e. active nanoparticle sensors that typically increase sensitivity and specificity for early disease detection.

2019-1-BG01-KA203-062371

These various types of biosensors help hormone levels, drugs, toxins, contaminants, heavy metals, pesticides, etc. to be identified with significant specificity.

Biosensors are tools that commonly estimate biological marker levels or any chemical reaction by creating signals that are primarily associated with an analyte's concentration in the chemical reaction. Typically, such biosensors help monitor diseases, drug discovery, pollutant detection, bacteria-causing disease detection, and markers that usually indicate diseased conditions, such as body fluids (saliva, blood, urine, sweat, etc.). A typical biosensor is shown in Figure 1.

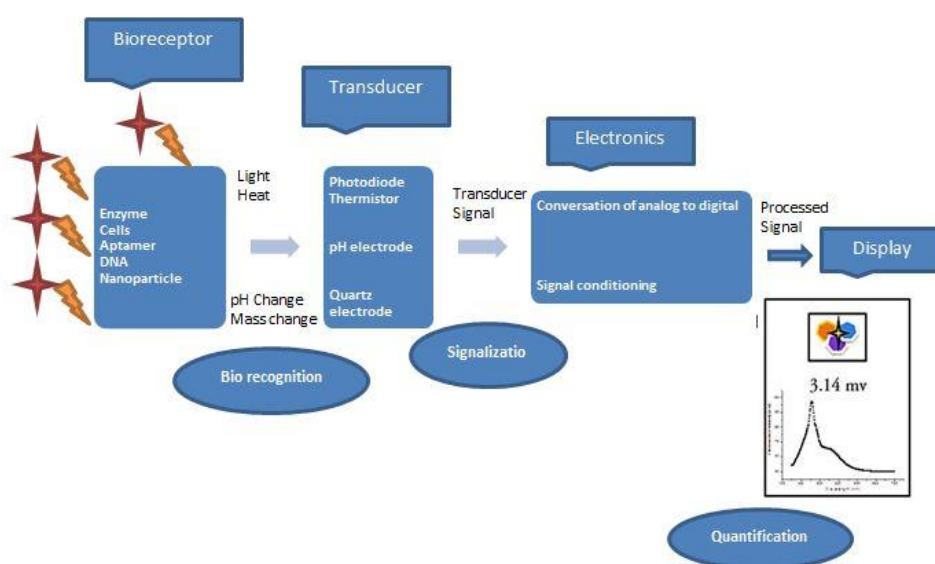


Figure 1. Schematic depiction of biosensor

A typical biosensor is composed of:

1. **Analyte:** A substance of interest, such as glucose for diabetes, that needs to be established.
2. **Bioreceptor:** A bioreceptor for enzymes may be a molecule which recognizes the analyte.
3. **Transducer:** Normally, a bio recognition event is converted into a detectable signal, known as signalization.
4. **Electronics:** In display form, it typically processes the transduced signal.
5. **Display:** Typically, the liquid crystal display results in a user-friendly manner in combination with hardware and software for biosensor generation.

There are several biosensor applications that have been introduced in different areas, such as medical science, the marine sector, the food industry, etc., and these biosensors are often programmed for improved sensitivity and linearity compared to conventional methods. However, the application of biosensors is growing increasingly in the field of medical science.

1. Glucose biosensors in diabetic management

Blood glucose monitoring has become a valuable tool in the management of diabetes and daily blood glucose levels are typically maintained by consulting clinicians who have developed a series of blood glucose sensors. *Diabetes mellitus* is the largest prevailing carbohydrate metabolism endocrine disorder with more morbidity and mortality in developing countries. Multiple tests are usual in diabetic patients for the investigation and monitoring of diabetic markers. The key diagnosis criteria for diabetes are the level of blood glucose, which includes diabetic patients' self-monitoring of glucose levels. Studies have shown that microvascular (nephropathy, neuropathy, and retinopathy) and macrovascular (coronary artery disease and stroke) complications can be improved by controlling the level of blood glucose in the normal range. Blood glucose is typically observed in healthy individuals in the range of 4.9-6.9 mM and can increase in diabetic patients up to 40 mM after glucose intake. Although different kinds of glucose sensors are commercially available, the third generation of glucose biosensors is shown in Figure 2 as an example.

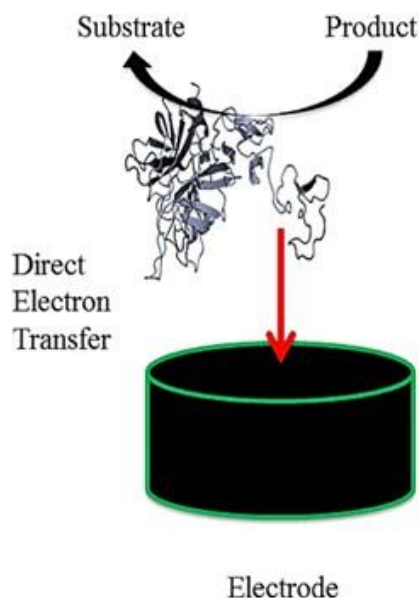


Figure 2. Third generation of glucose biosensor



2. Cardiovascular disease detection using biosensors

The number of deaths caused globally by cardiovascular disease (CVD) is significant and more people die of CVD than by any other disease. By 2015, about 17.7 million people had died from CVD, representing a total of 31% of all global deaths. 7.4 million of these were due to coronary heart disease and 6.7 million were due to stroke. By way of medication and therapy, a person with CVD needs earlier detection and management. The current CVD detection strategy relies on the traditional method, which is usually based on testing that can take many hours or even days. The WHO sets these diagnostic criteria, under which patients should follow at least one of the conditions, such as changes in the diagnostic electrocardiogram (ECG), elevation of biochemical markers in their blood samples, and characteristic chest pain. ECG is an important parameter for therapy management, but ECG is a poor diagnostic test in the case of CVD because half of CVD patients have a normal cardiogram, making it more difficult to diagnose this medical condition. Biosensor will aid in rapid diagnosis, providing excellent health care and reducing the delay time for the distribution of the results, which is immense stress for the patients.

3. Biosensor for detection of cancer

Cancer is one of the most lethal diseases, and several researchers have recently developed biosensors for early cancer detection. Most cancers are typically diagnosed by MRI, ultrasound or biopsy methods that rely on the physical properties and presence of the tumor and identify either advanced or invasive instruments. The variations in gene sequences, i.e. mutations, primarily cause cancer and thus require early diagnosis before the disease progresses. Early cancer detection makes treatment faster and more successful, opening up a biosensor platform for the detection of early cancer stages. Many experts assume that in the case of cancer, early detection could be possible because abnormalities in chemical and genetic composition may be identified long before the disease begins. Uncontrolled and irregular cell growth, commonly believed to be cancer, occurs due to the accumulation of unique genetic mutations and epigenetic defects. The tumor cells are shown to be resistant to apoptosis and the body's anti-growth defense mechanism. If it progresses and begins to expand to other body organs and systems, i.e. metastasize stage, the cancer becomes incurable. Oncogene stimulation and reducing the function of tumor suppressor genes (TSGs) are the two most important tumorigenesis mechanisms. Due to mutation or replication of normal gene (proto-oncogene), activation of oncogene takes place, which plays key roles including, control of cell growth, proliferation, and/or differentiation. Such genetic mutation guides the gene to produce an excess quantity of its gene product, resulting in dysregulation of cell division, cell growth and tumor establishment. Many oncogenes have been considered as promising cancer biomarkers for growth factor receptors. In ~ 33% of all breast cancers, the human epidermal growth factor receptor Her-2 is intensified, and cancers with strengthened Her-2 seem to develop and increase more rapidly. Awareness of Her-2 status is therefore essential in concluding the possible medication course. Trastuzumab is now a typical adjuvant therapy for patients with this type of amplified gene expression, a recombinant humanized monoclonal antibody targeted at Her-2 as a straight-forward treatment for breast cancer. TSGs are related to the control of insufficient cell growth and proliferation by minimizing



2019-1-BG01-KA203-062371

or preventing the division of cells. Retinoblastoma protein (Rb), BRCA1/2, and p53 are three of the well-studied TSGs in cancer. Rb is a master cell division regulator, and Rb mutation plays a significant role in various cancers. The most common causes of inactivation of the Rb1 gene are point mutations and deletions. BRCA1 is a DNA repair enzyme that is associated with newly replicated DNA 'proofreading' for fidelity and to search for any mutations. Until the cell divides, DNA repair enzymes normally work to excise replication errors. BRCA1 gene mutations are responsible for 50% of hereditary breast cancers and 80-90% of hereditary breast and ovarian cancers. Lastly, a main regulator of apoptosis or programmed cell death is the p53 protein. In the brain, breast, colon, lung, hepatocellular carcinomas, and leukaemia, p53 mutations are found. Another significant involvement with p53 loss is that it leads to the mechanism of resistance of chemotherapy drugs. The improvement of biosensors that can detect the existence of p53, Rb, and BRCA1 mutations is highly warranted and can enable us to evaluate the susceptibility of early cancer with detailed prognosis and treatment regimes.

4. Biochip in diagnostics

The DNA biochip opens up a new genetics-based field of diagnostics. The way the medical profession performs blood testing could be revolutionized by a newly developed DNA biochips. They are virtually immediate with the matchbox-sized biochip instead of a patient having to wait several days for results from a laboratory. And with no sacrifice of accuracy, it requires less blood. The DNA biochip reduces the need for radioactive labels used for detection, in addition to time savings. For technicians and laboratory workers handling samples and performing tests, this significantly decreases costs and future health effects. It also lowers disposal costs because, according to strict regulations, chemically labelled blood must be handled.

A biosensor must be highly sensitive and able to differentiate between, for example, bacteria, viruses or other chemical or biological species to be useful for detecting compounds in a real-life sample. According to Vo-Dinh, who clarified that the biochip mimics the sophisticated recognition capabilities of a living system, DNA biochips do that. The DNA biochip is a gene probe-based biosensor, as opposed to other biosensors based on enzyme and antibody probes. Gene probe-based biosensors provide exceptional selectivity and sensitivity, making them valuable tools for diagnosing genetic diseases and infectious species.

5. Biochip in Tuberculosis epidemic

The development of new biochip technologies by Russian and American scientists could bring some hope of halting the global resurgence of tuberculosis. Established by the U.S. Department of Energy's Argonne National Laboratory and the Russian Academy of Sciences' W. A. Engelhardt Institute



2019-1-BG01-KA203-062371

of Molecular Biology (Moscow), the technology is intended to help combat the current variety of drug-resistant strains of the disease.

The World Health Organization reports that tuberculosis kills more young people and adults, including AIDS and malaria combined, than any other infectious disease. The biggest challenge of the ongoing tuberculosis epidemic is that the disease can be caused by several different bacterial species, and each one is resistant to various drugs. The critical element in controlling the disease is to define the strain that affects a given patient and to determine the best antibiotic for combating that strain. To differentiate between numerous tuberculosis strains, Argonne intends to use biochip technology in research. Testing on segments of genetic material removed from tuberculosis bacteria would initially be carried out. Biochips are designed to simultaneously conduct a number of biochemical reactions and have been found to perform satisfactorily in laboratory testing. Since the detection of specific tuberculosis strains takes weeks or months, patients are frequently prescribed several antibiotics simultaneously.

6. Biochip in cancer

The biosensor chip technology also provides fast and simple access to crucial information about cancer-producing compound DNA damage, moving researchers a step closer in the fight against cancer. Unlike traditional methods of biosensing, a laser-based, high-resolution and low-temperature fluorescence method offers a precise fingerprint of the molecule. It is possible that its ease of use encourages the replacement of invasive endoscopic procedures and helps to detect colon cancer early on.

BIOSENSORS & BIOCHIPS APPLIED IN FOOD AND AGRICULTURE

The current food production faces immense challenges from the increasing human population, the maintenance of clean resources and food quality, and the protection of the environment and climate. Food sustainability is mainly a cooperative effort that results in the development of technology funded by both governments and companies. Several attempts have been supported to overcome challenges and improve the drivers in food production. Via their applications, biosensors and biosensing technologies are widely used to solve the major challenges of food production and its sustainability. As a result, there is a rising need for biosensing technology in the area of food sustainability. A technological system combining several technologies is defined by microfluidics. Nanomaterials, with its biosensing technology, is known to be the most innovative tool strongly associated with world populations in dealing with health, energy, and environmental issues. The need for point of care (POC) technology in this area focuses on analytical tools that are fast, simple, precise, compact, and low-cost.



2019-1-BG01-KA203-062371

For our existence and lives, food with its production industry is essential; and its sustainability is essential in continuous human growth on the planet. Current food production is facing immense difficulties from increasing human population, maintaining clean resources and food quality, and protecting environment and climate. Some of these issues stem from food production itself; others stem from other food production-related industries. Food recalls, for example, trigger major damage to food brands' credibility and prestige, with an estimate of \$15 million per incident over the last few years. 48 million sick cases are responsible for 3000 fatalities annually due to foodborne illnesses.

Food safety is largely a cooperative effort arising from both governments and companies in technology development. In order to pose new challenges in food safety issues, information technologies such as blockchain technology can accelerate communication between food quality, media and consumers. Five challenges can be summarized as the main challenges in the sustainability of food production: the production challenge of food safety and security; the quality challenge of food diversity and quality; the economic challenge in the leading food system, including its packaging and supply chain; the environmental challenge, including the processing of food waste; and the engineering challenge in the creation and generation of novel food.

Basically, a biosensor is an analytical instrument used to measure a sample's molecule of interest (target). In general, a bio-recognition factor (aptamer, antibody, enzyme, etc.) that is unique to the target is used. A physiochemical or biological signal is elicited by molecular recognition events between the recognition element and the target compound, which is transformed into a measurable quantity by the transducer. Signals are shown in either optical (colorimetric, fluorescence, chemiluminescence and plasmon surface resonance) or electrical (voltammetry, impedance and capacitance) or any other chosen format (Figure 3).

2019-1-BG01-KA203-062371

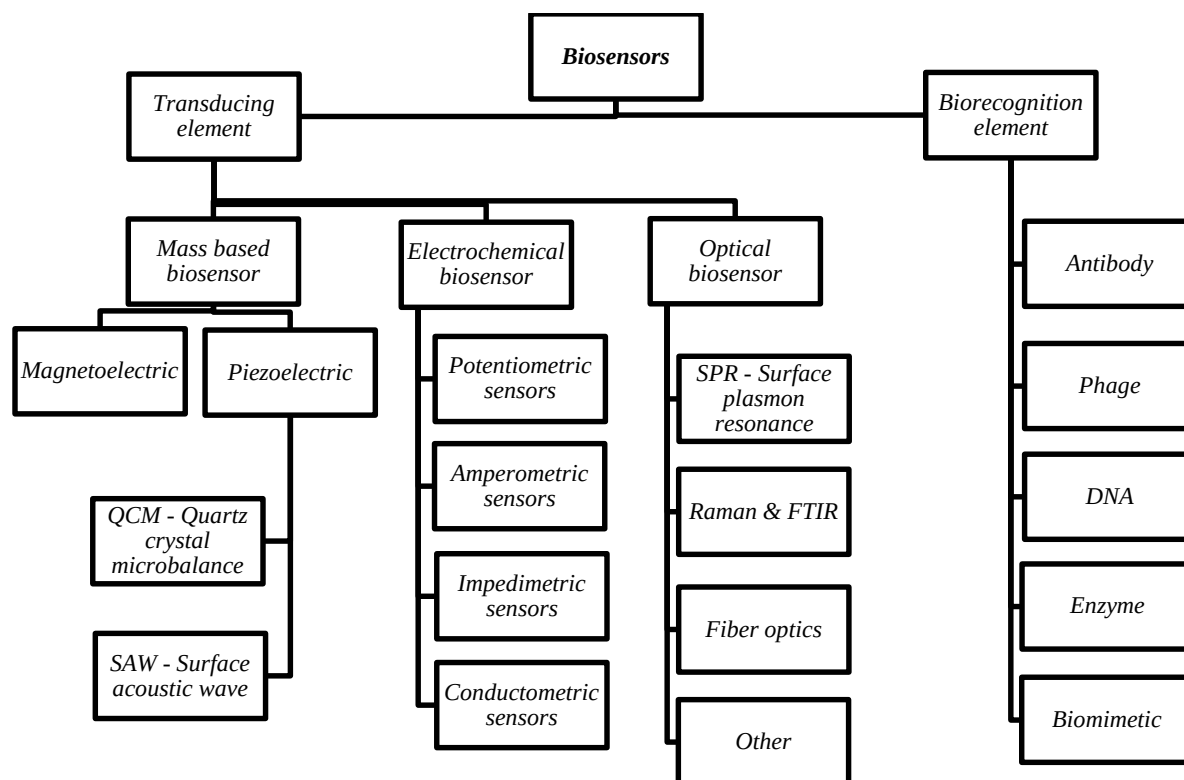


Figure 3. Classification of biosensors based on transducer and bio-recognition elements used in food analysis

As one of the primary objectives of food analysis, food safety is a major health issue in both animal and human lives. The advancement of food safety analytical technology means that it thrives in line with the rising interest in and emphasis on food supply safety issues. In food safety analysis, traditional approaches are labour-intensive, time-consuming, and need trained technicians. The application of microfluidics in food safety analysis provides fresh insight about how to detect foodborne toxins, allergens, pathogens, hazardous substances, heavy metals, and other contaminants effectively and rapidly. Microfluidics' features, such as its miniaturization capability, compact and reducible quantities of samples and reagents, make it a perfect technology for the development of food sustainability. Complex food matrix preparation and difficult manufacturing steps are the current challenges in the application of microfluidics to food sustainability. These challenges can be addressed by leveraging physical properties dependent on specific test targets, designing complex real food analysis microfluidic platforms, and incorporating into microfluidic systems biomolecules such as food proteins and DNA.



2019-1-BG01-KA203-062371

1. Nanomaterials in biosensing technology

With its biosensing technology, nanomaterials are the most promising tool in dealing with health, energy and environmental problems associated with population in the world. Particles smaller than 100 nm in at least one size dimension are known as nanomaterials. These nanomaterials are biocomposite polymers based on metal, metal oxide and carbon, and different types of nanoparticles have been established, such as magnetic iron, aluminum, gold, silver, copper, silica, zinc, zinc oxide, cerium oxide and titanium dioxide nanoparticles, and single/multiple walled carbon nanotubes (CNTs). Nanotechnology and its agricultural development have been greatly extended in different fields. These fields include food production, crop protection, detection of pathogens and toxins, purification of water, food packaging, disposal of wastewater, and environmental remediation. Improving the productivity and performance of applications is the priority of these agricultural fields.

In the field of food safety and protection, biosensing technologies have been developed for nutrient and quality detection, detection of pathogens and detection of toxins, as listed below.

2. Nutrient and quality detection

Food protection measures can be split down into two categories: post-harvest loss and food biosecurity. Food biosecurity means food contamination and degradation, which is addressed in the later sections, by environmental, political, unfair economic gain, warfare, or exacting revenge. Post-harvest loss, on the other hand, suggests the nutrients and edible conditions in food that need to be maintained between the harvest period and the moment of consumption by technologies. Since time differs from minutes to years, in maintaining and reducing losses, technologies focusing on reducing post-harvest losses are important.

To maintain food quality and to avoid post-harvest losses, new technology such as biosensing can be used. Biosensors have been developed, for example, to detect and analyse quantities of sweeteners in foods that can be used to detect both natural and artificial sweeteners. Sweeteners are widely used in food production and processing, but they have recently been identified in humans as causing health problems. A multi-channel biosensor has been developed to use electro-physiological sensing from taste epithelia to detect and analyse both natural and artificial sweeteners. To detect long-term signals from sucrose, glucose, cyclamate, and saccharin, respectively, the signals are studied through spatiotemporal techniques. The biosensor can distinguish between different concentrations with dose-dependent increased responses of the taste epithelium from different sweeteners. It can also distinguish between two natural sweeteners: sucrose and glucose, with two signal patterns. For glucose, the detection range is 50-150 mM, and for saccharin, 5-15 mM.

3. Detection of pathogens

Due to their reduced format, biosensors targeting pathogen detection such as bacteria (Table 1) and fungi (Table 2) started more than two decades ago; one device to address multiple problems, and a multi-panel signal detection. The ligand motif is a crucial element in the biosensor design for pathogen detection since it determines the sensitivity and efficiency of the device. The aim is to establish a fast, specific, and sensitive platform to detect in food samples the presence or absence of pathogens. It has been discovered that there is no ideal ligand, and various ligands have different advantages. The combination of bioreceptors to detect a large variety of microbes in different samples poses current challenges in pathogen biosensor detection; new synthetic ligand designs such as aptamers, small molecules, and peptides; and the incorporation of different ligands into a portable device to achieve rapid, effective, and low-cost detection.

Table 1. Conditions for numbers of bacteria grown in milk

<i>Temperature °C</i>	<i>24 h</i>	<i>48 h</i>	<i>96 h</i>	<i>168 h</i>
0	2400	2100	1850	1400
4	2500	3600	218,000	4,200,000
8	3100	12,000	1,480,000	
10	11,600	540,000		
15	180,000	28,000,000		
30	1,400,000,000			

Table 2. Temperature and water activity requirements for fungal growth

<i>Species</i>	<i>Minimum</i>	<i>Optimum</i>	<i>Maximum</i>	<i>Minimum</i>	<i>Optimum</i>
<i>Aspergillus ruber</i>	5	24	38	0.72	0.93
<i>A. amstelodami</i>	10	30	42	0.70	0.94
<i>A. flavus</i>	12	35	45	0.80	0.99
<i>A. fumigatus</i>	12	40	52	0.83	0.99
<i>A. niger</i>	10	35	45	0.77	0.99
<i>Penicillium martensii</i>	5	24	32	0.90	0.99

4. Detection of toxins

The mainstream of development in food safety is electrochemical biosensors for rapid detection and assessment of food toxins. Numerous platforms have been developed to allow customized and individualized devices to meet particular environmental and organizational requirements and to reach the nM to fM detection limit levels. For example, to encourage unique binding profiles, bioreceptor arrays address individual electrodes functionalized with different bioreceptors with binding targets. In addition to electrochemical biosensing, toxin and chemical detection in food production have been applied to other biosensors such as optic and piezoelectric sensing (Figure 4). In order to sense toxins, fluorescent nanoparticles have been produced in foods and bodies, including on-surface, inter- and intra-cellular foods.

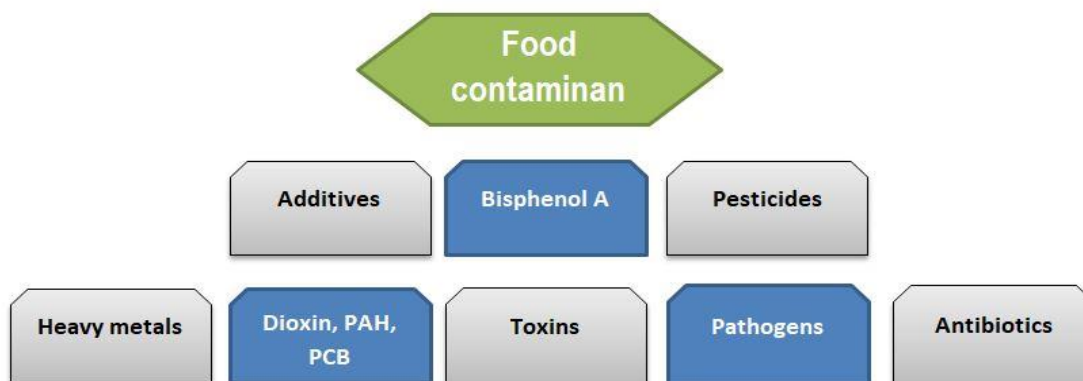


Figure 4. Predominant food contaminants and the target analytes in the food manufacturing industries

Toxin extraction from complicated food samples is one of the main obstacles in creating a fully automated toxin detector. To automatically assess their harmful levels from food and water samples, potential systems are expected to extract, process, and measure toxins. In identifying, discriminating, and quantifying chemical toxins in food matrices, sophisticated separation strategies have been coupled with SERS. In addition, even though they are typically in lower amounts, chemical contaminants from food processing can be a challenge. Lower stability, selectivity and sensitivity are another challenge in food toxin detection, where MIPs can be a solution to provide stable and low-cost alternatives.

Heavy metals like Ag^+ , As^{3+} , Cd^{2+} , Hg^{2+} , Pb^{2+} , and Zn^{2+} are known as chemical pollutants that form stable states of oxidation and interfere with metabolic pathways, resulting in health problems. Aptamer and DNA-based biosensors can detect heavy metals at both nanoscale and very large-scale levels, which are appropriate for food safety screening and monitoring. In order to detect arsenate in food, a heavy metal detecting biosensor is based on genetically modified bacterial cells and a green, fluorescent signal amplifier. With a detection range of 5-140 $\mu\text{g/L}$ of arsenic, its arsenic detection lasts just one hour and can be integrated with optical power output for its future biosensing optical fibre. Other biosensing technologies like aptamers, nanoparticles and graphene electrodes have been successfully applied to the identification and evaluation of arsenic, with the potential to be produced as fast, simple, easy-to-use, and low-cost devices.

Nanotechnology has been adapted to two separate fields of agri-food pesticides: as a pesticide delivery vector for pesticide management and as a trace-amount detector for pesticides. In the first field, nanoparticles are able to slowly modify pesticides to target insect pests, which helps prevent groundwater



2019-1-BG01-KA203-062371

and topsoil pollution, reduce pesticide levels and improve efficiency. In the second field, bio- or biomimetic-based nanotechnology, like antibodies, enzymes, aptamers, and MIP-like macromolecules, improves stability, selectivity, sensitivity, and speed of detection. Furthermore, bacterial, fungal, algal, and mammalian cells are all cell-based biosensors used in pesticide and herbicide detection, helping to establish fast, reliable, real-time, and cost-effective tools for decontamination procedures and preventive casualty damage.

Carcinogens, odorants, and marine contaminants are other toxins that are significant in food production. Carcinogens are a complex group of trace amount of toxins, like pesticides, heavy metals, mycotoxins, and acrylamide, in which the difficulty of identifying trace-amounts is a challenge; and imprinted aptamers, nanotechnology, and biosensing are optimistic for promising future use. Sensitive and soluble molecules effective in odour detection for olfactory animal systems are odorant binding proteins. A nanosensor combining localized SPR and small odorant binding proteins from honeybees has been established in which the detection range is 10 nM - 1 mM using a quantitative array of nanocups. To monitor and preserve a stable environment for marine food systems, marine contaminant detection is used. Finally, through their sensitive detection capabilities, miniaturized devices, wireless communication, and small-scale networks, biosensors can be applied to marine food safety to be established as advanced analytical and monitoring tools.

Another development kit for food safety biosensors focuses on the detection of genetically modified organisms (GMOs) in food products. Since the 1990s, GMOs in all fields of agricultural products have been considered a biotechnology revolution. To present, more than 45 percent of the world's soybeans, 40 percent of corn, and 50 percent of cotton are GM products; and GM is also used in livestock. Recent research, however, indicates that GMO products can affect human and animal bodies through gastrointestinal problems, antibiotic resistance, allergenicity, diversity of farm products degradation, and undesired gene flow to other species. Biosensors are designed to measure GMOs in foods and feeds using isothermal DNA amplification and fast detection signal detection to identify GM genes. Detecting unidentified DNA genes that can be resolved by high-throughput technology like the combination of biosensing and arrays, and the development of databases of GMO genes are the key challenges in GMO detection.

BIOSENSORS & BIOCHIPS FOR ENVIRONMENTAL MONITORING

Due to the strong connection between environmental pollution and human health/ socioeconomic progress, environmental monitoring has become one of the priorities on a European and global scale. Biosensors have been commonly used as cost-effective, rapid, *in situ*, and real-time analytical techniques in this field. The recent development of biosensors with new transduction materials obtained from nanotechnology and for multiplexed pollutant detection, involving multidisciplinary experts, explains the need for compact, fast, and smart biosensing devices. Several recent developments exist in the monitoring of air, water, and soil contaminants by biosensors under real conditions, like pesticides,



highly toxic components and small organic molecules, including toxins and endocrine disrupting chemicals.

Biosensors used in environmental monitoring can be categorized as optical (including optical fibre and surface plasmon resonance biosensors), electrochemical (including amperometric and impedance biosensors) and piezoelectric (including quartz crystal microbalance biosensors) based on their transduction or as immunosensors, aptasensors, genosensors and enzymatic biosensors based on their recognition elements, respectively when are used antibodies, aptamers, nucleic acids, and enzymes. The majority of biosensors in environmental monitoring are recognized as immunosensors and enzymatic biosensors, but the development of aptasensors has recently increased due to the beneficial characteristics of aptamers, like ease of modification, thermal stability, *in vitro* synthesis and the ability to design their structure, to differentiate targets with different functional groups and to rehybridize.

Study on the design of biosensors for the monitoring of organic pollutants, potentially toxic elements and pathogens in the environment has led to the sustainable development of civilization due to the environmental pollution issues confronting human health. Various chromatographic techniques (such as gas chromatography and high-performance liquid chromatography combined with capillary electrophoresis or mass spectrometry) are conventional analytical methods used for environmental monitoring of pollutants, but they require costly reagents, time-consuming sample pre-treatment and costly equipment. Therefore, for monitoring pollutants responsible for adverse effects on habitats and human health, more sensitive, cost-effective, fast, easy to function, and compact biosensing devices are desperately needed to overcome the magnification of environmental issues. In the case of accidental release of pesticides or acute poisoning, for example, common methods are not appropriate for *in situ* measurements where fast, miniaturized, and portable equipment like environmental monitoring biosensors is required. In this regard, the role of nanotechnology in the creation of rapid and intelligent biosensing devices is crucial for the success of environmental pollutant detection; most recent biosensors include nanomaterials and novel nanocomposites in their systems, which are beneficial for improving analytical performance, such as sensitivity and detection limits.

For the detection and monitoring of different environmental pollutants, biosensors, including immunosensors, aptasensors, genosensors and enzymatic biosensors have been documented using antibodies, aptamers, nucleic acids, and enzymes as recognition elements.

1. Pesticides

Pesticides are among the most significant environmental pollutants because of their large presence in the environment. Organophosphorus insecticides, for example, are commonly used in agriculture and represent a group of pesticides which, due to their high toxicity, are of immense environmental concern. Easy, responsive, and miniaturized *in situ* methodologies like biosensors have



2019-1-BG01-KA203-062371

therefore been established as analytical strategies for their detection and monitoring, without the need for comprehensive sample pre-treatment.

Disposable amperometric enzymatic (acetylcholinesterase) biosensors were proposed for the detection of organophosphorus insecticides using paraoxon as a model analyte applying a cysteamine self-assembled monolayer on gold screen-printed electrodes. The disposable biosensors showed a linear spectrum of up to 40 ppb with a 2 ppb detection limit and a $113 \mu\text{A mM cm}^{-2}$ sensitivity. Using the self-assembled monolayer, good analytical output could be due to the highly oriented enzyme immobilization. Recoveries of 97 ± 5 percent ($n = 3$) were reported after being tested in river water samples spiked with 10 ppb of paraoxon, indicating the effectiveness of such enzymatic biosensors. Furthermore, the use of disposable screen-printed electrodes dispenses with time-consuming methods like the reactivation of immobilized enzymes utilizing, for example, obidoxime solution and pralidoxime iodide (PAM) or the use of the renewable enzyme membrane needed for the second application of biosensors.

Nanoparticles based on iridium oxide have been used in the disposable enzymatic biosensor with tyrosinase based on low-cost screen-printed carbon electrodes for the detection of chlorpyrifos in river water samples. Linear biosensor response ($0.01\text{--}0.1 \mu\text{M}$) and low detection limit (3 nM) were reported, which could be due to the high conductivity of nanoparticles of iridium oxide and tyrosinase efficiency. Recovery tests were carried out in river water samples with the addition of $0.1 \mu\text{M}$ of chlorpyrifos and recoveries of 90 ± 9.6 percent were obtained with a residual standard deviation (RSD) smaller than 10 percent ($n = 3$) to demonstrate the applicability of the biosensor.

Acetamiprid was detected by colorimetric aptasensors and water samples by impedimetric aptasensors in real environmental samples, like fresh surface soil samples. A linear range of 75 nM to $7.5 \mu\text{M}$ and a detection limit of 5 nM were observed with the colorimetric aptasensor, while a wider linear range (50 fM to $10 \mu\text{M}$) and a lower detection limit (17 fM) were observed with the impedimetric aptasensor. Gold nanoparticles, multi-walled carbon nanotubes (MWCNT) and reduced graphene oxide nanoribbons were used in that biosensor as a composite to sustain the electrode surface acetamiprid aptamer, which could be responsible for higher electron transfer and improved analytical performance of the biosensor. A related detection limit (33 fM) was observed by an aptasensor based on silver nanoparticles anchored on nitrogen-doped nanocomposite graphene oxide constructed for acetamiprid detection in wastewater samples.

2. Pathogens

The existence of pathogens in environmental matrices, and especially in water compartments, could pose a serious risk to human health, and some biosensors have recently been suggested for monitoring the environment. For example, for the detection of metabolically active *Legionella pneumophila* in complex environmental water samples, rapid and precise optical biosensors based on surface plasmon resonance have been proposed. In one study, the detection principle was based on the

identification of bacterial RNA by the immobilized RNA detector probe on the gold surface of the biochip. For signal amplification, streptavidin-conjugated quantum dots were used, and the detection period was approximately three hours, indicating the viability of the biosensing device for successful bacteria detection in the range of 10^4 - 10^8 CFU mL⁻¹.

3. Potentially toxic elements

The contamination by heavy metals and corresponding ions of the natural waters can pose significant risks to human health, and compact, low-cost, and rapid heavy metal analyses are a global priority concern. As a model target for testing an optical DNA biosensor for the detection of heavy metal ions that are extremely toxic and common pollutants in the environment, mercury ions (Hg^{2+}) were used. The biosensor was compact, low-cost, and rapid with *in situ* screening of Hg^{2+} in natural waters in less than 10 min. The detection principle is focused on the capacity of certain metal ions to bind selectively to certain bases to form stable metal-mediated DNA duplexes; in the case of Hg^{2+} , thymine bases can be selectively coordinated to form stable thymine- Hg^{2+} -thymine complexes. In the detection range between 0 and 1000 nM, a detection limit of 1.2 nM was achieved, which is lower than the maximum value requested by the United States Environmental Protection Agency (10 nM)

For the detection of Pb^{2+} in water samples (pond and lake water samples) using DNAzymes/carboxylated magnetic beads and DNA aptamers, two fluorescence based optical biosensors have recently been suggested. The detection limits of 5 nM and 61 nM, respectively, were observed by biosensors based on DNAzymes and DNA aptamers, with a respective linear detection range of 0 to 50 nM and 100 to 1000 nM. The use of label-free unique dye (SYBER Green I), which was intercalated with double stranded DNA, showing strong fluorescence intensities, as seen in Figure 5. Moreover, the absence of biosensor fluorescence intensity is observed only with the dye (curve a). With the DNAzyme + Pb^{2+} (curve b) the fluorescence intensity increases with the addition of the dye + DNAzyme + Pb^{2+} , illustrating the sensitivity of the biosensor towards Pb^{2+} .

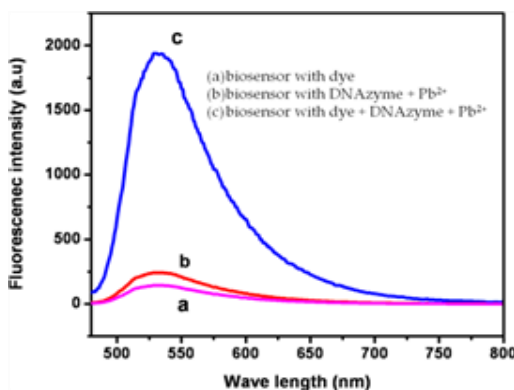


Figure 5. Fluorescence emission spectra for detection of Pb^{2+}



2019-1-BG01-KA203-062371

4. Toxins

Harmful toxins like brevetoxins and microcystins are created by the eutrophication of aquatic systems by the algal blooms of cyanobacteria, and thus accurate and cost-effective systems are needed for the early detection of such toxins. For the sensitive detection of brevetoxin-2, a marine neurotoxin, an electrochemical aptasensor has been used composed by gold electrodes functionalized with cysteamine self-assembled monolayers. A detection limit of 106 pg mL^{-1} was achieved and strong selectivity was observed for brevetoxin-2 against other toxins of various groups, like okadaic acid and microcystin. The feasibility of the aptasensor for detecting brevetoxin-2 in real samples was achieved by analysing shellfish and strong recoveries (102-110 percent) were reported, indicating no interaction with the aptasensor response from the shellfish matrix.

5. Endocrine disrupting chemicals

In water samples, bisphenol A was detected as an endocrine disrupting chemical by aptasensors based on the fluorescence principle with functionalized aptamers (fluorescein amidite) and gold nanoparticles and based on evanescent-wave optical fibre. The evanescent-wave optical fibre aptasensor was compact and found to be rapid, cost-effective, sensitive and selective for the detection of bisphenol A in water samples, with the benefit of no requirement of any pre-concentration or treatment steps. Furthermore, the aptasensor can be reused for 90 s by regeneration with a 0.5% sodium dodecyl sulphate (SDS) solution and further washing with a phosphate buffered saline (PBS) solution (pH 7.2) for over a hundred assay cycles without any noticeable loss of efficiency. Similar detection limits (0.1 and 0.45 ng mL^{-1}) were observed in both optical biosensors where the DNA molecule probe, which is the complementary sequence of a small fraction of the bisphenol A aptamer, was adsorbed by electrostatic interaction in the surface of gold nanoparticles and covalently immobilized on the surface of the fibre. Lately, for the detection of bisphenol A in river water samples using molybdenum carbide nanotubes, another fluorescence-based aptasensor was proposed. With such a label-free, inexpensive, and easy to use aptasensor, a low detection limit of 0.23 ng mL^{-1} has been obtained. The specificity of the aptasensor was evaluated by analysing other molecules with structures similar to that of bisphenol A (e.g., 4,4J-biphenol, bisphenol AF, and 4,4J-sulfonyldiphenol) and only background signals showing high specificity for bisphenol A were identified for these molecules.

A disposable and label-free electrochemical immunosensor based on a field effect transistor with SWCNT has recently been employed in seawater samples for assessing another endocrine disrupting chemical - 4-nonylphenol. The immunosensor has a high reproducibility ($0.56 \pm 0.08\%$), an average recovery of 97.8% to 104.6% and a low detection limit ($5 \mu\text{g L}^{-1}$), which is lower than the recommended maximum concentration of $7 \mu\text{g L}^{-1}$ specified by the corresponding regulations. In seawater samples such as 4-nonylphenol, the biosensor could be used to detect hazardous priority substances, even at low concentrations and with an easy and low-cost methodology.



6. Other environmental compounds

New, fast, and accurate analytical methodologies have been needed for the early detection and monitoring of various other hazardous compounds liberated during algal blooms. Due to the excellent sensitivity and specificity of nucleic acid probes to their complementary binding partners, biosensors have been developed to detect algal RNA. For the enhanced selective and sensitive detection of RNA from 13 harmful algal organisms, an electrochemical genosensor based on screen-printed gold electrode was recently reported; the genosensor could distinguish RNA targets from environmental samples (spiked seawater samples) containing 10^5 cells, considered to be the limit of detection.

References

- Adami A; Mortari A; Morganti E; Lorenzelli L. 2018. Microfluidic Sample Preparation Methods for the Analysis of Milk Contaminants. Available online: <https://www.hindawi.com/journals/js/2016/2385267/>
- Al-Mawali A. 2015. Non-communicable diseases: shining a light on cardiovascular disease, Oman's biggest killer. *Oman Med. J.*, 30 (4): 227.
- Arduini F, Guidone S, Amine A, Palleschi G, Moscone D. 2013. Acetylcholinesterase biosensor based on self-assembled monolayer-modified gold-screen printed electrodes for organophosphorus insecticide detection. *Sens. Actuators B Chem.*, 179: 201–208.
- Arduini F, Cinti S, Scognamiglio V, Moscone D. 2016. Nanomaterials in electrochemical biosensors for pesticide detection: advances and challenges in food analysis. *Microchim. Acta*, 183: 2063–2083.
- Arugula MA, Simonian AL. 2016. Biosensors for Detection of Genetically Modified Organisms in Food and Feed. In *Genetically Modified Organisms in Food*, Elsevier: Amsterdam, The Netherlands, pp. 97–110, ISBN 978-0-12-802259-7.



2019-1-BG01-KA203-062371

- Ayari-Jeridi H. et al. 2015. Mutation spectrum of RB1 gene in unilateral retinoblastoma cases from Tunisia and correlations with clinical features. *PLoS One*, 10 (1): e0116615.
- Bahadır EB, Sezgentürk MK. 2017. Biosensor technologies for analyses of food contaminants. In *Nanobiosensors*, Elsevier: Amsterdam, The Netherlands, pp. 289–337, ISBN 978-0-12-804301-1.
- Baldwin CJ. 2015. Introduction to the Principles. In *The 10 Principles of Food Industry Sustainability*, John Wiley & Sons, Ltd.: Hoboken, NJ, USA, pp. 1–14, ISBN 978-1-118-44769-7.
- Beck MB, Walker VR. 2013. On water security, sustainability, and the water-food-energy-climate nexus. *Front. Environ. Sci. Eng.*, 7: 626–639.
- Belkhamssa N, da Costa JP, Justino CIL, Santos PSM, Cardoso S, Duarte AC, Rocha-Santos T, Ksibi M. 2016. Development of an electrochemical biosensor for alkylphenol detection. *Talanta*, 158: 30–34.
- Bhalla N. et al. 2016. Introduction to biosensors. *Essays Biochem.*, 60 (1): 1–8.
- Bohunicky B, Mousa SA. 2011. Biosensors: the new wave in cancer diagnosis. *Nanotechnol. Sci. Appl.*, 4: 1.
- Bourne MC. 2014. Food Security: Postharvest Losses. In *Encyclopedia of Agriculture and Food Systems*, Elsevier: Amsterdam, The Netherlands, pp. 338–351, ISBN 978-0-08-093139-5.
- Bruen D et al. 2017. Glucose sensing for diabetes monitoring: recent developments. *Sensors*, 1866.
- Burris KP, Stewart CN. 2012. Fluorescent nanoparticles: Sensing pathogens and toxins in foods and crops. *Trends Food Sci. Technol.*, 28: 143–152.
- Byrne B et al. 2009. Antibody-based sensors: principles, problems and potential for detection of pathogens and associated toxins. *Sensors*, 9 (6): 4407–4445.
- Cash KJ, Clark HA. 2010. Nanosensors and nanomaterials for monitoring glucose in diabetes. *Trends Mol. Med.*, 16 (12): 584–593.
- Chao R, Mishra S, Si T, Zhao H. 2017. Engineering biological systems using automated biofoundries. *Metab. Eng.*, 42: 98–108.
- Chen Y, Li H, Gao T, Zhang T, Xu L, Wang B, Wang J, Pei R. 2018. Selection of DNA aptamers for the development of light-up biosensor to detect Pb(II). *Sens. Actuators B Chem.*, 254: 214–221.
- Eissa S, Siaj M, Zourob M. 2015. Aptamer-based competitive electrochemical biosensor for brevetoxin-2. *Biosens. Bioelectron.*, 69: 148–154.
- Elmore S. 2007. Apoptosis: a review of programmed cell death. *Toxicol. Pathol.*, 35 (4): 495–516.



2019-1-BG01-KA203-062371

Enrico DL, Manera MG, Montagna G, Cimaglia F, Chesa M, Poltronieri P, Santino A, Rella R. 2013. PR based immunosensor for detection of *Legionella pneumophila* in water samples. *Opt. Commun.*, 294: 420–426.

EPA. National Recommended Water Quality Criteria-Aquatic Life Criteria Table. Available online: <http://www.epa.gov/wqc/national-recommended-water-quality-criteria-aquatic-life-criteria-table>

Fei A, Liu Q, Huan J, Qian J, Dong X, Qiu B, Mao H, Wang K. 2015. Label-free impedimetric aptasensor for detection of femtomole level acetamiprid using gold nanoparticles decorated multiwalled carbon nanotube-reduced graphene oxide nanoribbon composites. *Biosens. Bioelectron.*, 70: 122–129.

Foudeh AM, Trigui H, Mendis N, Faucher SP, Veres T, Tabrizian M. 2015. Rapid and specific SPRi detection of *L. pneumophila* in complex environmental water samples. *Anal. Bioanal. Chem.*, 407: 5541–5545.

Garnet TF. 2013. Food sustainability: Problems, perspectives and solutions. *Proc. Nutr. Soc.*, 72: 29–39.

Gheorghe I, Czobor I, Lazar V, Chifiriuc MC 2017. Present and perspectives in pesticides biosensors development and contribution of nanotechnology. In *New Pesticides and Soil Sensors*, Elsevier: Amsterdam, The Netherlands, pp. 337–372, ISBN 978-0-12-804299-1.

Ghorashi M. 2018. Technology's Role in Eradicating Foodborne Illness. Available online: <https://www.foodsafetymagazine.com/signature-series/technologye28099s-role-in-eradicating-foodborne-illness/>

Giacinti C, Giordano A. 2006. RB and cell cycle progression. *Oncogene*, 25 (38): 5220-5227.

Guo L, Li Z, Chen H, et al. 2017. Colorimetric biosensor for the assay of paraoxon in environmental water samples based on the iodine-starch color reaction. *Anal. Chim. Acta*, 967: 59–63.

Hameed I, et al. 2015. Type 2 diabetes mellitus: from a metabolic disorder to an inflammatory condition. *World J. Diabetes*, 6 (4): 598.

Hassani S, Momtaz S, Vakhshiteh F, et al. 2017. Biosensors and their applications in detection of organophosphorus pesticides in the environment. *Arch. Toxicol.*, 91: 109–130.

He MQ, Wang K, Wang J, Yu YL, He RH. 2017. A sensitive aptasensor based on molybdenum carbide nanotubes and label-free aptamer for detection of bisphenol A. *Anal. Bioanal. Chem.*, 409: 1797–1803.

Holford TR et al. 2012. Recent trends in antibody based sensors. *Biosens. Bioelectron.*, 34 (1): 12–24.



2019-1-BG01-KA203-062371

- Hughes RA, Ellington AD. 2017. Synthetic DNA synthesis and assembly: Putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.*, 9.
- Husu I, Rodio G, Touloupakis E, et al. 2013. Insights into photo-electrochemical sensing of herbicides driven by *Chlamydomonas reinhardtii* cells. *Sens. Actuators B Chem.*, 185: 321–330.
- Jain KK. 2004. “Applications of biochips: from diagnostics to personalized medicine.” *Curr Opin Drug Discov Devel*, 7(3): 285-289.
- Jiang D, Du X, Liu Q, Zhou L, Dai L, Qian J, Wang K. 2015. Silver nanoparticles anchored on nitrogen-doped graphene as a novel electrochemical biosensing platform with enhanced sensitivity for aptamer-based pesticide assay. *Analyst*, 140: 6404–6411.
- Justino CIL, Freitas AC, Duarte AC, Santos TAPR. 2015. Sensors and biosensors for monitoring marine contaminants. *Trends Environ. Anal. Chem.*, 6–7: 21–30.
- Justino CIL, Freitas AC, Pereira R, Duarte AC, Rocha-Santos TAP. 2015. Recent developments in recognition elements for chemical sensors and biosensors. *Trends Anal. Chem.*, 68: 2–17.
- Kazemi-Darsanaki R et al. 2012. Biosensors: functions and applications. *J. Biol. Today's World*, 2 (1): 20–23.
- Khot LR, Sankaran S, Maja JM, Ehsani R, Schuster EW. 2012. Applications of nanomaterials in agricultural production and crop protection: A review. *Crop Prot.*, 35: 64–70.
- Kost GJ, Tran NK., 2005. Point-of-care testing and cardiac biomarkers: the standard of care and vision for chest pain centers. *Cardiol. Clin.*, 23 (4): 467–490.
- Lang Q, Han L, Hou C, Wang F, Liu A. 2016. A sensitive acetylcholinesterase biosensor based on gold nanorods modified electrode for detection of organophosphate pesticide. *Talanta*, 156: 34–41.
- Lee EY, Muller WJ. 2010. Oncogenes and tumor suppressor genes. *Cold Spring Harb. Perspect. Biol.*, 2 (10): a003236.
- Li Z, Yu Y, Li Z, Wu T. 2015. A review of biosensing techniques for detection of trace carcinogen contamination in food products. *Anal. Bioanal. Chem.*, 407: 2711–2726.
- Liao W, Lu X. 2016. Determination of chemical hazards in foods using surface-enhanced Raman spectroscopy coupled with advanced separation techniques. *Trends Food Sci. Technol.*, 54: 103–113.
- Long F, Zhu A, Shi H, Wang H, Liu J. 2013. Rapid on-site/in-situ detection of heavy metal ions in environmental water using a structure-switching DNA optical biosensor. *Sci. Rep.*, 3: 2308.
- Loo C, et al. 2005. Immunotargeted nanoshells for integrated cancer imaging and therapy. *Nano Lett.*, 5 (4): 709–711.



2019-1-BG01-KA203-062371

Maduraiveeran G, Jin W. 2017. Nanomaterials based electrochemical sensor and biosensor platforms for environmental applications. *Trends Environ. Anal. Chem.*, 13: 10–23.

Marcellin E, Nielsen LK 2018. Advances in analytical tools for high throughput strain engineering. *Curr. Opin. Biotechnol.*, 54: 33–40.

Martín-Timón I et al. 2014. Type 2 diabetes and cardiovascular disease: have all risk factors the same strength? *World J. Diabetes*, 5 (4): 444.

Mayorga-Martinez C, Pino F, Kurbanoglu S, et al. 2014. Iridium oxide nanoparticles induced dual catalytic/inhibition based detection of phenol and pesticide compounds. *J. Mater. Chem. B*, 2: 2233–2239.

McPartlin DA, Loftus JH, Crawley AS, et al. 2017. Biosensors for the monitoring of harmful algal blooms. *Curr. Opin. Biotechnol.*, 45: 164–169.

Mehrotra P. 2016. Biosensors and their applications - a review. *J. Oral. Biol. Craniofac Res.*, 6 (2): 153–159.

Meriç S, Çakır Ö, Turgut-Kara N, Arı S. 2014. Detection of genetically modified maize and soybean in feed samples. *Genet. Mol. Res.*, 13: 1160–1168.

Moran KLM, Fitzgerald J, McPartlin DA, Loftus JH, O’Kennedy R. 2016. Biosensor-Based Technologies for the Detection of Pathogens and Toxins. In *Comprehensive Analytical Chemistry*, Elsevier: Amsterdam, The Netherlands, Volume 74, pp. 93–120, ISBN 978-0-444-63579-2.

Mungroo NA, Neethirajan S. 2014. Biosensors for the Detection of Antibiotics in Poultry Industry—A Review. *Biosensors*, 4: 472–493.

Ngoepe M et al. 2013. Integration of biosensors and drug delivery technologies for early detection and chronic management of illness. *Sensors*, 13 (6): 7680–7713.

Omidfar K. et al. 2013. New analytical applications of gold nanoparticles as label in antibody based sensors. *Biosens. Bioelectron.*, 43: 336–347.

Orozco J, Villa E, Manes C, Medlin LK, Guillebault D. 2016. Electrochemical RNA genosensors for toxic algal species: Enhancing selectivity and sensitivity. *Talanta*, 161: 560–566.

Patra S, Roy E, Madhuri R, Sharma PK. 2017. A technique comes to life for security of life: The food contaminant sensors. In *Nanobiosensors*, Elsevier: Amsterdam, The Netherlands, pp. 713–772, ISBN 978-0-12-804301-1.



2019-1-BG01-KA203-062371

Pola-López LA, Camas-Anzueto JL, Martínez-Antonio A, et al. 2018. Novel arsenic biosensor “POLA” obtained by a genetically modified *E. coli* bioreporter cell. *Sens. Actuators B Chem.*, 254: 1061–1068.

Ragavan KV, Selvakumar LS, Thakur MS. 2013. Functionalized aptamers as nano-bioprobes for ultrasensitive detection of bisphenol-A. *Chem. Commun.*, 49: 5960–5962.

Rapini R, Marrazza G. 2016. Biosensor Potential in Pesticide Monitoring. In *Comprehensive Analytical Chemistry*, Elsevier: Amsterdam, The Netherlands, Volume 74, pp. 3–31, ISBN 978-0-444-63579-2.

Ravikumar A, Panneerselvam P, Radhakrishnan K, et al. 2017. DNzyme based amplified biosensor on ultrasensitive fluorescence detection of Pb(II) ions from aqueous system. *J. Fluoresc.*, 27: 2101–2109.

Rocchitta G, et al. 2016. Enzyme biosensors for biomedical applications: strategies for safeguarding analytical performances in biological fluids. *Sensors*, 16 (6).

Rotariu L, Lagarde F, Jaffrezic-Renault N, Bala C. 2016. Electrochemical biosensors for fast detection of food contaminants—Trends and perspective. *TrAC Trends Anal. Chem.*, 79: 80–87.

Saucedo NM, Mulchandani A. 2016. Sensing of Biological Contaminants. In *Comprehensive Analytical Chemistry*, Elsevier: Amsterdam, The Netherlands, Volume 74, pp. 73–91, ISBN 978-0-444-63579-2.

Shi H, Zhao G, Liu M, Fan L, Cao T. 2013. Aptamer-based colorimetric sensing of acetamiprid in soil samples: Sensitivity, selectivity and mechanism. *J. Hazard. Mater.*, 260: 754–761.

Singh M, del Valle M. 2015. Arsenic Biosensors. In *Handbook of Arsenic Toxicology*, Elsevier: Amsterdam, The Netherlands, pp. 575–588, ISBN 978-0-12-418688-0.

Sinha K, Ghosh J, Sil PC. 2017. New pesticides: A cutting-edge view of contributions from nanotechnology for the development of sustainable agricultural pest control. In *New Pesticides and Soil Sensors*, Elsevier: Amsterdam, The Netherlands, pp. 47–79, ISBN 978-0-12-804299-1.

Sodano V, Gorgitano MT, Quaglietta M, Verneau F. 2016. Regulating food nanotechnologies in the European Union: Open issues and political challenges. *Trends Food Sci. Technol.*, 54: 216–226.

Specht K, Siebert R, Hartmann I, et al. 2014. Urban agriculture of the future: An overview of sustainability aspects of food production in and on buildings. *Agric. Hum. Values*, 31: 33–51.

Stadler RH. 2016. Foreword for Food Processing—Derived Contaminants in Food Analysis. In *Reference Module in Food Science*, Elsevier: Amsterdam, The Netherlands, ISBN 978-0-08-100596-5.



2019-1-BG01-KA203-062371

- Tabish SA. 2007. Is diabetes becoming the biggest epidemic of the twenty-first century? *Int J. Health Sci.* 1 (2), V–VIII.
- Templier V, Roux A, Roupioz Y, Livache T. 2016. Ligands for label-free detection of whole bacteria on biosensors: A review. *TrAC Trends Anal. Chem.*, 79: 71–79.
- Thomas S, et al. 2015. The expression of retinoblastoma tumor suppressor protein in oral cancers and precancers: a clinicopathological study. *Dent. Res. J.*, 12 (4): 307.
- Tian L, Hires SA, Mao T, et al. 2009. Imaging neural activity in worms, flies and mice with improved GAAmp calcium indicators. *Nat. Methods*, 6: 875–881.
- Turner AP. 2013. Biosensors: Sense and sensibility. *Chem. Soc. Rev.*, 42: 3184–3196.
- USEPA. Mercury Update: Impact of Fish Advisories, EPA Fact Sheet EPA-823-F-01-011, EPA, Office of Water: Washington, DC, USA, 2001.
- Valastyan S, Weinberg RA. 2011. Tumor metastasis: molecular insights and evolving paradigms. *Cell*, 147 (2): 275–292.
- Verma N, Kaur G. 2016. Trends on Biosensing Systems for Heavy Metal Detection. In *Comprehensive Analytical Chemistry*, Elsevier: Amsterdam, The Netherlands, Volume 74, pp. 33–71, ISBN 978-0-444-63579-2.
- Vu T, Claret FX. 2012. Trastuzumab: updated mechanisms of action and resistance in breast cancer. *Front. Oncol.*, 2: 62.
- Way JC, Collins JJ, Keasling JD, Silver PA. 2014. Integrating biological redesign: Where synthetic biology came from and where it needs to go. *Cell*, 157: 151–161.
- Weng X, Neethirajan S. 2017. Ensuring food safety: Quality monitoring using microfluidics. *Trends Food Sci. Technol.*, 65: 10–22.
- Xiao Y, Lubin AA, Heeger AJ, Plaxco KW. 2005. Label-free electronic detection of thrombin in blood serum by using an aptamer-based sensor. *Angew. Chem. Int. Ed. Engl.*, 44: 5456–5459.
- Yildirim N, Long F, He M, Shi HC, Gu AZ. 2014. A portable optic fiber aptasensor for sensitive, specific and rapid detection of bisphenol-A in water samples. *Environ. Sci. Process Impacts*, 16: 1379–1386.
- Yoshida K, Miki Y. 2004. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Sci.*, 95 (11): 866–871.
- Zhang D, Lu Y, Zhang Q, et al. 2015. Nanoplasmonic monitoring of odorants binding to olfactory proteins from honeybee as biosensor for chemical detection. *Sens. Actuators B Chem.*, 221: 341–349.



2019-1-BG01-KA203-062371

Zhang F, Zhang Q, Zhang D, Lu Y, Liu Q, Wang P. 2014. Biosensor analysis of natural and artificial sweeteners in intact taste epithelium. *Biosens. Bioelectron.*, 54: 385–392.

Zhang W, Asiri AM, Liu D, Du D, Lin Y. 2015. Nanomaterial-based biosensors for environmental and biological monitoring of organophosphorus pesticides and nerve agents. *Trends Anal. Chem.*, 54: 1–10.



Project website: www.digit-biotech.eu

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

DIGIT~Bio~TECH



LO4 GREEN ENERGY & ICT: FROM SMART TO WISE STRATEGIES

Basic level

AUTHOR:
TRAYANA NEDEVA & ANNA KUJUMDZIEVA



Contents

The Green energy	3
What is green energy?	3
1. Green energy products work	4
2. The types of green energy	4
Energy Consumption in ICT Sector	6
1. Green Energy Provisioning for ICT	7
2. Green ICT and ways of Greening ICT	7
Reduction of energy consumption and gas emission	12
The future ICT trends: from green to wise.....	17
Conclusions	19
References	20



The Green energy

Nowadays, new regulations are in preparation aiming to propose contact code information for specific fields and also for individuals to find alternative sources of energy. These alternative sources have to provide power for private and public buildings, and at the same time generate a low number of toxic compounds. The so-called "going green" approach. Different alternative types of energy have been designed, like solar and nuclear power, with the purpose to save the planet because the toxic emissions accompanying the production of the traditional ones are a huge problem since they affect badly world life.

Along with the widely investigated toxic actions of global warming in recent years, it is vague in parallel to the harm provided by other resources used in the production of food and maintenance of clean water. Society has referred to materials like coal, oil, and even kerosene to ensure the needed energy. The fossil fuels, coal, oil, and other resources used for power production emit harmful side effects. These fuels are non-renewable and contaminate the environment and the atmosphere, impacting the sources needed for the survival of the species inhabiting our planet. As these sources are naturally restricted, troubles for their shortages and access are growing, the worst among them is their harmful effect on the environment. So, the use of these conventional sources of energy contributes to global warming. Coal and oil outflow toxic gases into the environment, and in this way endanger general health rising respiratory problems, and diminish the quality of life.

Green energy will help to relieve and smoothen at least some of these problems, and the faster we move to renewable energy sources the better.

WHAT IS GREEN ENERGY?

The human effect on the environment becomes a substantial value of transfer from standard energy resources to renewable ones and it gets up obvious to the users. Thus, green energy is a judgment for better sustainability in the power grid, and this meaning differs from the term renewable resource. Respecting the Environmental Protection Agency (EPA), green energy provides profound environmental value and includes power obtained from solar, wind, geothermal, biogas, low-impact hydroelectric, and certain eligible biomass sources.

Renewable energy implies the same sources as green energy; the energy produced is realized with technologies and products that pose a considerable impact on the local and global environment. When green power is linked with application in different renewable energy projects, they have an input in these technologies.

Green energy is any type of energy produced based on natural resources, like sunlight, wind, or water. It frequently goes out from renewable energy sources minding variations established for renewable and green energy mentioned above. The basic rule important for these energy resources is that they are not harmful to the environment by releasing greenhouse gases into the atmosphere. Green energy produces power from natural resources without negative side effects polluting the Earth's



2019-1-BG01-KA203-062371

atmosphere and surface. Besides solar and wind power, many other environmentally friendly alternatives might ensure the power needed to keep the planet life. Water, wind, and the sun have been used to generate power, heat water, and operate machinery since ancient times. Recently, technologies have enlarged in respect of designs, types, and uses of alternative sources. Though the energy source used can produce some pollution, the greener options minimize harmful side effects. So, alternatives that are currently exploited are based on the new ways for extracting energy from actual fossil fuels and coal but in less harmful manners. This happens through shifting the harmful by-products. However, at present, this has been less than effective and still cannot be achieved in cost-efficient ways. Nevertheless, there are a lot of alternatives that are viable opportunities.

1. Green energy products work

It is accepted that to be recognized as a green energy resource it should not produce pollution, such as is found with fossil fuels. This evidence indicates that not all sources used by the renewable energy industry are green. Thus, power generation that burns organic material from sustainable forests is mind as renewable but due to the CO₂ production by the burning process, it is not green. Green energy sources are usually naturally completed, as set against fossil fuel sources like natural gas or coal, which are developed millions of years. Green sources also often avoid mining or drilling operations which can impair the ecosystems.

The future in energy consumption is focused on the exploitation of a mix of green, renewable, and conventional energy, regardless of the purchased product. Thus, all energy sources in the electric grid are mixed alongside the power transmission grid. For those keen to become green at home and do not possess opportunities for a solar panels array, the mentioned mix is the best way to reduce the carbon footprint associated with energy consumption. It is the most presumable way to rise the large-scale renewable energy investment and it gives more households and businesses access to green energy.

2. The types of green energy

The variety of green energy types are linked with the wide variety of sources (Fig. 1). Some of these types are better suited to specific environments or regions. As a source of energy, green energy often comes from renewable energy technologies such as solar energy, wind power, geothermal energy, biomass, and hydroelectric power. These technologies are acting through different processes.

Find an alternative energy source is the focus of many countries around the world. It is of strategic importance to discover natural and renewable options as an energy source. In some cases, it may be a simple decision to find appropriate architectural designs that keep buildings cool in the summer and warm in the winter. In another, an anaerobic design is used in energy-producing systems to replace fossil fuels with other resources. Thus, reducing carbon emissions, preventing environmental harm, and job-creating are just some of the advantages provided by investing in green energy.

2019-1-BG01-KA203-062371



Solar power

- Solar power is a clean source of energy that comes directly from the sun through a process known as nuclear fusion. Through nuclear fusion, smaller atoms are fused by heat and pressure and formed heavier atoms conserving the energy emitted in the process. This energy reaches our planet via solar radiation, and we can collect and convert it into usable electricity. The most common collecting devices are the solar panels comprising a net of photovoltaic cells that generate energy in a suitable form for utilization.



Wind power

- This energy source is also powered by the sun since the reason for winds is linked with the unequal heating of the atmosphere. This inequality is influenced by the planet's topology, its spin, and the orbit around the sun. Further, winds are modulated by the planet's surface which they are passing over. The spike wind turbines are most commonly associated with wind power. Unlike solar energy, they work on a rather simple principle that creates electricity: the working mode is that the strength in the wind is captured and channeled to electricity generators. Nowadays, the wind is an innovative concept that offers the advantage of no pollutants and use of the generated power in many settings. However, they are noisy and expensive.



Hydropower

- Since ancient times water is used to generate a power called hydropower. The water from a moving source, like a stream or river, rotates turbines generating energy. Before the appearance of commercial power plants, hydropower was used everywhere from irrigation to milling. This process is expensive, it yields less energy than other sources, and creates concerns due to the impact on wildlife and water quality. Nevertheless, because of its efficiency, hydroelectricity is one of the most popular forms of green energy. Due to this energy in 2017 alone, nearly 4 billion tons of greenhouse gases were not emitted into the atmosphere by generating electricity from hydropower.



2019-1-BG01-KA203-062371



Figure 1. The variety of green energy types

Going green means greater funding to solar, wind, and other renewable energy projects, creating technologies to better harness the renewable sources and make them more admissible for people.

ENERGY CONSUMPTION IN ICT SECTOR

Currently, it is estimated that ICT consumes 1.15% of the total electricity supply. It is accounted that the total annual operational electricity consumption for ICT is 242 TWh in 2015. This sum comprises on-site generated electricity (27 TWh) and grid electricity (215 TWh). Besides, the global operational carbon emitted by the ICT sector in 2015 is about 169 M tonnes CO₂. It is equivalent to 0.53% of the whole carbon emission by the energy sector (32 G tonnes) and 0.34% of global carbon emission (50 G tonnes) in 2015. The electricity consumption in the ICT network increased by 31% from 2010 to 2015.



2019-1-BG01-KA203-062371

This subsumes 185 TWh rise for 5 years, which is corresponding to 1% of the total electricity grid supply. The operational carbon emission growth has been 17% for this period.

Applying widely the 5G in near future, the rate of energy consumption increase is going to be even greater.

1. Green Energy Provisioning for ICT

Today, reducing greenhouse gas (GHG) emissions is getting up one of the most important research subjects in Information and Communication Technologies (ICT) due to the disturbing growth of indirect GHG emissions coming from the tremendous use of ICT electrical devices. Thus, solving the ICT GHG problem is linked to improving energy efficiency through energy consumption reduction at the micro-level.

Research in this field is focused on microprocessor design, computer design, power-on-demand architectures, and virtual machine consolidation techniques. The micro-level energy efficiency method will bring an overall rise in energy consumption due to the Khazzoom-Brookes postulate (also known as Jevons paradox) that says: “*energy efficiency improvements that, on the broadest considerations, are economically justified at the micro-level, lead to higher levels of energy consumption at the macro-level*”. Therefore, it could be reasonable that reducing GHG emissions at the macro level is a more appropriate solution. Large ICT companies, like Microsoft which consumes up to 27MW of energy at any given time, have built their data centers near green power sources. Unfortunately, a lot of computing centers are not placed close to green energy sources. For this reason, green energy distributed network is an emerging technology, allowing that losses incurred in energy transmission over power utility infrastructures are much higher than those caused by data transmission. This makes relocating a data center near a renewable energy source a more efficient decision unless bringing the energy to an existing location.

Management and technical policies will be a decision to arrange virtualization, which helps to move virtual infrastructure resources from one site to another based on power availability. This will facilitate the use of renewable energy within the ICT network providing an ‘Infrastructure as a Service (IaaS)’ management tool.

2. Green ICT and ways of Greening ICT

Green ICT is a broad concept, which lacks a general definition. This concept includes the energy efficiency of equipment such as computers, servers, and monitors. Occasionally, it mentioned the production of ICT equipment as well as the recycling. In other cases, it includes ways in which ICT can be used to mitigate the environmental impact of other sectors.



- **Defining Green ICT**

Green ICT has been defined in the literature as “*the using of IT resources in an energy-efficient and cost-effective manner*” or “*an initiative to encourage individuals, groups, and organizations engaged in the use of ICT to consider environmental problems and find solutions to them*”. Green ICT is dealing with the environmental impact of the ICT sector itself. ICT for Green reveals how ICT can be applied to make other sectors greener. The ICT sector gives about 2% of the world's GHG emissions, mainly due to the emissions from the aviation segment. Although, this figure does not look so much, some authors consider that ICT sector emissions are the highest growing one with rising rates of 6% annually. Besides, the environmental impact of ICT is largely neglected, aviation begins to pay attention to the environment decades ago, but the ICT sector has only started to take care of the environment nowadays. Concerning the Green ICT, the main ICT used are presented in Fig. 2. For each of these categories of equipment, there are different options, which can be made to reduce the environmental impact. When discussing the environmental impact of ICT, the most debated topic is energy efficiency. Meanwhile, there is also a question about such estimation of materials used for production, and the way to be done.

Measuring the energy efficiency is the easiest metric to estimate the effect of green ICT, because of the simple way to determine the energy consumption. More energy can at times be used during the equipment production in comparison to its entire lifespan and this is the case for PCs. When producing computers, several metals like aluminum, arsenic, copper, and lead are used. Some of them are hazardous and create needs for handling the equipment, especially during the recycling. It is claimed that 70% of all hazardous wastes are e-wastes. There are regulations for e-waste recycling in a lot of countries, but informal recycling offers a cheap way to deal with this waste and unfortunately is practiced a lot.

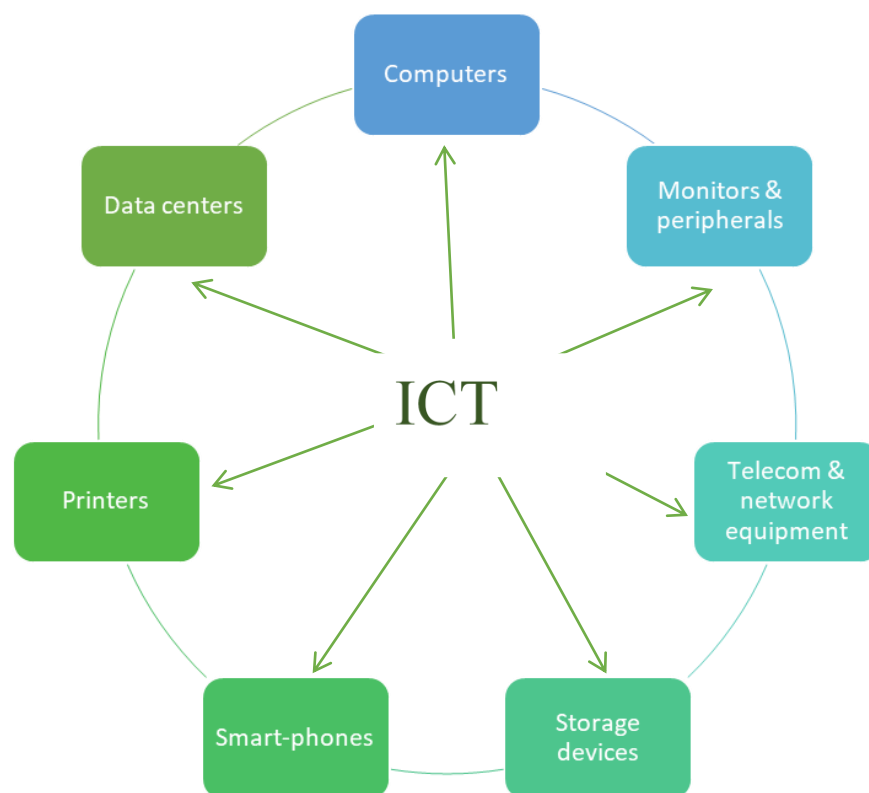


Figure 2. Main categories of ICT used

- **Going green in ICT**

The term ‘green’ is used in everyday language to refer to environmentally sustainable activities. Also, the term ‘green’ repeatedly is used as being sustainable. Both concepts are closely linked but are not identical and to be green is only a part of being sustainable. The United Nations commonly discuss three aspects of sustainable development: the social, economic, and environmental dimensions. The term ‘green’ fits the environmental aspect because both are often interrelated. From an environmental point of view, it is well understood and reasonable especially in the case of linking ‘green’ ICT with economic growth. Thus, there is a reason that sustainability becomes a factor for the rapid growth of the ICT sector; it favors economic growth, ensures wide access to new technologies, and enhances the efficiency of other sectors.

The sustainability or its environmental issues definition given in the UN report "Our Common Future" said that sustainable are those processes that "*meet the needs of the present without compromising the ability of future generations to meet their own needs*".



2019-1-BG01-KA203-062371

To be '*green*' in the ICT sector is partly about taking informed solutions for the way of the use of natural resources.

A good illustration of the term "sustainability" is the way of how today's living can impact the lives of future generations in terms of lack of some natural resources, health problems for those working with production and recycling, as well as the impact on the future generations. So, dealing with ICT there are different ways to become green.

It is necessary to define '*green*' technologies and '*green*' behavior. Green technologies involve parts like virtual servers, allowing a higher rate of utilization to realize a saving of energy of up to 85%, compared to standard PCs. However, the technology that is used inefficiently will not be green. In this way, the employees can have a big importance on the environmental trace. Turning off the equipment after leaving the office, and using technology to make green other parts of life, like hosting video conferences instead of traveling, or sharing different equipment like printers between departments, are examples of green bearing in the workplace.

Being '*green*' in the ICT sector and information management is not precisely the same as being '*green*' in other sectors. Each sector possesses its environmental issues and mitigation strategies. The ICT sector is peculiar, as causing greening of other sectors and it is not sure how the large contribution can be obtained for reduction of the emissions and achieving a better environment. Nevertheless, this environmental impact of ICT should not be overestimated. To be '*green*' in the ICT sector is necessary to become aware of production way and the use and recycling the ICT equipment to be obligatorily sustainable. One of the most important aspects of '*green*' ICT are those concerned with conserving energy, and other resources, such as paper and rare elements. Rare and hazardous materials are used in the production of ICT equipment. Hazardous materials offend not only the environment but also the people working with the manufacture of ICT equipment and its recycling. All these things are linked with ICT and information storage having the biggest impact on the environment. If all equipment is used optimally, it is reasonable to diminish the impact on the environment as well as possibly to save finances.

- **The Importance of being '*green*'**

Nowadays to be '*green*' is significant as a business strategy, which is due partly to consumer understanding of environmental protection. Environmental management is a part of the strategic approach for several successful companies, to involve a new concept of '*green*' management to answer this strategy. Besides, investing in '*green*' innovations and environmental protection is beneficial to companies from a financial point of view. Being '*green*' could increase a company's competitive advantage, as well as bring new market opportunities, and thus make '*green*' companies more profitable.

At the same time, the dangers of the so-called 'green-washing': "*the focus must move away from an emphasis on image to an emphasis on substance*" are a serious warning. Thus, 'green-washing' is the term assigned to companies that try to get the face of being '*green*', but apply cosmetic measures, rather than actual changes in the way the organization operates. So, if organizations want to make a real '*green*' change, it can get a positive effect on society. The greening of organizations can bring new investments to improve the environmental situation, creating jobs and wealth. It is considered that investments in '*green*' ICT can give short-term economic support.

- **Benefits of going green**

A major chance to support greening is cost scanting. Observing the ‘green’ information technology, it becomes clear that data centers and servers dispose a large amount of energy for keeping run and cooing down. ICT resources cannot be used in their full capacity (a utilization rate is about 12.5%). This figure indicates lack of efficiency in the ICT sector. It is found that 86% of the needs can be arranged by 26% of the current energy utilization and this indicates a big room for cost improvement. Also, another part of ICT can propose considerable saving opportunities in usage in an environmentally friendly way. When turning off printers, as well as other energy-consuming equipment, whenever they are not exploited, resources can be saved and the environmental impact can be reduced.

Besides these options, it is acknowledged that moving towards a ‘green’ knowledge society will require structural changes, suggesting that governments should help to enable this process, and doing so benefits for the society as a whole will be realized.

No conventional definition of ‘green’ ICT exists. ‘Green’ topics could be approached from the perspective of a ‘problem’ (focusing on reducing the emission) or from the perspective of a ‘solution’ (focusing on new green solutions). This is reflected in the approach to ‘green’ ICT. This approach is presented in Fig. 3. It encompasses two segments: greening of ICT and greening with ICT.

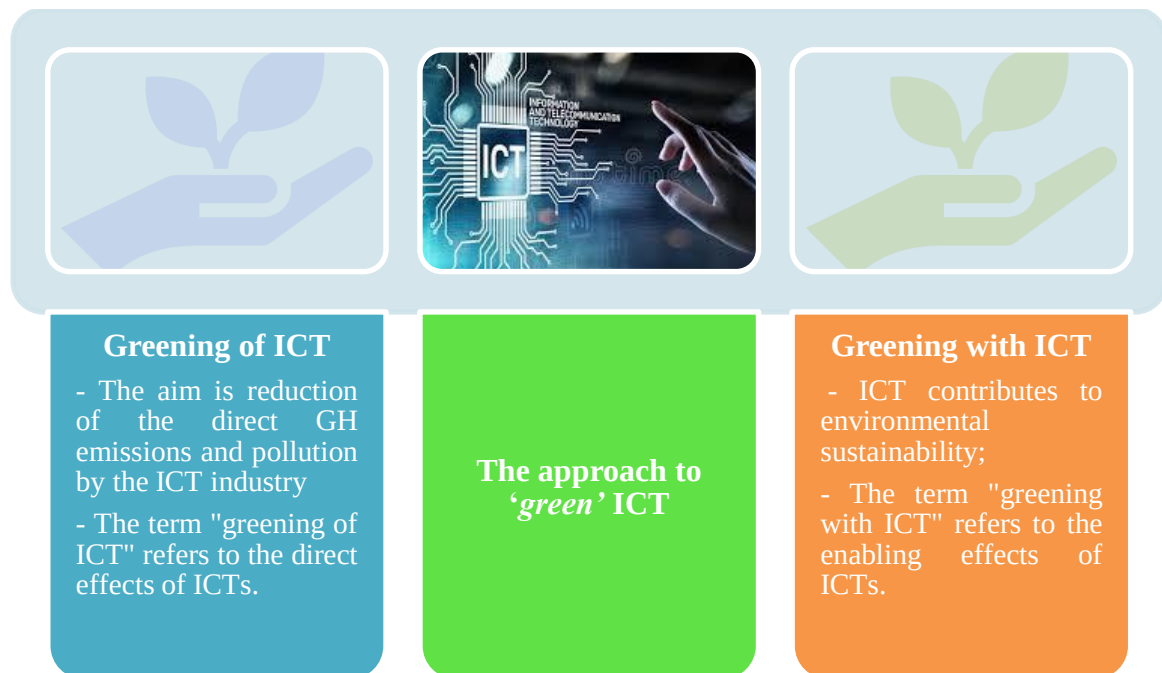


Figure 3. The approach to ‘green’ ICT



2019-1-BG01-KA203-062371

Greening of ICT in a narrower sense refers to ICTs with low environmental burdens, but using ICT as an enabler reduces environmental impact across the economy outside of the ICT sector. Green ICT, as greening with ICT, is a new concept and even the leading countries and stakeholders have only about few years of experience. It is important to note that both the traditional ‘problem’ approach and the new ‘solution’ approach are needed. Pollution needs to be regulated and companies need incentives to address their emissions. However, ensuring that the new generation of solution providers get the right incentives is equally important.

REDUCTION OF ENERGY CONSUMPTION AND GAS EMISSION

ICT can contribute to the reduction of energy consumption and gas emissions through:

- Inventing innovative energy saver systems, technologies and ‘smart’ devices, using ‘smart energy management’;
- Applying energy saver policies using renewable sources, solar energy and photovoltaic, wind energy, bio-fuel, bio-climatic technology, anti-pollutants technology, etc.
- Recycling and reducing e-waste such as old IT systems, chips, PC, hardware, printers, mobile phones, etc.

About 40% of the total energy consumption is due to households. That is why, innovative ‘smart houses’, constructed by green materials, and green architecture exploiting innovative energy sensors are needed. In this context, ICT systems can achieve to measure, manage and reduce electricity consumption and air-conditioning requirements. During the last decade, technical and industrial product manufacturers were essentially obliged to change the direction of their energy consumption, as a result of the economic crisis in addition to the increased environmental awareness of the public. The concern is taken by the producers towards energy reduction via every computing device, from the laptops and mobiles to the data centers, and will be presumably successful. Consumers show their preference for smart devices, new less energy-consuming technologies, renewable energy sources, and updated, more efficient cooling systems with improved energy management software. These products are equipped with official certification to meet or exceed efficiency guidelines.

- **Ways of Greening ICT**

It is possible to use ICT in a way that reduces stress on the environment in comparison to the traditional ways. This can be done through new technologies, techniques, and strategies that allow the consumption of less energy and resources. An important part of the ICTs job is to save information. The need for storage is rapidly increasing because of the growth of Internet usage, new laws and regulations, arranging the rules for keeping the information, and scientific computing. The final part of the stored information is collected in data centers. These data centers use enough energy to partially shift the positive effects that ICT can have on society, for which it is claimed that "*a fraction of energy savings in ICT and networks could lead to significant financial and carbon savings*". To reach energy efficiency in information storage, it is necessary to choose hardware with better energy efficiency or saving energy methods for the equipment use, which are suggested below.

During the last few years, there is a tendency of awareness rising regarding the impact of modern societies on the environment. A lot of environmentally important factors like energy consumption or e-waste are caused by the application of ICT but at the same time, this technology also could solve some other environmental issues. The dual nature of the issue is approached from an integrative perspective with the creation of the Green Computing concept. It is introduced in 1992 by the United States Environmental Protection Agency starting the 'Energy Star' programme. In the beginning, it was enforced on various products such as computer monitors, television sets, and air conditioners. The first well-known result of green computing was the sleep mode option of computer monitors consuming low energy in case user activity is lacking after a certain period. Today green computing gathers a lot of other concepts like confirming computer hardware using, virtualization software, cloud computing or, magnify the energy efficiency of data centers. Green Computing gathers technologies that can contribute to both decreasing the environmental impact of ICT ('Green IT' – greening of ICT) and applying information systems (IS) diminishing the environmental impact of ICT consumables ('Green IS' – greening by ICT). This integrative vision combines two complementary approaches, presented in Fig. 4.

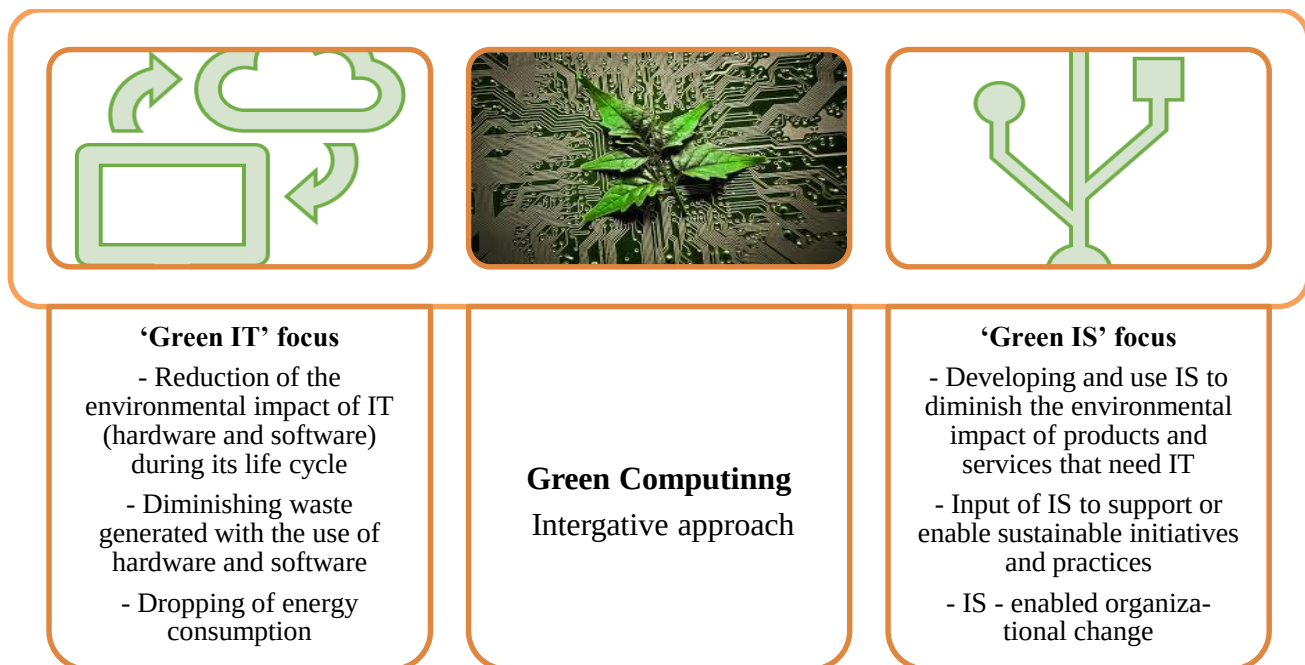


Figure 4. The integrative vision for Green Computing

2019-1-BG01-KA203-062371

The practical implementation of the Green IT and Green IS can be clarified through the following conceptual value models.

Green IT: Value Model

The Green IT value model can assist to reach the goal of environmental sustainability. This conceptual model is grounded on four elements, as shown in Fig. 5.

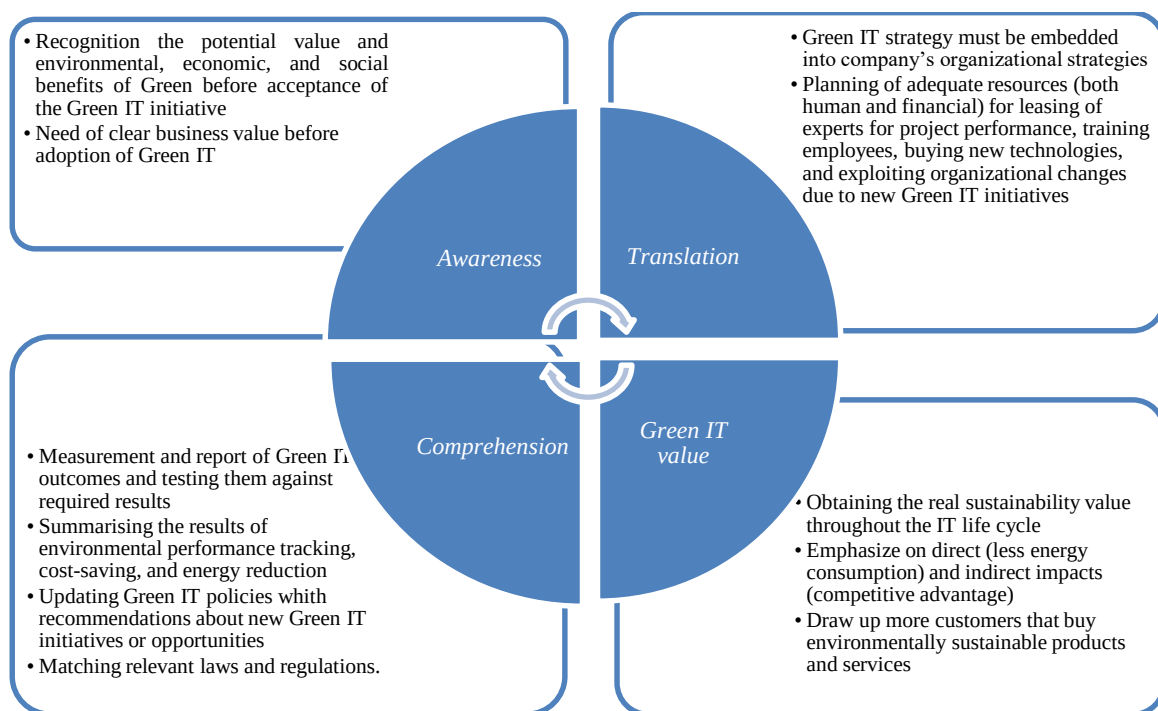


Figure 5. The Green IT value model

Green IS value Model

The Green IS value Model is important for Green IS acceptance and its impact on the firm's environmental enforcement. While companies are under permanent suspense from different regulators, clients, and competitors, some of them are ready to cancel efficiency and effectiveness for environmental issues. Green IS investments to fulfill sustainable business experiences are intended to raise turnover and/or income.

Green IS adoption by a company staff (senior managers) is investigated and explained through a model that describes perception as determined by three basic factors. Three types of strategic initiatives have to be taken into account for IS adoption. All of them are depicted in Fig. 6.

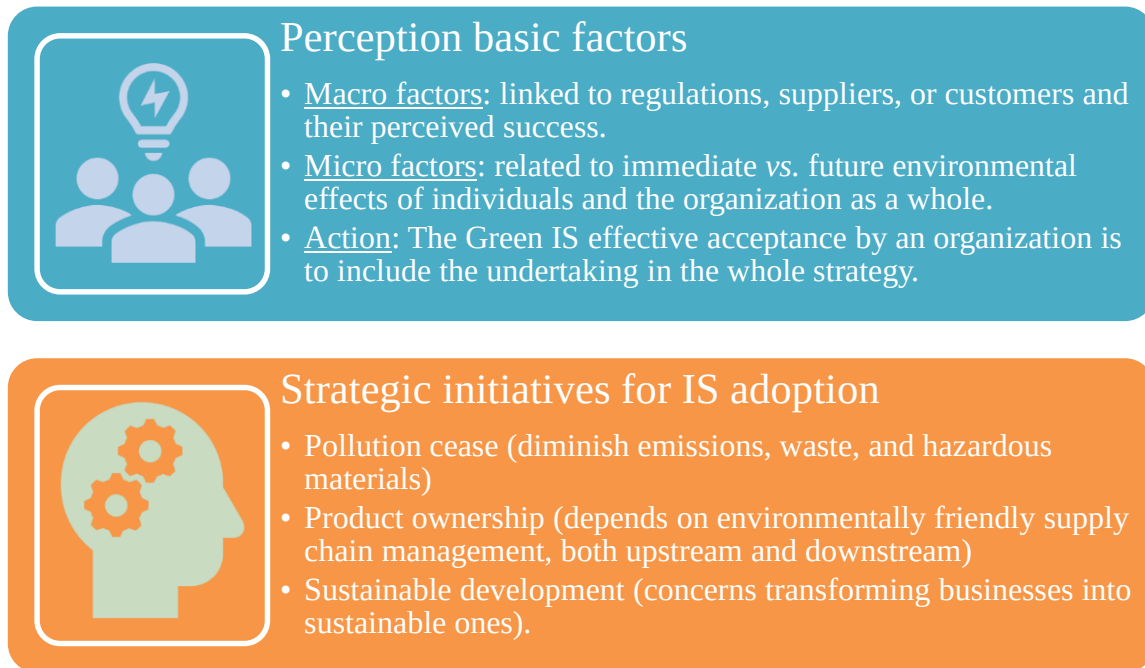


Figure 6. The Green IS value Model

This model could be put into action through a strategy for application of variety of technologies and techniques, like video and teleconferencing, emission management systems, etc.

A framework for management and application of Green IT and Green IS can be established that encompass variety of technologies offering opportunities for reducing the negative environmental impacts in activities by producers and consumers that use ICT. Among these Green Technologies are the following.

Cloud computing

A lot of organizations apply a new computing paradigm - cloud computing, to optimize usage and minimize the cost of their computer servers. In this way, previous costs linked with setting up the IT infrastructures are omitted. Cloud computing allows linking shared infrastructure and balancing IT resources for computing tasks in real-time while reducing gas emissions and maintaining the levels of service. Cloud computing relies on sharing of hardware and software resources that are shared by multiple users and dynamically reallocated per demand. This 'green' technology removes ultimately the need for a company to have an on-premise data center, which has a positive effect on the natural environment and its resources. To date, many vendors provide Green IS-focused cloud services and are experiencing significant financial growth rates *per annum*.



2019-1-BG01-KA203-062371

Computer power management

To save energy in computing, a common action is turning off the equipment in case it is not in operation. There is a so-called open industry standard "Advanced Configuration and Power Interface" (ACPI 2) that represents an advanced Green IT practice. It allows direct power control of the computer operating system and its underlying hardware. This standard automatically rules out components like monitors and hard drives when the inactive period of the equipment takes place. The computer power savings device includes different sleep modes of the monitor, hard disk, system standby and hibernation, and different CPU power states.

Data Center Design

Data centers consume a big percentage of energy, more than 100 times than standard commercial buildings. Thus, the efficient energy recycling design of data centers, like recycling of waste heat, gives a considerable positive impact on energy safety and can be used in the following areas:

- IT systems: increase hardware usage (through virtualization, see below); allow computer power savings modes; buy energy effective tools (e.g., computer power supplies, computer processors, solid-state storage devices, terminal servers);
- Main power systems: power management parts (hardware and software devices for optimizing work and power); rising the usage of renewable energy; use of natural light instead of electricity;
- Cooling systems;
- Air management.

IT Virtualization

The IT virtualization refers to the separation of IT resources through server virtualization and storage virtualization. The server virtualization means operation of many logical computer systems on one physical hardware, and the storage virtualization means summing physical storage from many network memory-tools, on which to be put a single storage device managed from a central console. The IT virtualization decreases costs for hardware, diminishes energy consumption and physical space use. At the same time, it refines software testing and spreading and rises the versatility of hardware investments. Also, it can help in work-sharing: the servers are either busy or put in a low-power sleep state. Thus, virtualization is one of the main approaches for organizations to implement environmental sustainability into IT practices and business.

Material recycling and e-waste

Exploring proper recycling of IT techniques, e-waste is reduced and harmful elements like lead, mercury, and cadmium are discarded from the natural environment. Then it they can be reused and *de novo* production might be omitted. The process of materials recycling is feasible nearly for any computing device and accessory, such as hard disks, printer cartridges, and batteries. In this way, persons and business organizations can extend the life of IT equipment by upgrading tools instead of just changing them.



2019-1-BG01-KA203-062371

Smart grid and smart matters

The Smart Grid technology encompasses hardware and software that ensures more effective exploitation of actual infrastructures for the generation of electricity, its transmission, and distribution. This is an updated version of the electricity delivery system, automatically acting on information flows. The task is to make better the efficacy, safety, and sustainability of the yielding and spreading of electricity by load balancing and peak load management. Devices on the network possess sensors to collect data (power meters, voltage sensors, fault detectors, etc.). Such a device is the Smart Meter, that makes use of digital communication between devices connected to the grid. The collected data relates to energy production and consuming behaviors of both suppliers and consumers. In that sense, real-time information exchange between producers and consumers helps to better control energy demands and reduces the need for energy surplus during peak hours. Tools on the network have sensors to gather data (power meters, voltage sensors, fault detectors, etc.). They read the energy (electricity, gas, etc.) consumption in defined periods and daily sends data back to the public utility e.g., for monitoring and cost estimation. It can also be used to provide information about energy consumption and to set real-time energy prices to consumers. A key feature is automation technology that lets the public utility adjust and control each device from a central location. Benefits include handling alternative sources of electricity (e.g., solar and wind power), smart control for eco-friendly buildings, and in due course integrating electric vehicles onto the grid.

The future ICT trends: from green to wise

Within the ICT industry, there are two main segments:

- Telecom infrastructure, i.e., telecom networks comprised of base stations and access points that provide connectivity to devices (human-centered or machine-centered);
- Mobile devices/terminals, that are communicated by connecting to the infrastructure (e.g., mobile phones, tablets, sensors, actuators, vehicles, drones, etc.).

In the future, billions of devices/terminals will be connected to millions of base stations. Although both segments (network infrastructure as well as devices/terminals) would benefit from renewable energy, the role of renewable energy is much greater in the network segment due to the following reasons:

- Energy consumption of base stations (telecom infrastructure) is much greater than the energy consumption of devices/terminals;
- Global network coverage is important to realize the envisioned networked society and internet of things;
- Base stations have a larger size and are expensive. It is affordable to integrate a renewable energy system with each base station.
- Base stations need reliable and continuous power provision unlike devices (e.g., mobile phones) that can be charged whenever power is available.



To comprehend the future green ICT trends, it is important to understand not only the history of green but also the history of ICT. For the last 20–30 years the ICT infrastructure has been built, the performance and density will continue to improve and increase. However, a turning point has now been reached. The current threshold is similar to the turning point all industrial eras have experienced. During the installation phase, new solutions are used to increase the efficiency in the old system, during the deployment phase the new system reaches maturity, allowing it to deliver entirely new solutions. Initially, the transformation happened in the "information sectors", within e.g., music, video or book sectors, etc., and now we start to see the first signs of a serious change in the "heavy sectors", such as car and mobility, construction, agriculture, and retail sectors, as well as in basic business models.

The shift from improving existing systems to providing new solutions is supported by two trends of ICT development and '*green*' ICT that are important to understand.

- (1) the ICT companies are now influential economic players. For the first time in history, an ICT company – Apple – was the largest company in the US. Apple overtook Exxon Mobile, demonstrating that ICT companies can no longer be ignored by policymakers.
- (2) the ICT companies are now part of a ubiquitous network that is connecting almost everyone and almost everything on the planet. Nowadays, more people are connected than during any other time in human history. By 2020, there are about 50 billion connected devices, and the society is gaining access to data and experience transparency that is fundamentally different from what any society has ever experienced before.

One of the major challenges is that the new '*green*' ICT solutions have to compete in a regulatory environment encompassing regulations created for the 19th-century industrial structure. It has also to deal with the predispositions among people unfamiliar with the speedy development of ICT solutions. In fact, the current technological development is so fast that society members of any kind - from policymakers to business leaders, to economic experts are witnessing how the whole procedure – from an idea to full-scale implementation, is taking place just for few decades. To clarify the situation, the phases of a long-term disruptive solution must be studied. At the beginning, something triggers an idea that spreads, e.g., the first personal computers. These brought about next ideas, such as paperless offices, virtual meetings, etc. All these activities excited people. After a while, working prototypes were introduced and many companies invested in very expensive prototypes of videoconference equipment. The technology was too new and no viable business model was used, instead, these prototypes were bought and managed by the companies themselves. Progressively, many decision-makers had been excited by the idea of ICT as a disruptive force in different areas, that they finally thought it would never come to happen.

A trend today is that policymakers do not consider significant changes that will be made by '*green*' ICT. Many are still planning to invest in new coal power plants because they use the same economic model as they used previously. For instance, Siemens Germany has announced that they support a 25% target for renewables and are ending their nuclear power business.



Conclusions

The history of ICT's development shows that society now is at an inflection point where ICT solutions move towards creating new solutions instead of making old systems better. Two important trends concerning ICT development could be pointed out:

- (1) The fact that ICT companies now are economically powerful and serve as a source of both economic and political capital.
- (2) ICT solutions and companies now are so abundant that new clusters of solutions providers can emerge. It is important to understand what ICT solutions actually can deliver.

'Green' ICT includes the use of ICT solutions to support smart growth. Consequently, to observe 'green' ICT, the context of current 'green' ICT trends should be comprehended. While until the 21st century, 'green' was seen as nature conservation or pollution control, a significant shift took place at the beginning of the 2000s. A focus moved from a problem perspective (pollution control) to a solution perspective and a new generation of business leaders saw the opportunity to link the need for resource efficiency and sales of new products and services.

A parallel shift is observed in the ICT development. It moves from improving existing systems to providing new solutions. New ICT solutions were created to support energy efficiency and 'green' growth. One of the most popular examples for ICT-driven solutions are the e-books, smart grids, electric cars, online meetings, etc.

It's positive that many citizens have begun to realize the concept of human-caused climate change and resource depletion. They consider as well the imperative necessity of acting on this matter. This understanding has led many people to make changes of their personal lifestyle. The "living green" tendency includes many aspects such as green constructions, renewable energy use, energy-saving at home, the extended use of eco-friendly products, a recycling approach. Here to add are the so-called sustainability checklists, designed to help households to assess how sustainable they are, and to offer suggestions for increasing home sustainability. Moreover, clever use of e-services can be a tool for less energy consumption in everyday life and at work. For instance, this is the case of paper use - the nowadays correspondence is substituted by digital formats using the Internet and smart devices. The production and distribution of new products and services show the tendency for minimization of the needed energy, estimated by carbon footprint. Another example is the substitution of traditional conferences with online ones that impose a direct positive effect on environmental protection and reduction of the GHG emissions because of reduced transportation services. It is reported that teleconferences can avoid the production of approximately 540.000 tn CO₂ per year; this is the cost of the air transportation of people.

The use of 'Green' ICT tools and services through broadband/5G Internet contributes to the environmental and societal welfare with the decrease of cost and time to access government offices (24/7 services), energy savings (no transportation), and restriction of pollute emission (carbon footprint). Similarly, in the sector of e-commerce and e-business new innovative business solutions are in favor of either the entrepreneur or the final customer.



References

- Andreopoulou ZS. 2012. Green Informatics: ICT for Green and Sustainability. *Journal of Agricultural Informatics*. 3, 2, 1-8.
- Berl A, Gelenbe E, Di Girolamo M, Giu G. 2010. Energy-Efficient Cloud Computing. *The Computer Journal* 53(7), DOI: 10.1093/comjnl/bxp080
- Brush, K, Kirsch B. Virtualization. <https://searchservvirtualization.techtarget.com/definition/virtualization> 22/03/21
- EU ENERGY STAR programme; https://ec.europa.eu/energy/topics/energy-efficiency/energy-efficient-products/energy-star_en
- Environmental Technology. http://en.wikipedia.org/wiki/Environmental_technology 22/03/21
- Gholami R, Sulaiman A, Ramayah T, Molla A. 2014. Senior Managers' Perception on Green Information Systems (IS) Adoption and Business Value: Results from a Field Survey. *Information & Management* 50(7):431-438, DOI: 10.1016/j.im.2013.01.004.
- ICT for Sustainable Growth: Energy Efficiency of the ICT Sector. DAE Actions. https://ec.europa.eu/information_society/activities/sustainable_growth/ict_sector/index_en.htm 22/03/21
- Joumaa C, Kadry S. 2012. Green IT: Case Studies. *Energy Procedia*, 16, 1052 – 1058.
- Klimova S. 2016. Systematic literature review of using knowledge management systems and processes in green ICT and ICT for greening. Conference: International SEEDS Conference, Leeds Beckett University, UK, 1-21.
- Malmodin J, Lundén D. 2018. The Energy and Carbon Footprint of the Global ICT and E&M Sectors 2010–2015. *Sustainability*, 10, 3027; doi:10.3390/su10093027
- OECD Towards Green ICT Strategies: Assessing Policies and Programs on ICT and the Environment. <http://www.oecd.org/dataoecd/47/12/42825130.pdf> 22/03/21
- OECD countries agree to tackle global environmental challenges through information and communication technologies (ICTs). http://www.oecd.org/document/26/0,3343,en_2649_33757_45073498_1_1_1_1,00.html 22/03/21



2019-1-BG01-KA203-062371

OECD. Towards Green ICT Strategies: Assessing Policies and Programmes on ICT and the Environment. <http://www.oecd.org/dataoecd/47/12/42825130.pdf> 22/03/21

Porter ME, Kramer MR. 2006. Strategy and Society: The Link Between Competitive Advantage and Corporate Social Responsibility. Harvard Business Review, <https://hbr.org/2006/12/strategy-and-society-the-link-between-competitive-advantage-and-corporate-social-responsibility>

Report of the World Commission on Environment and Development: Our Common Future. Brundtland, 1987

UN ECONOMIC and SOCIAL COUNCIL. Sustainable Development
<https://www.un.org/ecosoc/en/sustainable-development>

UN Report “Our Common Future”
<https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf>

White paper: GREEN COMPUTING. 2016. <https://portail-qualite.public.lu/dam-assets/fr/publications/normes-normalisation/information-sensibilisation/white-paper-green-computing/white-paper-green-computing.pdf> 22/03/21

Williams G, Duncan A, Landell-Mills P, Unsworth S. 2010. Politics and Growth. Dev. Policy Rev., 5-31, <https://doi.org/10.1111/j.1467-7679.2011.00519.x>

Yadav K. 2014. Green Computing - A Necessity Now.
<https://www.acecloudhosting.com/blog/green-computing-a-necessity-now/> 22/03/21



Project website: www.digit-biotech.eu

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

DIGIT~Bio~TECH



LO4 GREEN ENERGY & ICT: FROM SMART TO WISE STRATEGIES

Advanced level

AUTHOR:

ANNA KUJUMDZIEVA & TRAYANA NEDEVA



Contents

Introduction	3
The global trends in green ICT development	4
Local conservation	4
Pollution control and corporate social responsibility	4
Solutions, transformative change, and profit	5
World directions in green ICT policies	7
Green ICT advancement: global level policy	7
The leaders in greening with ICT	10
Alternative approaches to greening with ICT	11
The green informatics contribution to green energy	12
Reduction of energy consumption and carbon footprint along production and use	12
Diffusion of information, education, and training to rise environmental awareness	13
Projects promotion through environmental networks communication	13
Sustainable environmental governance	14
Advanced cooperation for green ICT solutions	15
ICT impacts green and sustainability	15
ICT and the Economy-Defining Technologies (EDTs): KBE/KBBE	16
Green ICT and education	17
Green ICT impact at governmental level	19
References	21



Introduction

Nowadays, ICT plays an important role in the environment protection and fighting climate changes. It has attracted considerable attention of diverse types of international forums. Temperature and sea level rising, as well as the floods incidents and storms are undoubtedly impacting climate change, and influencing also the balance of the ecosystems, water and food supply, public health, industry, agriculture and infrastructure. The measures to combat climate change are focused on strategical aims like: i) energy efficiency enhancement; ii) increase the part of energy used from renewable sources and assure the trustworthiness of energy supplies; iii) ensure the supply of energy products and services, and sustainable production of green products.

The energy market today is undergoing serious reforms due to the introduction of new advanced energy technologies. They cause continuous environmental problems, rising needs for European and international cooperation. In this aspect, various intergovernmental agreements have been concluded to sharp and harmonize the organizational and legislative framework of the energy markets. Along the increasing attention for global climate change and related to the energy markets, the green Information and Communication Technologies (ICT) has been proposed as one, in which the environmental impact is taken into account in the design of new systems and technologies.

The topics of “Green informatics” and “Green ICT” are frequently discussed and the interest in ICT’s potential needs to be better appreciated and to focus the attention it deserves. The Green Informatics includes design, construction, and information diffusion techniques and offers optimization of environmental governance, in its efforts to save the natural environment. In this way, it contributes to successful management of the natural resources regarding sustainability, taking into account as well the energy requirements, in particular the alternative energy sources.

The web technology and broadband Internet along with web-based projects are penetrating in a great speed our society and a huge amount of information moves across the WWW worldwide. The Green Informatics are ICT tools, services and technologies deal with green practices and green manners either in the ICT industrial sector or with the ICT users. This can contribute also to the preservation and recovery of the environment as well as to the promotion of the quality of human life. Thus, the concept for “Green Informatics” has turned into a synonym to eco-friendly technology and software tools like Virtualization, Recycling and Telecommuting.

At present, ICT unifies the electronic services (e-services) - broadband network infrastructure - mobile services, and wireless technologies. This mergence led to the development of instruments, products, services and technologies with increased social network opportunities, available 24/7, worldwide in all sectors of human life.

Broadband has been the entrance to the networking economy. Its abilities to convert the daily processes into work and life opens new business prospects for development nowadays, when a lot of countries are fighting to save their economies during global economic crisis.

The services for stabile permanent access to Internet assure reliable delivery with a great speed of Internet of high quality almost all over the world. Part of the web-based products commonly used to assure e-services are e-learning, e-working, e-banking, e-voting, e-government, e-commerce, e-shop, e-



2019-1-BG01-KA203-062371

research, e-medicine and e-payment. Recently, useful mobile broadband services (m-services) adapted to the needs of the people were set up. They combine elements of user-generated content with network-based promotion. Through modification of economic and social measures, the mobile technologies support the sustainable development through green banking, green commerce, green governance, green constructions, etc.

The global trends in green ICT development

In order to understand the global green ICT trends, the current dynamics and directions of different sub-trends should be understood. The context of current green ICT trends should be explained, including the evolution of green trends and the background of ICT. From the perspective of green ICT, it is important to distinguish three different green trends, each of them alive today, but with different logic and history.

LOCAL CONSERVATION

The first green movement (1860-1960) followed the idea that nature was static and should be protected against Industrialism. The main focus was on the creation of national parks, nature was primarily seen as an object of study and a place for recreation. Protecting nature just because it is beautiful and because it offers a place for recreational time spending is still a major part of the green agenda in many countries. This type of link to green ICT can be seen even today when companies plant trees to improve their image. There are still companies who think that a donation to a conservation project is a key part of their green work and many of the major environmental NGOs still approach ICT companies as a source of funding for conservation projects.

POLLUTION CONTROL AND CORPORATE SOCIAL RESPONSIBILITY

During the period 1960-2000, a different trend is shaped, as instead of protecting individual natural areas, companies and policymakers considered the industrialization as on a collision course with the planet. This trend saw companies mainly as a source of environmental problems therefore, rules and regulations were created to minimize the negative impact of companies. The response of most companies was to establish environmental health and safety (EHS) staff and corporate social responsibility/public relations (CSR/PR) staff. Focus was on the end-of-pipe technologies and communication. Many companies, and especially business associations, considered environmental regulations as a threat to their business and this perspective still exists in many processes related to green issues. Several governments still define green or environmental technology as the end-of-pipe technology. The renewable energy is ever increasingly similarly regarded. But some include transformative low-carbon



2019-1-BG01-KA203-062371

ICT solutions, such as teleworking, e-books, smart control systems for buildings, even though these are of key importance to reduce emissions and the need for natural resources.

SOLUTIONS, TRANSFORMATIVE CHANGE, AND PROFIT

At the early 21st century the need for transformative change and sustainability moved to a new phase. Instead of biologists and environmental organizations identifying problems, a new generation of stakeholders started to present solutions. This new trend of green thinking was the result of a number of converging trends (see Fig. 1). In addition, the new generation of entrepreneurs and business leaders see the opportunity to link the need for dramatic resource efficiency with the sales of new products and services. Instead of approaching green as a threat that only demands an incremental improvement in existing systems, these entrepreneurs have realized that new smart solutions, which challenge existing business models and ways of providing services, are ready. Underlying this shift is ICT development and its targeting of new areas. E-books, smart grids, electric cars, video conferencing and mobile applications are just a few examples of ICT driven solutions.

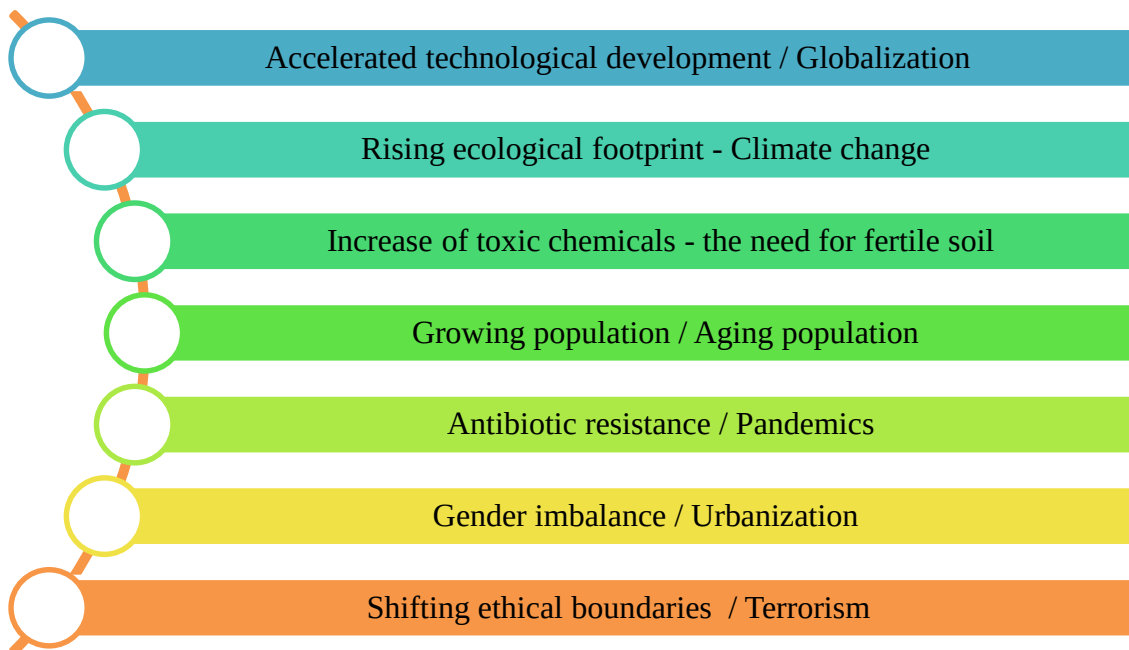


Figure 1. Converging trends of green thinking

This trend of green thinking often creates significant pressure within existing structures in respect to the older trends on multiple levels. An overview of the different scopes and approaches related to

business and policymakers is presented in Fig. 2. It is a matrix that illustrates the strains between the different institutions with different roles, as well as the pressure due to the discrepancy between the problem perspective and the solutions perspective.

In the upper half of the matrix, companies focus on society needs and apply sustainability as a driver for innovation and profit. The actions within this part would be possible because of the unified efforts of entrepreneurs, business leaders and strategic players within the government. However, green alone is rarely the key driver due to the discrepancy among various stakeholders: governments, NGOs, media, etc. It is hard to implement greening with ICT solutions, since they require collaboration between multiple stakeholders. The result is a lot of green initiatives that focus only on ICT companies, as the source of emissions. Besides, companies that are solution providers implement a number of green solutions, but neither the solutions are called green nor the people using these solutions are aware of the green benefits from them. Thus, a shift from the traditional ‘problem’ approach to the new ‘solution approach’ is needed.

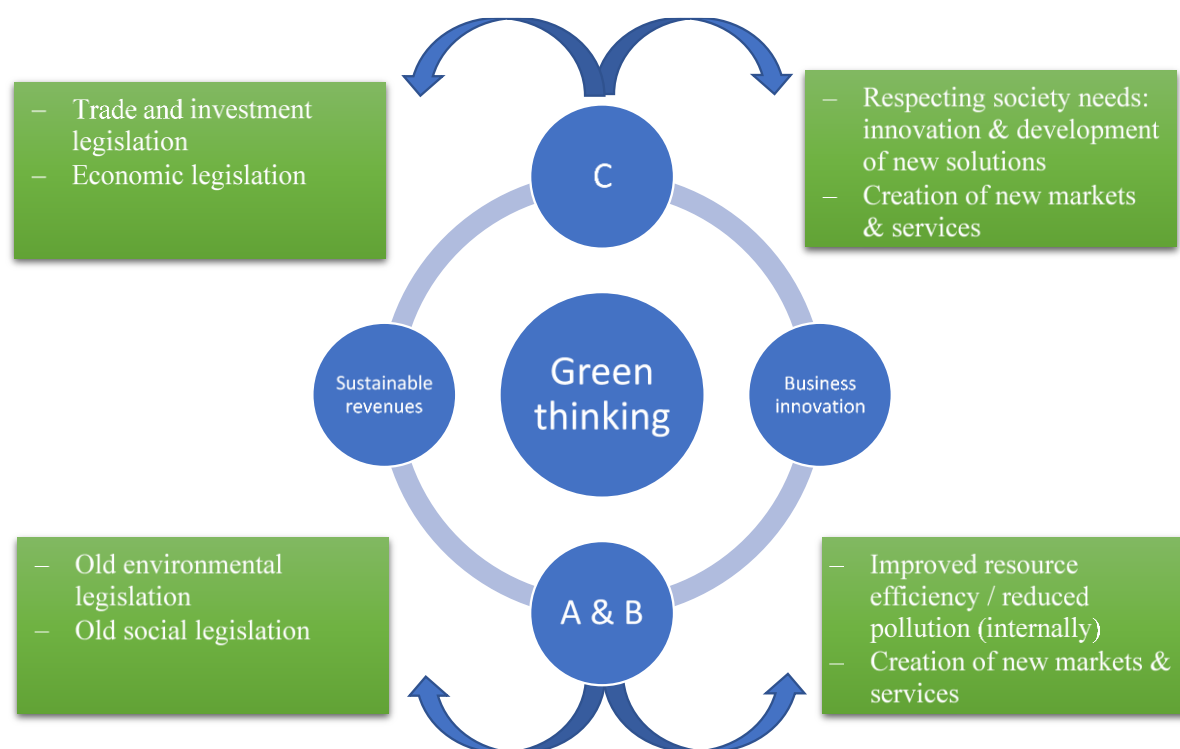


Figure 2. Current and future green thinking trends. Legend: A – Local conservation; B – Pollution control and corporate social responsibility; C - Solutions, transformative change, and profit



The first green movement (1860-1960) followed the idea that nature was static and should be protected against Industrialism. The main focus was on the creation of national parks, nature was primarily seen as an object of study and a place for recreation. Protecting nature just because it is beautiful and because it offers a place for recreational time spending is still a major part of the green agenda in many countries. This type of link to green ICT can be seen even today when companies plant trees to improve their image. There are still companies who think that a donation to a conservation project is a key part of their green work and many of the major environmental NGOs still approach ICT companies as a source of funding for conservation projects.

World directions in green ICT policies

An analysis of the global trends in green ICT has been performed with the purpose to find out what are the predominant trends in green ICT policies at global scale from medium- and long-term point of view. The analysis was carried out applying a desk research approach. First, relevant documentation related to green ICT was identified and, second, based on this documentation, the corresponding green ICT related policy trends at the EU and global level were outlined. The documentation set encompasses variety of EU and OECD strategy documents (e.g., OECD Green Growth Strategy, EU 2020 strategy, Digital Agenda etc.), as well as on researches and case studies performed before.

GREEN ICT ADVANCEMENT: GLOBAL LEVEL POLICY

Both green ICT and greening with ICT comprise a new concept that has been on the economic and societal agenda since a decade. The area of greening with ICT is so new that it is hardly possible to launch a good practice for a single country to be securely followed. Most policymakers, major researches and business groups unambiguously declare that the greening with ICT is significantly important. However, the real programmes and policies at governmental level, NGOs' work, research at universities, business initiatives, etc. are rather concentrated on the direct effects then on the long-term impacts.

A number of studies have shown that the majority of the initiatives at both governmental and business level focus on greening of ICT not greening with ICT. A study, performed by the OECD, published in June 2009 showed that most *“Green ICT” initiatives concentrate on the direct effects of ICTs themselves rather than tackling climate change and environmental degradation through the use of ICTs as an enabling or “smart” technology*. The OECD analysis showed that the government programmes include initiatives to consider the enabling effects of ICTs. The ICT applications used for the dissemination of environmental information, for smart transportation, grids, and buildings are the most commonly promoted. However, software for energy optimization or smart engines have been less indorsed.



2019-1-BG01-KA203-062371

There is a gap between the policymakers and ICT companies that want to support a good image in front of the society, who focus on greening of ICT, and the needs for implementation actions that will lead to greening with ICT. In this context, solutions are necessary to be implemented by the institutions and frameworks in the form of new rules and regulations to support greening with ICT.

There is a variety of opportunities and challenges ahead the policy makers and businesses stakeholders along this implementation. To be best understand, these possibilities and challenges must be considered in respect to the main ICT players contributing to both greening of ICT and greening with ICT ideas. These interrelations are depicted in Fig. 3.

It can be seen that the majority of policy makers, foremost researches and business groups definitely indicate that the greening with ICT is meaningfully more important than greening of ICT.

The new concept of greening with ICT most commonly does not require only business relations but policy makers that create a new legislative framework and relevant guidelines for its implementation. However, currently the real initiatives of the business and the actual policies still focus on greening of ICT, persuading the direct effects of it.

2019-1-BG01-KA203-062371

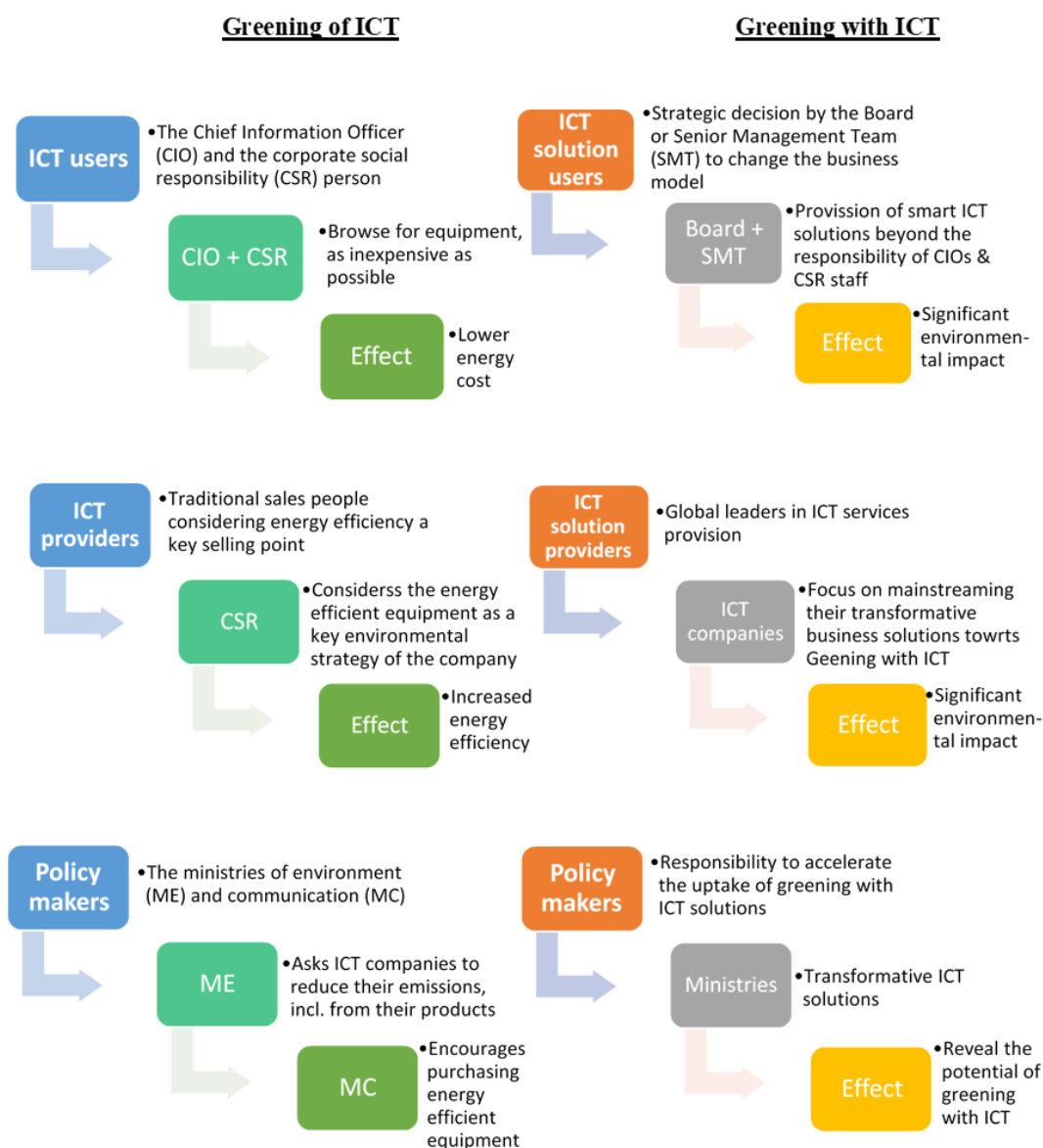


Figure 3. Main ICT players contributing to greening of ICT and greening with ICT ideas

A global trends analysis has been performed to make an overview of the long-term tendencies in global policy making in respect to green ICT. A representative selection of the most relevant and comprehensive documents worldwide has been evaluated. It is assumed that within last two decades, greening with ICT has moved ahead from being an almost fiction idea into being an important subject. In the 1990s, greening with ICT matter has not been included in global strategy documents or policy



2019-1-BG01-KA203-062371

making. The role of ICT in sustainable development had shaped clearly among thought leaders at the beginning of the new 21st century. However, it was still more or less overlooked by both the policy makers and the business. The reason for this underestimation was grounded on the fact that the executive power was given to leaders from ministries of environment, and their efforts, and solutions, were sector-focused.

However, at that time the changes in policies on strategy level has started to officially recognize the businesses as an active party in solutions development; the power of ICT in supporting a more environmentally sustainable development was also acknowledged. Most governments and organizers that had mainly focused on the building of ICT infrastructure and the standards and rules its operation requires, embarrassed the new idea that ICT could be used to support an environmentally sustainable development. Accordingly, the companies have started to be seen as solution-generators that contribute to accelerate sustainable development instead of pollution-makers, which should reduce their emissions.

Starting from 2007, first attempts to include governments and businesses in a discussion about greening with ICT were organized. A year later, the chair of G8 brought green ICT to the forefront of the discussions' agenda for the first time. The G8 leaders united around the immediate necessity the world to diminish carbon emissions, contributing strongly to global warming, by at least 50 percent by 2050. It was then the issue of green ICT was opened and OECD and EU confirmed the tendency through increased focus on greening with ICT.

In this way, greening with ICT has been acknowledged as a policy area. However, between the new ideas generated and their practical realization, there is still an "implementation gap. Since not all the ideas become reality and the shift from words to actions does not take place very smoothly, the businesses are facing serious challenge how and when to make their contribution to the greening with ICT trend.

Greening with ICT strategy is growing more stable to become a significant element in the mainstream policy making. The global climate meeting in Durban, 2011 and the Rio+20 conference in 2012 were the first forums that indicated this. More and more leading countries want greening with ICT to become a part of the global agenda for planet saving. They recognize competitive advantage in developing and exporting greening with ICT solutions, and the greening with ICT is already deeply integrated into their policy making.

THE LEADERS IN GREENING WITH ICT

In general, all main international stakeholders are key players in the greening with ICT area. At present, however, there is no a single leader; by contrast, there are many players that are focused on the process of greening with ICT with varying intensity and success. Indisputably, OECD and the European Commission are very important stakeholders but many innovative ideas have been emerging from independent business coalitions. Thus, there is a tendency for greening with ICT by economic clusters that focus on implementation at a local level.

The modern industrial society is built through organization of the society around certain economic sectors. These sectors were created by establishing new clusters in the pre-industrial society,



2019-1-BG01-KA203-062371

and nowadays are called industries, even though they are combinations of several different skills. Crossing the era of the post-industrial / knowledge-based economy, the society needs new clusters to guarantee the new supply chains with new materials and new productions methods, the to secure the new business models with fresh ideas.

An example for this transition and the advancement of the clusters to a leading position is the energy sector, in particular the renewable energy. In renewable energy sector traditional utilities failed to deliver solar and/or wind solutions. It is because a collaboration between construction companies and architecture one and new solutions are necessary, and they often require new business models.

Usually, the ICT companies are very often at the center of the new clusters. However, their largest customers most commonly are one of the biggest polluters. This is a great challenge to the ICT companies and they have to be transformative to issue new solutions that support the business clusters with innovative greening with ICT solutions.

ALTERNATIVE APPROACHES TO GREENING WITH ICT

Traditional green strategies are hardly effective for generation of greening with ICT solutions. In fact, most greening with ICT solutions result from smart strategies that are in search of resource efficiency and innovation in diverse economic areas, e.g., construction, transport, power supply, etc. This requires an approach to greening with ICT subordinated to the global development, discarding synergies and old stakeholders / methods. This very approach is a non-green one, although it also targets greening with ICT. A good example for a non-green approach is the use of procurements, both public and done by companies. As a rule, the companies face “green” and “ICT” trough a one-sided and limited approach; they focus on greening of ICT. This is because ICT companies (those who are selling) have products to put on the market and make profit from them, and CIOs (those who purchase) are trapped by old-fashioned ways of thinking and are besides, are not responsible for solutions like teleworking, virtual meetings and development of new business models.

Both the providers and the requesters of “greening with ICT” solutions very often do so without knowing and/or taking care about the green benefits of the action. Two of the most typical examples of such lack of knowledge/care are Amazon and Apple. Bothe companies are great promoters of greening with ICT; Amazon has contributed largely to the dematerialization of world’s books and magazines, and Apple has dramatically changed world’s music industry and made it more resource efficient. In the same way, smaller companies like Skype and Viber have progressively transforming the business models and habits without regarding their green activities.

The lack of cooperation and understanding between those that know and value the ICT opportunities (ICT companies and CIOs) and those that make new products and services is a big challenge. The problems are multiplied by the lack of strategies and definitions of greening with ICT. Some innovative “smart” solutions are considered “green” regardless of the actual results of their performance. For instance, the development and exploitation of smart grids is a trend where the transformative potential of the new ICT solutions, such as smart grids, is jeopardized by the traditional



stakeholders that rule the agenda: the power utilities label their solutions as “smart”, although they continue to use old business models and large-scale utility structures.

In brief, greening with ICT solutions is mainly directed by smart strategies not by green demand or green policies.

The green informatics contribution to green energy

Green informatics contribute to the environment and environmental sustainability in the following areas:

- Reduction of energy consumption and carbon footprint along production and use;
- Diffusion of information, education, and training to rise environmental awareness;
- Environmental projects and networks promotion through communication;
- Sustainable environmental governance.

REDUCTION OF ENERGY CONSUMPTION AND CARBON FOOTPRINT ALONG PRODUCTION AND USE

The massive introduction of ICTs in everyday life has resulted among others in the increase of greenhouse effect, due to the ‘carbon footprint’ increasing. As per definition, carbon footprint (CF), known as well as ‘Carbon profile’ comprises the overall amount of carbon dioxide (CO₂) and other greenhouse gas (GHG) emissions (e.g., methane, laughing gas, etc.) associated with a product, along its supply-chain, end-life recovery and disposal. In respect to ICT, it refers to the energy needed, and the pollution generated in ICT production processes and within the ICTs use (Fig. 4). The total amount of CO₂ emissions from the ICT industry progressively counts upstream. At the same time, ICT applications are recognized to possess enormous potential in contributing to different performances across the economy and society. They are the right tool of the strategies for the global environmental protection.



	2009	2015	2020
Data Centers	121.30	229.87	369.48
PCs	126.69	222.41	516.55
Mobiles	1.54	3.74	6.58
Gaming Consoles	11.23	26.04	40.22
Carbon Conversion Number (CCN)	1.3	1.265	1.23
Total	260.77	482.06	932.84

Figure 4. Carbon footprint in megatonnes of CO₂ from ICT sectors. Source: Bronk et.al., 2010

DIFFUSION OF INFORMATION, EDUCATION, AND TRAINING TO RISE ENVIRONMENTAL AWARENESS

Diffusion of information, education, and training with the purpose to rise environmental awareness is an approach applied worldwide to help people be up-to-date and understand environmental issues and environmental policies.

The wireless/mobile) internet access is irreplaceable as a tool of information delivery for populations that are isolated or remoted and that lack an access to traditional channels, such as TV, radio, newspapers, magazines, etc.

There are numerous of internet sites, blogs, forums, social network groups, internet polls, etc. dedicated to delivery and sharing environmental information. These information sources are operating from local to international scale, and act as an open tribune for everyone to participate, to offer and share opinion.

Learning and training contribute to enhancement of people's knowledge, skills, and awareness. Suitable learning/training software packages are those offering presentations and educational games, and educational e-services such as e-classrooms, e-learning, ODL, web-based learning, LLL, etc.

PROJECTS PROMOTION THROUGH ENVIRONMENTAL NETWORKS COMMUNICATION

The application of ICT for communicating different projects within environmental networks can be a useful approach for their successful implementation at local, regional, national, and transnational level. The accomplishment of environmental projects requires as a prerequisite effective communication

among participants of various stakeholders, protected through innovative green informatics tools and services. Namely, green informatics makes safe information flow for the purposes of quick and reliable communication. Environmental networks comprise various stakeholders, all integrated through ICT-mediated communication (Fig. 5).

SUSTAINABLE ENVIRONMENTAL GOVERNANCE

Green ICTs have turn into a key factor for public sector's performance because it promotes the advancement in the delivery of information and services and encourages the citizen participation in the decision-making process. In this way ICT helps the government to become transparent, responsible, and operative. The e-governance strategies, initiatives, and developments are grounded on the ICTs. In particular, the governance of natural ecosystems, natural resources and agriculture has to manage a wide range of connections between different environmental elements and decisions of local, regional, national, and international importance, and has to coordinate diverse administrative objects and players, and ICTs helps a lot for execution of these complex tasks.

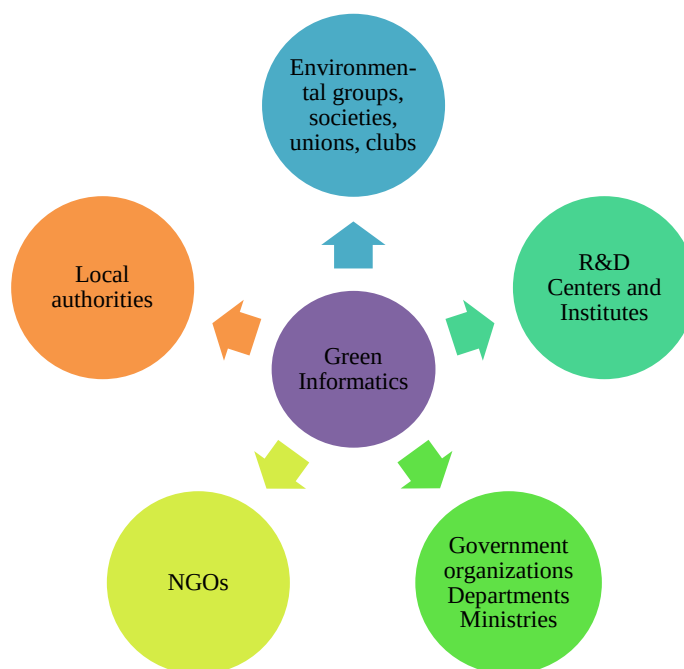


Figure 5. Impact of green informatics to environmental networks communication



Advanced cooperation for green ICT solutions

ICT IMPACTS GREEN AND SUSTAINABILITY

EU has launched the initiative ‘ICT for sustainable growth’, a specific process that focuses on greening with ICT (in addition to greening of ICT). It has determined six policy areas of major priority that focus on Energy Efficiency, Water Management and Climate Change Adaptation (Fig. 6). Thus, it contributes to the development of a more sustainable Europe solving environmental problems and ensuring the sustainable environmental management.

The contribution of Green Informatics to the preservation and improvement of natural environment and resources is focused on the building of surveillance systems that aim to protect and restore natural ecosystems. In addition, it promotes ecosystems potential deployment and introduces prevention actions for its maintenance. For instance, forests and agricultural land are important to climate change mitigation because of the significance of their carbon stock and also for their exchange of greenhouse gases between the atmosphere, soil, and vegetation. The harvest of trees to supply book and newspaper industries leads to the emission of millions metric tons of CO₂ annually. Thus, the innovative tele-detection for forest fires, monitoring and alarm systems, GIS technology, etc., all contribute to sustainable forestry management.

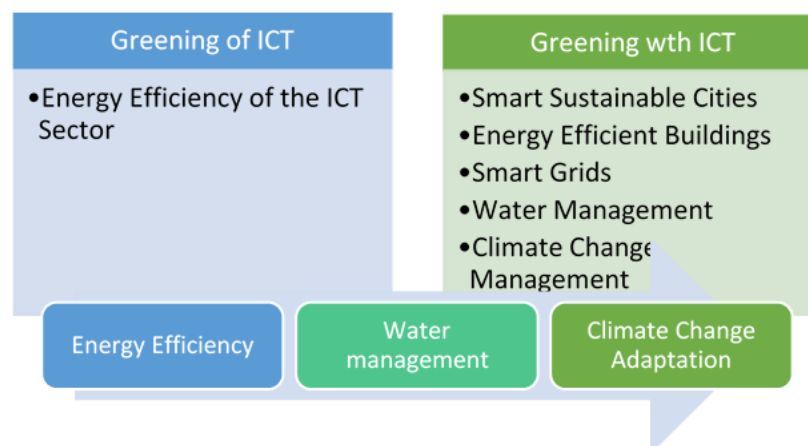


Figure 6. ‘ICT for sustainable growth’ six policy areas

The organization, access and management of the information in the environmental databases is an important factor within decision-making process. Since environmental projects have to manage huge multivariable data sets of interdisciplinary character (meteorological, geographic, biological, economic, etc. data), this has been successfully achieved through ICTs applications and is called environmental monitoring. The main procedure of environmental monitoring is presented in Fig.7. It is a useful tool

2019-1-BG01-KA203-062371

that integrates geospatial technologies aiming to sustain agricultural and environmental observation networks and deploy agricultural and environmental applications

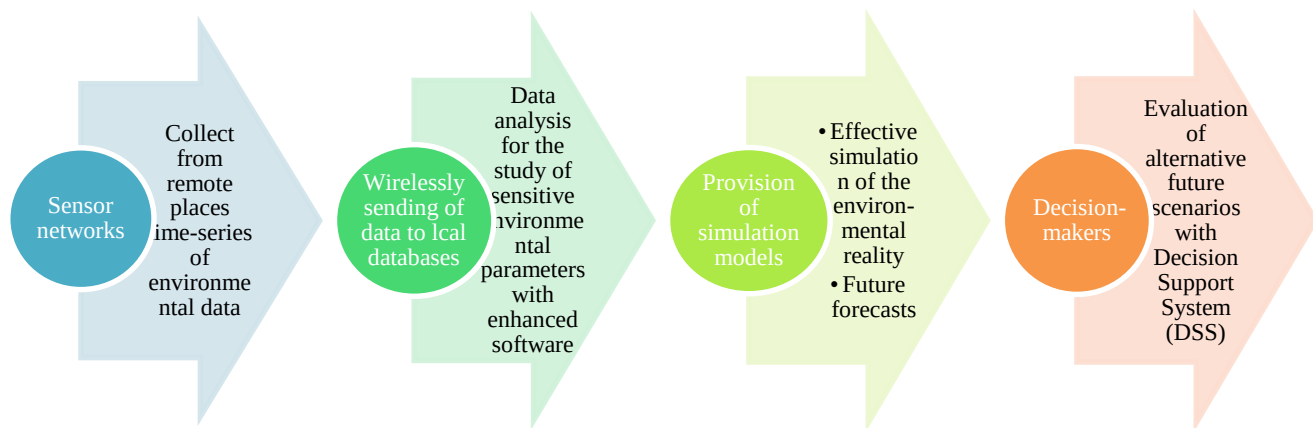


Figure 7. Wise management to environmental sustainability – Green Informatics contribution

The DSSs use defined parameters to provide wise management aiming to environmental sustainability and helping decision making process towards sustainable environmental management. These are Environmental databases, GIS, time-series, multi-variant, and multi-criteria analysis, expert systems, etc.

ICT AND THE ECONOMY-DEFINING TECHNOLOGIES (EDTs): KBE/KBBE

The innovative technologies that constitute the principal technological basis of an economy, are called **Economy-Defining Technologies (EDTs)**. EDTs are always inherent to the corresponding economy, and ICTs are definitely such technology.

Knowledge has always played a central role in any economy. However, in the **Knowledge-Based Economy (KBE)** knowledge has been liberated from its temporal and spatial constraints because of the ICTs. This is what allows knowledge to unfold its powers as a universal resource, leading to the KBE.



2019-1-BG01-KA203-062371

Grounded on this concept, OECD defines the KBE in the following way: “*The knowledge-based economy is an expression coined to describe trends in advanced economies towards greater dependence on knowledge, information and high skill levels, and the increasing need for ready access to all of these by the business and public sectors.*”

Knowledge takes predominance for both individuals and organizations in the KBE.

The **Knowledge-Bases Bio-Economy KBBE** presents a wide range of challenges to ICT. However, the area with potential highest impact of ICT for KBBE is the general and synthetic biology along the dimensions of education, research, and industry applications.

The needs of ICT-driven innovations, which are able to reduce energy and materials used while enhancing the efficiency of business systems, can generate wide opportunities for companies' businesses. The said innovations include software applications (e.g., virtualization technology to implement power savings), and hardware applications (e.g., server with energy-saving properties). In addition, essential industrial infrastructure must be active in order to capitalize on the expanding global market for ICT-based solutions planned for improving the energy efficiency as well as tackling the climate change concerns.

Besides the economic benefits, adopting Green ICT practices in business operations can easily deal with climate change issues that are associated with greenhouse gas emissions. Additionally, it also described that Green ICT can play a crucial role in helping to promote the low carbon economy around the world. The ICT industry can also produce a green image while behaving as a responsible global citizen.

GREEN ICT AND EDUCATION

Green ICT at Higher Education Institution

Currently global warming and climate change are in the front of societal agenda. They turned into a common subject of discussion globally. Climate change consequences arise huge environmental problems and impact energy and industrial policies worldwide.

Green ICT, as a system of initiatives and strategies that reduce the environmental footprint of technology, is able to respond to the needs for implementation of climate change adaptation and mitigation actions. Hence, Higher Education Institutions (HEI) are forced to implement more sustainable approaches to ICT use. This necessity is introduced by the government, stakeholders and society as a whole. Green ICT implementation at University level has developed as key factor to reach the cost-effective solutions and sustenance of ICT.

Moreover, the HEI have deep moral responsibility to increase knowledge, skill and awareness, with the aim to create a sustainable future. Their role in mainstreaming society towards sustainability is indisputable. However, at the same time, they are facing some barriers in Green ICT practical implementation. Therefore, Green ICT is a multifaceted subject, which importance is progressively increasing towards understanding the role of ICT is enabling sustainable practices. For instance, green research and development activities can contribute to reducing environment impact of society by dropping the impact of ICT installations in telecommunication and data-centers, customer offices, homes



2019-1-BG01-KA203-062371

through greening of ICT. At the same time, the impact of society can be reduced by providing various kinds of ICT services through greening by ICT. Thus, HEI need to transform education within sustainability prospect to be able to educate undergraduates to become the ICT engineers our future needs.

Green ICT practices in HEI

The SMART 2020 report stressed upon the capacity of ICT to monitor and maximize energy efficiency, not only within its own sector but also outside it, resulting in considerable emission and footprint reduction. Studies performed to analyze and review the evolution of Green ICT practices in different HEI have proven that the proper ICT deployment contributes to sufficient reduction of GHG emissions. HEI follow environmental sustainability (ES) practices to complete the strategic plans for environmental sustainability through organization of virtual classrooms, digitalization of paper documents, performance of video conference to reduce travel, use alternative clean sources of electrical power, etc.

Constraints for Green ICT Practices in HE

There are various factors that obstruct the implementation of sustainability initiatives of HEI, associated mainly with institutional barriers. These constraints are complex: from old-fashioned contracts of educators, to shortage of equipment, finance and trained tutors, to inadequate environmental teaching methods, lack of motivation towards ‘green behavior’ of teaching staff and students in their approach of using ICT (e.g., by reducing print volumes, using conference calls to reduce unnecessary travel, etc.).

The main constraints that act as barriers in implementation of Green ICT at HEI are listed in Fig. 8.

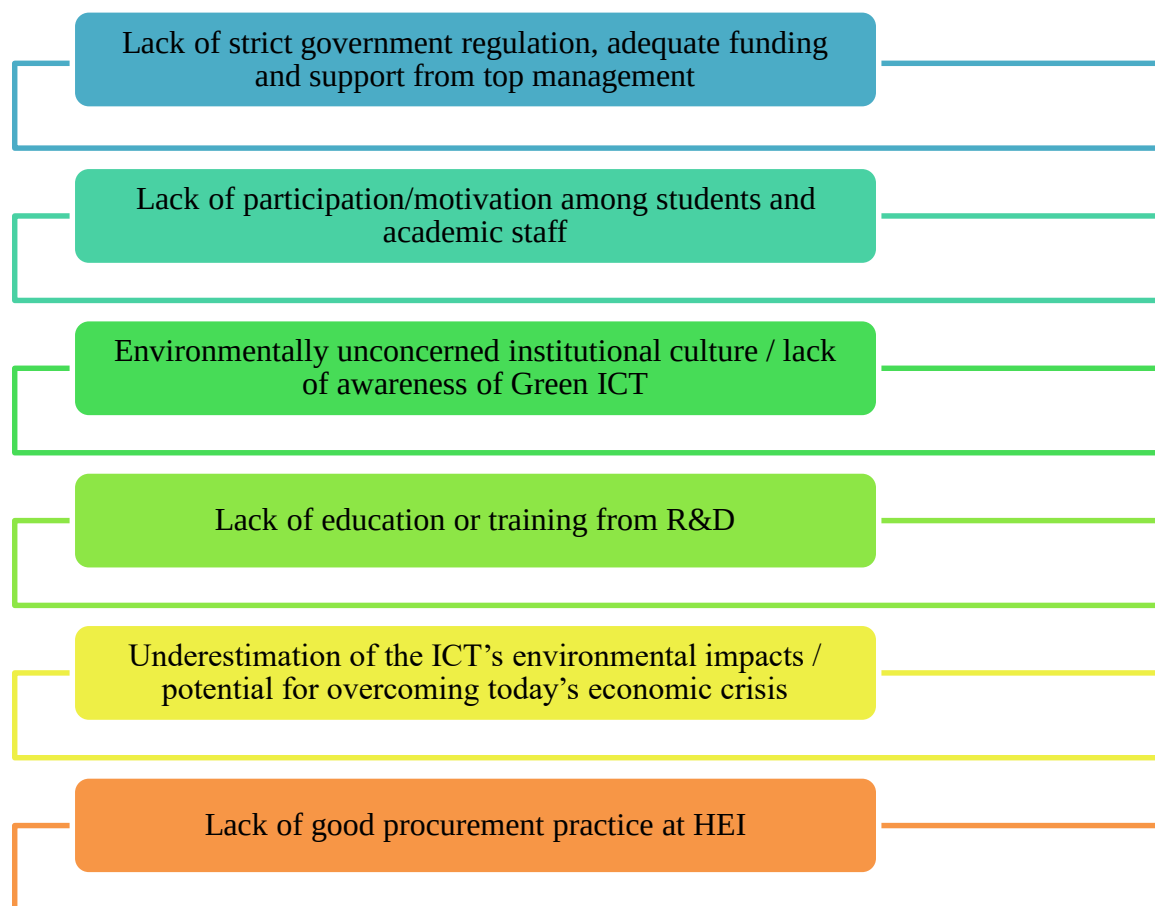


Figure 8. Constrains for Green ICT implementation at HEI

GREEN ICT IMPACT AT GOVERNMENTAL LEVEL

Governments are one of the primary users of ICT and impact substantially the ICT industry. Due to its potent influencing power, the governments must play a leading role in acceptance of Green ICT technologies, improving its operational efficiency, and encouraging societal environmentally-aware and sustainable culture. Thus, it is essential for the governments to establish close collaboration with the ICT industry, and there are several areas that the governments have to seek for better opportunities for the Green ICT implementation. These are outlined in Fig. 9.

2019-1-BG01-KA203-062371

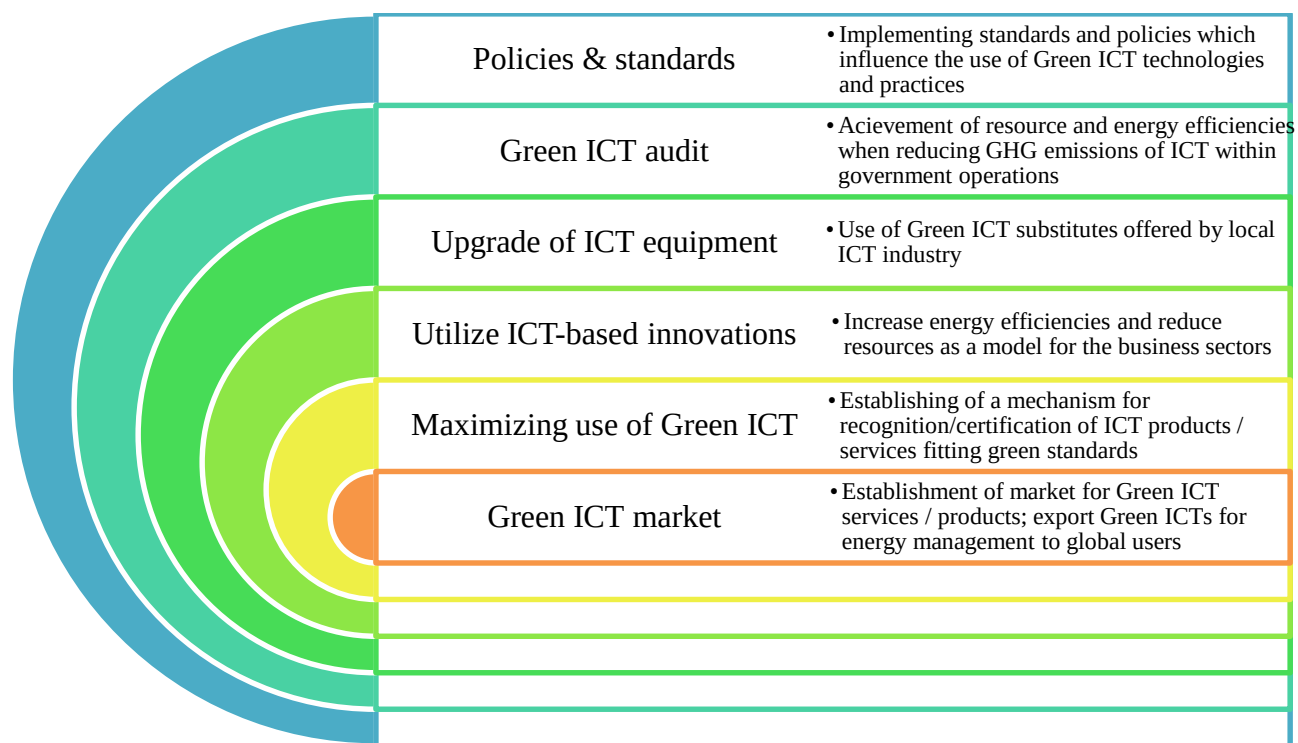


Figure 9. Areas for governmental collaboration with the ICT industry and Green ICT implementation



References

Andreopoulou, ZS 2009. Adoption of Information and Communication Technologies in public forest service in Greece. *Journal of Environmental Protection and Ecology*. 10(4): 1194-1204.

Bibri SE. The shaping of ambient intelligence and the internet of things: historico-epistemic, socio-cultural, politico-institutional and eco-environmental dimensions. Berlin: Springer; 2015.

Broadband Commission. 2012. The Broadband Bridge: Linking ICT with Climate Action for a Low-Carbon Economy. Available at: <http://www.broadbandcommission.org/Documents/Climate/BD-bbcomm-climate.pdf> (18/9/2012).

Bronk, C, Lingamneni A, Palem K. 2010. Innovation for sustainability in information and communication technologies (ICT). Internal report, Rice University, <http://www.rice.edu/nationalmedia/multimedia/2010-10-11-ictreport.pdf>

EC. 2007. Carbon Footprint – What it is and how to measure it. European Platform on Life Cycle Assessment. Available at: <http://lct.jrc.ec.europa.eu/pdf-directory/Carbon-footprint.pdf> (18/9/2012).

Ernst & Young Baltic AS. 2011. The Role of Green ICT in Enabling Smart Growth in Estonia. Available at: http://www.pamlin.net/new/wp-content/uploads/EY_MKM_Green_ICT_study_2011_FINAL-REPORT2.pdf (18/9/2012).

ICT ENVIRONMENTAL IMPACT EC Rolling Plan 2021
<https://joinup.ec.europa.eu/collection/rolling-plan-ict-standardisation/ict-environmental-impact>

Molla A, Cooper V, Pittayachawan S. 2011. The Green IT Readiness (G-Readiness) of Organizations: An Exploratory Analysis of a Construct and Instrument Communications of the Association for Information Systems 29(1):67–96; DOI: 10.17705/1CAIS.02904

OECD, 2005. Glossary of statistical terms, <https://stats.oecd.org/glossary/detail.asp?ID=6864>

OECD, 2009. Towards Green ICT Strategies: Assessing Policies and Programmes on ICT and the Environment. OECD Conference on “ICTs, the environment and climate change”, Helsingør, Denmark, 27-28 May 2009 <https://www.oecd.org/sti/ieconomy/towardsgreenictstrategies.htm>.

YPEKA, 2012. Climate change. Ministry of Environment and climate change. www.ypeka.gr



Project website: www.digit-biotech.eu

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

DIGIT~Bio~TECH



LO5 OPEN ACCESS SCIENTIFIC RESOURCES: DIGITAL DATABASES

Advanced level

AUTHOR:

FABIANO CHALHOUB & ZOI GEORGIU

Contents

Advanced structure of a Database	4
Advanced structure of a Database	4
Database Management Systems	4
Database models' languages	4
Database characteristics	5
Database model types.....	5
Phases of database design	6
Degree of abstraction	6
Database schemas.....	7
External level	7
Conceptual level.....	7
Internal level.....	7
Data independence	7
The Relational Data Model	8
Basic Elements of a Relational Data Model.....	9
Characteristics of a Table.....	9
Differentiating features of the Relational Database Model.....	10
The Entity-Relationship Model.....	10
Examples of entity types and relationships in biological databases.....	11
Modification Anomalies.....	12
Key Definitions	12
Structured Query Language (SQL)	12
SQL Commands List.....	13
Commercial and Free Databases used in the real world	18
Commercial Databases.....	18
SAP HANA	18
IBM Db2 Database.....	19
Oracle Database	20
Free Databases	21
MySQL.....	21



2019-1-BG01-KA203-062371

PostgreSQL	21
Microsoft SQL	22
MariaDB.....	22
Oracle	23
Firebirdsql	23
Databases in the scientific world advanced module.....	24
Overview of databases in the scientific world.....	24
DNA databases.....	25
RNA databases	26
Protein databases.....	26
Disease databases	26
Expression databases.....	26
Pathway databases.....	26
The Medical Information Mart for Intensive Care (MIMIC) database	29
PCORnet	29
Open NHS	30
Database de-identification.....	30
The application of blockchain in Digital health	30
References	32



Advanced structure of a Database

ADVANCED STRUCTURE OF A DATABASE

This part deals with the advanced design of a Database. It explains the structure of a Database and how to make relations between Database tables. It also presents the specific language used to make queries (SQL) to retrieve data from a Database.

DATABASE MANAGEMENT SYSTEMS

A modern database can be defined as a structured collection of information (data) that is representative of the real world. Database Management Systems (DBMS) are used for the creation, management and query of databases. At present, relational database management systems (RDBMS) are the most mature and widely operated database systems in production. Almost all online transactions and most online content management systems (e.g. blogs and social networks) rely on these types of database systems, which are central to the world's application infrastructure. The focal point of a DBMS is the compilation of services that offer the persistence of data in the database and the functionality to ensure that the data is correct and consistent and that transactions follow the ACID properties. ACID refers to four essential properties of a transaction:

- Atomicity
- Consistency
- Isolation
- Durability

DATABASE MODELS' LANGUAGES

All database models have a language for the specification of the database's structure and content. The specification is known as the schema design and represents the logical view of information that will be managed by a certain DBMS. This database specification language needs to be flexible so as to be useful and lasting. The most visible element of a database, which is identifiable by database professionals and application developers, is the data manipulation language. It can exhibit many forms, with the most common being a programming-language-like interface. Today, the textual and procedural languages, including Structured Query Language (SQL) and Object Query Language (OQL), remain the most widespread forms of data manipulation language.



DATABASE CHARACTERISTICS

A database can be characterised as coherent, logical and internally consistent. It can also be characterised as self-describing, as it includes metadata, which define and describe the data and relationships between tables in the database. It is designed to contain data for a specific purpose. Each data item is stored in a field; a combination of fields is referred to as table. A number of tables may exist in a database.

In contrast to the file-based system, in database systems the data structure is stored in the system catalogue and not in the application programs. This separation between the programs and data is named program-data independence.

The architecture of a database system is composed of a set of services that are constructed on top of basic operating system services, system file storage services and primary memory buffer management services. This set of services is comprised of the following: catalogue management, integrity management, transaction management, concurrency control, lock management, deadlock management, recovery management, security management, query processing, communications management and log management.

DATABASE MODEL TYPES

Data models can be divided into two types:

- High-level conceptual data models
- Record-based logical data models

High-level conceptual data models propose concepts for presentation of data in ways that are similar to how people perceive data. An example of this data model is the entity-relationship (ER) model, which is based on concepts, such as entities, attributes and relationships. An entity corresponds to a real-world object, attributes represent properties of the entity and a relationship indicates an association among entities.

Record-based logical data models propose concepts that users can comprehend, but are similar to the way data is stored in the computer. Relational data models, network data models and hierarchical data models are three of the most prevalent record-based logical data models.

- In the relational model, data are represented in the form of relations, or tables.
- In the network model, data are represented as record types. Also represented by this model is a set type, defined as a limited type of one-to-many relationships.
- In the hierarchical model, data are represented as a hierarchical tree structure, each branch of which is representative of a number of related records.



PHASES OF DATABASE DESIGN

Data modelling constitutes the first step of database design. This step is at times though to be a high-level and abstract design phase, known as the conceptual design. This phase aims to describe the following:

- The data present in the database
- The relationships between data items
- The constraints on data

At this initial phase of the database design process, information-requirement analysis is essential. It is the most important phase because the overall effectiveness of the system relies on how accurately the information requirements and user views are specified in the beginning. The specifications about information requirements made at this stage affect the final form and content of the database system.

After the specifications have been determined and developed, they must be structured into an integrated, cohesive system, a procedure called logical design. Logical design includes the following steps:

- i) developing a data model for each user view
- ii) integrating the entities, attributes and relationships into a composite logical schema that describes the database for that module in terms that are not related to the software package being used
- iii) transforming the logical schema into a software schema expressed in the language of the chosen database management package

The final step of designing a database is physical design. This step is required in order to change the software schema into a form that can be implemented with the specific hardware, operating system and database management system of an organisation. Involved in physical design is the implementation of integrity and security requirements and the design of navigation paths.

DEGREE OF ABSTRACTION

Data abstraction signifies the concealing of certain details of the way data are stored and maintained. In terms of their degree of abstraction, database models can be divided into three levels, which are:

- The external or view level, which is the highest level of abstraction and represents only part of the entire database
- The logical level, which describes what data are stored in the entire database



2019-1-BG01-KA203-062371

- The physical level, which is the lowest level of abstraction and describes how the data are stored in the database

DATABASE SCHEMAS

The database schema can be defined as the early-stage database description that is not expected to frequently change. Numerous schemas exist in a database system. The database architecture consists of three levels of schemas.

External level

This is the highest level of schemas. The external level data view is concentrated on specific data-processing applications or user views. It contains several views and represents a fragment of the actual database. Each view is offered for a user or group of users so that it helps make the interaction between the user and the system simpler.

Conceptual level

This level describes the logical structure of the entire database, which is, in turn, described by simple logical concepts, including objects, their properties or relationships. Therefore, the intricacy of the implementation details of the data will not be visible by the users. Only one conceptual level view is maintained in the database. In order for entities or attributes to be referred to in the database system, they must first be defined in the conceptual level view, formally described as the logical schema. This level view has to be highly stable, since it is considered to be the basis for the development of external and internal level views.

Internal level

The way the data are stored and the way to access the data are described in this schema. The internal level represents the internal or physical state of the database. Its objective is to increase the efficiency of the database system, while fulfilling the required needs.

DATA INDEPENDENCE

Data independence refers to the ability of user applications to remain unaffected by changes made in the definition and organisation of data. Two types of data independence exist: logical and physical.



2019-1-BG01-KA203-062371

Logical data independence is the ability to alter the logical (conceptual) schema without affecting the external schema or user view. Adjustments to the logical schema, such as changes to the database structure like adding tables, should not have an effect on the function of the application (external views).

Physical data independence is the ability of the conceptual level schema to remain unaffected by changes made to the internal schema. Alterations to file organisation or storage structures, storage devices or indexing strategy do not bring about changes in the conceptual level.

THE RELATIONAL DATA MODEL

The relational data model was developed by Dr. Edgar F. Codd in 1970. It represents data in a tabular form, which is a familiar to many people way of representing data. The logical simplicity of flat file structures is maintained in this model. The relational model is based on a set theory, which provides the foundation for several of the operations that are performed on relations. It offers the most flexible access to data and, thus, is useful in dynamic decision-making environments.

SQL is a relational transformation language; it offers ways to form relations and manipulate the data. The outcome of a transformation operation is always another relation, which may contain just one row and column.



Basic Elements of a Relational Data Model

Table 1. Basic components of a relational data model

Database component	Description
Table	includes columns and rows; a subset of the Cartesian product of a list of domains characterised by a name
Columns	main storage units; contain the basic elements of data into which the content can be divided
Rows	contain columns that are associated; together with columns form the basis of all databases
Domain	a set of acceptable values that can be included in a column
Degree	the number of columns present in a table

A relation, which is also called a table or file, can be characterised as a two-dimensional table that consists of data regarding an entity class or the relationships between entity classes. In each row of a table, data referring to a specific entity is included, and, in each column, a specific attribute is included. The rows, or records, of a relation can be referred to as tuples. A record within a table represents an instance of an entity. The number of rows in a relation are indicative of its cardinality. The number of columns, also known as fields or attributes, in a relation corresponds to the degree of the relation. The basic elements of a relational data model are described in Table 1. A unary relation consists of only one attribute; a binary relation consists of only two attributes; a ternary relation consists of only three attributes.

CHARACTERISTICS OF A TABLE

- Each table in a database has a unique name
- No duplicate rows exist; each row is different
- Each row has a different name
- The sequence of rows and columns is not important
- Entries from columns are derived from the same domain according to their data type, including: date, logical (true/false), character (string) and number (numeric, integer, float, ...)



2019-1-BG01-KA203-062371

DIFFERENTIATING FEATURES OF THE RELATIONAL DATABASE MODEL

Essentiality: A data structure is considered essential if it results in a loss of information in the database, when it is removed.

Integrity Rules: These ensure that the database content remains accurate and consistent. There are two types of integrity:

1. **Entity integrity:** Allows the unique identification of each entity in the relational database. This ability ensures access to all data. Requires that no primary key has a null value.
2. **Referential Integrity:** Allows the reference of tuples using foreign keys. Requires that the values assumed by a foreign key either match a primary key that is present in the database or are completely null.

Data manipulation: A method to manipulate the data; principal approach for creating information for decision making.

THE ENTITY-RELATIONSHIP MODEL

The entity-relationship (ER) data model has been available for more than 35 years. It is relatively abstract and easy to explain. ER models are readily translated to relations and represented by ER diagrams. Relationships and entities are the fundamentals of this model. An entity may be an object that exists physically or has conceptual existence. If its tables are existence-dependent, then an entity is characterised as weak. Conversely, if it can exist separately from all of its associated entities, then an entity is referred to as being strong.

Different kinds of entities exist:

- **Independent entities or kernels:** The building blocks of the database. They are strong entities. The primary key is not a foreign key and can be simple or composite. The different types of keys are described in Table 2.
- **Dependent or derived entities:** They are existence-dependent on two or more tables. They are used to bring two kernels together and may include other attributes. Each related table is identified by the foreign key. Three options are available for the primary key: i) use a composite of foreign keys of related tables, if unique, ii) use a composite of foreign keys and a qualifying column, or iii) create a new simple primary key.
- **Characteristic entities:** These entities offer additional information about another table. They describe other entities and are representative of multivalued attributes. The foreign key is used for further identification of the characterised table. Two options are available for the primary key: i) use a composite of foreign keys and a qualifying column, or ii) create a new simple primary key.

**Table 2. Types of keys**

Types of keys	Description
Candidate key	simple or composite key that is unique, because no two rows in a table can have the same value at any time, and minimal, since every column is needed to achieve uniqueness
Composite key	must be minimal; composed of two or more attributes
Primary key	candidate key chosen by the database designer for use as an identifying mechanism for the whole entity set; must uniquely identify tuples in a table and not be null; indicated in the ER model by underlining the attribute
Secondary key	attribute strictly used for retrieval purposes; can be composite
Alternate key	all candidate keys not selected as the primary key
Foreign key	attribute in a table that references the primary key in another table OR it can be null

Null values: Different from zero or blank values; do not depend on data type. A null value means that either the actual value is unknown or that the attribute is not applicable.

EXAMPLES OF ENTITY TYPES AND RELATIONSHIPS IN BIOLOGICAL DATABASES

An entity type describes the characteristics that are shared by a collection of entities in a domain. For example, Protein can be considered as an entity type, with attributes, which include sequence, name, molecular weight, species and accession number. A single entity type will likely have several instances, each of which provides values to the attributes that are specified in the corresponding type. For example, the names of two instances of the entity type Protein are human α -haemoglobin and whale myoglobin. The values of their attribute species would be human and whale, respectively.

Relationships indicate that two or more entity types are associated. For example, a Protein may interact with many other Proteins, or may be a member of a family. Different categories of relationship may describe the nature of the relationship. For example, one entity type could be represented as a part of another (e.g. a Beta strand is part of a Sheet in the secondary structure of a Protein) or as a kind of another (e.g. an Enzyme is a kind of Protein).



MODIFICATION ANOMALIES

Unintentional mistakes may occur in a database during the processes of insertion, deletion or modification of data. If the mistake is a result of the database design, then this is called a modification anomaly.

There are three types of modification anomalies:

1. Deletion anomaly: the removal of one logical entity that leads to loss of information about an unrelated logical entity
2. Insertion anomaly: the insertion of data about one logical entity that necessitates the insertion of data about an unrelated logical entity
3. Update anomaly: the alteration of the information for one logical entity that necessitates more than one alteration to a relation.

KEY DEFINITIONS

Centralised database system: data in this system is stored at a single site

Distributed database system: database and DBMS software are distributed in different sites connected by a computer network.

Database: a shared collection of associated data to be used for supporting the activities of organisations.

Data Definition Language (DDL): used to define the conceptual and internal schemas

Database Management System (DBMS): computer programs used for the creation, management and query of databases

Data Model: a collection of concepts used for the description of the database structure

Data redundancy: storage of the same data piece in two or more places in the database system

Normalisation: a method that structures data in such a way that problems are decreased or avoided

Recovery: the procedure of using logs and backup copies to recreate a database that has been damaged

STRUCTURED QUERY LANGUAGE (SQL)

SQL stands for Structured Query Language, which is a computer language for storing, manipulating and retrieving data stored in a relational database. It is the most widely used database



2019-1-BG01-KA203-062371

language. It offers ways to construct relations and manipulate data. SQL is the standard language for Relational Database Systems. All the Relational Database Management Systems (RDMS), like MySQL, MS Access, Oracle, Sybase, Informix, Postgres and SQL Server, use SQL as their standard database language although, they use different “dialects”:

- MS SQL Server uses T-SQL
- Oracle uses PL/SQL
- MS Access uses a version of SQL called JET SQL (native format) etc.

SQL Commands List

A list of SQL commands that covers all the necessary actions with SQL databases follows. However, as previously mentioned, there might be some differences between different types of databases, including the use of different “dialects”. Each SQL command is provided with its syntax and description.

The commands in SQL are called Queries and they are of two types:

1. Data Definition Query: The statements that define the structure of a database, create tables, specify their keys, indexes and so on,
2. Data manipulation queries: These are the queries that can be edited.



2019-1-BG01-KA203-062371

SQL Commands List¹:

Command	Syntax	Description
ALTER table	ALTER TABLE table_name ADD column_name datatype;	It is used to add columns to a table in a database
AND	SELECT column_name(s)FROM table_nameWHERE column_1 = value_1 AND column_2 = value_2;	It is an operator that is used to combine two conditions
AS	SELECT column_name AS 'Alias'FROM table_name;	It is a keyword in SQL that is used to rename a column or table using an alias name
AVG	SELECT AVG(column_name)FROM table_name;	It is used to aggregate a numeric column and return its average
BETWEEN	SELECT column_name(s)FROM table_nameWHERE column_name BETWEEN value_1 AND value_2;	It is an operator used to filter the result within a certain range
CASE	SELECT column_name,CASEWHEN condition THEN 'Result_1'WHEN condition THEN 'Result_2'ELSE 'Result_3'ENDFROM table_name;	It is a statement used to create different outputs inside a SELECT statement
COUNT	SELECT COUNT(column_name)FROM table_name;	It is a function that takes the name of a column as argument and counts the number of rows when the column is not NULL
Create TABLE	CREATE TABLE table_name (column_1 datatype, column_2 datatype, column_3 datatype);	It is used to create a new table in a database and specify the name of the table and columns inside it
DELETE	DELETE FROM table_nameWHERE some_column = some_value;	It is used to remove the rows from a table
GROUP BY	SELECT column_name, COUNT(*)FROM	It is a clause in SQL used for aggregate functions in collaboration with the SELECT statement

¹ Source <https://intellipaat.com/blog/tutorial/sql-tutorial/sql-commands-cheat-sheet/>

2019-1-BG01-KA203-062371

	table_nameGROUP BY column_name;	
HAVING	SELECT column_name, COUNT(*)FROM table_nameGROUP BY column_nameHAVING COUNT(*) > value;	It is used in SQL because the WHERE keyword cannot be used in aggregating functions
INNER JOIN	SELECT column_name(s)FROM table_1JOIN table_2 ON table_1.column_name = table_2.column_name;	It is used to combine rows from different tables if the JOIN condition goes TRUE
INSERT	INSERT INTO table_name (column_1, column_2, column_3) VALUES (value_1, 'value_2', value_3);	It is used to add new rows to a table
IS NULL/ IS NOT NULL	SELECT column_name(s)FROM table_nameWHERE column_name IS NULL;	It is an operator used with the WHERE clause to check for the empty values
LIKE	SELECT column_name(s)FROM table_nameWHERE column_name LIKE pattern;	It is a special operator used with the WHERE clause to search for a specific pattern in a column
LIMIT	SELECT column_name(s)FROM table_nameLIMIT number;	It is a clause to specify the maximum number of rows the result set must have
MAX	SELECT MAX(column_name)FROM table_name;	It is a function that takes number of columns as an argument and returns the largest value among them
MIN	SELECT MIN(column_name)FROM table_name;	It is a function that takes number of columns as an argument and returns the smallest value among them
OR	SELECT column_nameFROM table_nameWHERE column_name = value_1 OR column_name = value_2;	It is an operator that is used to filter the result set to contain only the rows where either condition is TRUE
ORDER BY	SELECT column_nameFROM table_nameORDER BY column_name ASC DESC;	It is a clause used to sort the result set by a particular column either numerically or alphabetically

2019-1-BG01-KA203-062371

OUTER JOIN	SELECT column_name(s)FROM table_1LEFT JOIN table_2 ON table_1.column_name = table_2.column_name;	It is used to combine rows from different tables even if the condition is NOT TRUE
ROUND	SELECT ROUND(column_name, integer)FROM table_name;	It is a function that takes the column name and an integer as an argument, and rounds the values in a column to the number of decimal places specified by an integer
SELECT	SELECT column_name FROM table_name;	It is a statement that is used to fetch data from a database
SELECT DISTINCT	SELECT DISTINCT column_nameFROM table_name;	It is used to specify that the statement is a query which returns unique values in specified columns
SUM	SELECT SUM(column_name)FROM table_name;	It is function used to return the sum of values from a particular column
UPDATE	UPDATE table_nameSET some_column = some_valueWHERE some_column = some_value;	It is used to edit rows in a table
WHERE	SELECT column_name(s)FROM table_nameWHERE column_name operator value;	It is a clause used to filter the result set to include the rows in which the WHERE condition is TRUE
WITH	WITH temporary_name AS (SELECT *FROM table_name)SELECT *FROM temporary_nameWHERE column_name operator value;	It is used to store the result of a particular query in a temporary table using an alias

2019-1-BG01-KA203-062371

Commands and syntax for querying data from a single table or multiple tables²:

Single Table	Multiple Table
SELECT c1 FROM t To select the data in Column c1 from table t	SELECT c1, c2 FROM t1 INNER JOIN t2 on condition Select column c1 and c2 from table t1 and perform an inner join between t1 and t2
SELECT * FROM t To select all rows and columns from table t	SELECT c1, c2 FROM t1 LEFT JOIN t2 on condition Select column c1 and c2 from table t1 and perform a left join between t1 and t2
SELECT c1 FROM t WHERE c1 = 'test' To select data in column c1 from table t, where c1=test	SELECT c1, c2 FROM t1 RIGHT JOIN t2 on condition Select column c1 and c2 from table t1 and perform a right join between t1 and t2
SELECT c1 FROM t ORDER BY c1 ASC (DESC) To select data in column c1 from table t either in ascending or descending order	SELECT c1, c2 FROM t1 FULL OUTER JOIN t2 on condition Select column c1 and c2 from table t1 and perform a full outer join between t1 and t2
SELECT c1 FROM t ORDER BY c1 LIMIT n OFFSET offset To skip the offset of rows and return the next n rows	SELECT c1, c2 FROM t1 CROSS JOIN t2 Select column c1 and c2 from table t1 and produce a Cartesian product of rows in a table
SELECT c1, aggregate(c2) FROM t GROUP BY c1 To group rows using an aggregate function	SELECT c1, c2 FROM t1, t2 Select column c1 and c2 from table t1 and produce a Cartesian product of rows in a table
SELECT c1, aggregate(c2) FROM t GROUP BY c1 HAVING condition Group rows using an aggregate function and filter these groups using 'HAVING' clause	SELECT c1, c2 FROM t1 A INNER JOIN t2 B on condition Select column c1 and c2 from table t1 and join it to itself using INNER JOIN clause

² Source <https://intellipaat.com/blog/tutorial/sql-tutorial/sql-commands-cheat-sheet/>

Commercial and Free Databases used in the real world

• 4th Dimension • Adabas D • Airtable • Apache Derby • Aster Data • Amazon Aurora • Altibase • CA Datacom • CA IDMS • Clarion • ClickHouse • Clustrix • CockroachDB • CSQL • CUBRID • DataEase • DataFlex • Database Management Library • Dataphor • dBase • Derby (aka Java DB) • Empress Embedded Database • Exasol • EnterpriseDB • eXtremeDB • FileMaker Pro • Firebird • FrontBase	• Google Fusion Tables • Greenplum • GroveSite • H2 • Helix • HSQLDB • IBM Business System 12 (historical) • IBM DB2 • IBM Lotus Approach • IBM DB2 Express-C • Infobright • Informix • Ingres • InterBase • InterSystems Caché • InterSystems IRIS Data Platform • LibreOffice Base • Linter • MariaDB • MaxDB • Microsoft Access • Microsoft Jet Database Engine (part of Microsoft Access) • Microsoft SQL Server • Microsoft SQL Server Express • SQL Azure (Cloud SQL Server) • Microsoft Visual FoxPro • Mimer SQL	• MonetDB • mSQL • MySQL • Netezza • NexusDB • NonStop SQL • NuoDB • Omnis Studio • OpenLink Virtuoso (Open Source Edition) • OpenLink Virtuoso Universal Server • OpenOffice.org Base • Oracle • Oracle Rdb for OpenVMS • Panorama • Paradox • Percona Server for MySQL • Percona XtraDB Cluster • Pervasive PSQL • Polyhedra • PostgreSQL • Postgres Plus Advanced Server • Progress Software • Raima Database Manager (RDM) • RDM Server • R:Base • RethinkDB	• SAND CDBMS • SAP HANA • SAP Adaptive Server Enterprise • SAP IQ (formerly known as Sybase IQ) • SingleStore • Snowflake Cloud Data Warehouse • solidDB • SQL Anywhere (formerly known as Sybase Adaptive Server Anywhere and Watcom SQL) • SQLBase • SQLite • SQream DB • SAP Advantage Database Server (formerly known as Sybase Advantage Database Server) • Teradata • Tiberio • TiDB • TimesTen • Trafodion • Unisys RDMS 2200 • UniData • UniVerse • Vectorwise • Vertica • VoltDB • YugabyteDB
--	--	---	---

Figure 1. Non-exhaustive list of available databases

This part deals with the common databases found on the market, whether they are free or proprietary. However, there are so many databases available (figure 1) that we cannot mention all of them. A choice had to be made and the ones presented below are the “most popular” or the “most frequently used”.

COMMERCIAL DATABASES

From the vast number of databases available on the market, we chose to present three commercial databases commonly used by the major companies and organisations.

SAP HANA

This database is designed by the European company SAP SE, founded in Germany. SAP HANA is a database engine that is column-oriented and can





2019-1-BG01-KA203-062371

handle SAP and non-SAP data. The engine is designed to save and retrieve data from applications and other sources across multiple tiers of storage. SAP HANA can be deployed on-premises or in the cloud from a number of cloud service providers. This database is usually chosen by organizations that are pulling data from applications and are not under a terribly constrained budget.

Its main features are:

- It supports SQL, OLTP and OLAP.
- The engine reduces resource requirements through compression.
- Data is stored in memory, reducing access times, in some cases, significantly.
- Real-time reporting and inventory management are available.
- It can interface with a number of other applications.

As of January 2021, the currently supported hardware platforms³ for SAP HANA are:

- Intel-based hardware platforms
- IBM Power Systems

As of January 2021, the currently supported operating systems⁴ for SAP HANA are:

- Linux SUSE
- Linux Red Hat

IBM Db2 Database

IBM Db2 database traces its roots back to the beginning of the 1970s when Edgar F. Codd, a researcher working for the company, described the theory of relational databases, and in June 1970 published the model for data manipulation. Today, it is a database engine that has NoSQL capabilities, and it can read JSON⁵ and XML files⁶.



The current version of DB2 is LUW 11.1, which offers a variety of improvements. One, in particular, was an improvement of BLU Acceleration (BLink Ultra or Big Data, Lightning fast and Ultra-easy), which is designed to make this database engine work faster through data skipping technology.

³ Source SAP SE <https://help.sap.com/viewer/eb3777d5495d46c5b2fa773206bbfb46/2.0.01/en-US/d3d1cf20bb5710149b57fd794c827a4e.html>

⁴ For more information about supported operating systems for SAP HANA, see SAP Note 2235581 - SAP HANA: <https://service.sap.com/sap/support/notes/2235581>

⁵ JavaScript Object Notation is an open standard file format as XML and is considered as unstructured data.

⁶ XML is an open standard file format as JSON and is considered as unstructured data.



2019-1-BG01-KA203-062371

Data skipping is designed to improve the speed of systems with more data than can fit into memory. The latest version of Db2 also provides improved disaster recovery functions, compatibility and analytics.

Its main features are:

- BLU Acceleration can make the most of available resources for enormous databases.
- It can be hosted from the cloud, a physical server or both at the same time.
- Multiple jobs can be run at once using the Task Scheduler.
- Error codes and exit codes can determine which jobs are run via the Task Scheduler.

The currently supported hardware platforms⁷ as of January 2021 for IBM Db2 are:

- IBM z/Architecture mainframe
- Intel-based hardware platforms

The currently supported operating systems as of January 2021 for IBM Db2 are:

- z/OS
- Unix
- Linux
- Windows

Oracle Database

Oracle Database is commonly used for running online transaction processing (OLTP) or data warehousing (DW). It can also mix OLTP and DW database workloads. **ORACLE®** Oracle Database is available on-premises, on-cloud or as hybrid cloud installation. It may be run on third-party servers, as well as on Oracle Exadata hardware on-premises, on Oracle Cloud or on a private Cloud at customer premises.

The first version was released in 1979 and its development was influenced by the research of Edgar F. Codd on relational database design.

Its main features are:

- It is a cross-platform database. It can run on various hardware across operating systems, including Windows Server, Unix and various distributions of GNU/Linux.
- It has its networking stack that allows applications from a different platform to communicate smoothly with the Oracle Database, e.g. applications running on Windows can connect to the Oracle Database running on Unix.
- It is an ACID-compliant database that helps maintain data integrity and reliability.

⁷ Source IBM Support <https://www.ibm.com/support/pages/system-requirements-ibm-db2-linux-unix-and-windows#1155S>



2019-1-BG01-KA203-062371

The currently supported hardware platforms are:

- Proprietary Oracle Database Appliance
- Sparc
- IBM Power Systems
- X64-based hardware platforms

The currently supported operating systems⁸ are:

- Unix
- Linux
- Windows

FREE DATABASES⁹

If a database is free, this does not necessarily mean that no fees are charged to the user. It is true for some of the following databases, however, some developers choose to limit certain features and charge a fee to be able to unlock those features (refer to the first unit of the Basic Level).

MySQL

MySQL is an open-source relational database, which runs on a number of different platforms, including Windows, Linux, macOS, etc. A cloud version. MySQL can be used for packaged software, business-critical systems and high-volume websites.



Its main features are:

- It provides scalability and flexibility
- The tool has web and data warehouse strengths
- It provides high performance
- It has robust transactional support

PostgreSQL

PostgreSQL is an enterprise-class open source database management system. It supports both SQL for relational and JSON for non-relational queries. It is backed by an experienced community of developers who have made a tremendous



⁸ Source https://support.oracle.com/knowledge/Oracle%20Database%20Products/1369107_1.html

⁹ Source <https://www.guru99.com/free-database-software.html> updated on 2021



2019-1-BG01-KA203-062371

contribution to make it a highly reliable database management software. It runs on three different platforms, namely Windows, Linux and macOS. A cloud version is not available. PostgreSQL enables the creation of custom data types and a range of query methods. A stored procedure can be run in different programming languages.

Its main features are:

- It is compatible with various platforms using all major languages and middleware
- Standby server and high availability
- The tool has mature server-side programming functionality
- Log-based and trigger-based replication SSL
- It offers a most sophisticated locking mechanism
- Support for multi-version concurrency control
- It provides support for client-server network architecture
- The tool is Object-oriented and ANSI-SQL2008 compatible
- PostgreSQL allows linking with other data stores like NoSQL, which act as a federated hub for polyglot databases.

Microsoft SQL

SQL Server is an RDBMS developed by Microsoft. It supports ANSI SQL, which is the standard SQL (Structured Query Language) language. However, SQL Server comes with its implementation of the SQL language, T-SQL (Transact-SQL). It runs on Docker Engine, Ubuntu, SUSE Linux Enterprise Server and Red Hat Enterprise Linux. A cloud version is available.



Its main features are:

- It provides integration of structured and unstructured data with the power of SQL Server and Spark.
- The tool offers scalability, performance and availability for mission-critical, intelligent applications, data warehouses and data lakes.
- It offers advanced security features to protect your data.
- Access to rich, interactive Power BI reports, to make a faster and better decision.

MariaDB

MariaDB is a fork of the MySQL database management system. It was created by its original developers. This DBMS tool provides data processing capabilities for both small and enterprise tasks. It runs on three platforms, namely Windows, Linux





2019-1-BG01-KA203-062371

and macOS. A cloud version is available. MariaDB is an alternate software to MySQL. It provides high scalability through easy integration.

Its main features are:

- It operates under GPL, BSD or LGPL licenses.
- It comes with many storage engines, including the high-performance ones that can be integrated with other relational database management systems.
- It provides the Galera cluster technology.
- MariaDB can run on different operating systems and it supports numerous programming languages.

Oracle

Oracle is a self-repairing, self-securing and self-driving database designed to eliminate manual data management. It is an intelligent, secure and highly available database in the cloud that helps businesses grow. It runs on two platforms, namely Windows and Linux. A cloud version is also available.



Its main features are:

- Oracle Cloud is optimized for high-performance database workloads, streaming workloads and hyperscale big data.
- You can easily migrate to the Cloud.
- It provides the services based on how you like to operate, in order to run Oracle Cloud in your data center.

Firebirdsql

Firebird is an open source SQL RDBMS that runs on Microsoft Windows, macOS, Linux and several Unix platforms, including HP-UX, Solaris and AIX. A cloud version is available. Firebird has development-friendly language support, stored procedures and triggers.



Its main features are:

- Firebird allows you to build a custom version.
- It is free to download, register and deploy.
- The tool has enhanced multi-platform RDBMS.
- Provides a range of funding options from firebird memberships to sponsorship commitments.



Databases in the scientific world advanced module

This section is dedicated to further exploring the open access databases used in science and how to use and get benefit of the existing knowledge.

OVERVIEW OF DATABASES IN THE SCIENTIFIC WORLD

Existing databases dedicated to science and how to use them

As previously mentioned, sharing, integrating and annotating data is a crucial part of biological research as it allows researchers to reproduce the examination and interpretation of experimental findings. Although it is thought that bioinformaticians and computer scientists are responsible for these actions, life scientists have an equal role in promoting data integration, since they are the ones that generate these types of data and are usually the end users.

Data integration is defined as the process of combining data from different sources in order to offer users a unified view of such data. In computational sciences the theoretical frameworks for data integration have been categorised, based on the method used to integrate the data, into “eager” and “lazy”. According to the eager method, which is also known as warehousing, the data are copied over to a global schema and stored in a central data warehouse. The term “schema” refers to an organised and “queryable” approach for storing data. In the lazy method, the data are located in distributed sources and are integrated on demand in accordance with a global schema used for mapping the data between sources. The volume of data, the owner of the data and the existing infrastructure are the main factors that ultimately determine which of the two methods will be used for data integration. Moreover, in biological sciences, these methods may be applied in various different ways and used at a range of levels. As a result, six distinct and widely used schemata have been formulated for the integration of data:

- Data centralisation: The data reside in centralised resources. UniProt and GenBank are two examples of databases following this method.
- Data warehousing: data from a variety of sources reside on one central repository. Pathway commons is a database that follows this approach to integrate data.
- Dataset integration: In-house workflows access databases that are distributed and download data to a local repository.
- Hyperlinks: This approach enables users to access databases and tools in different fields of life science, thus promoting interoperability. ExPASy is an indicative example of a portal that is based on this data integration methodology.
- Federated databases: A translational layer is required in order for data to be integrated among heterogeneous databases. This means that data from the database are transformed to a commonly accepted format in such a way that they are able to be interpreted in the



2019-1-BG01-KA203-062371

same way from a mapping service. The Distributed Annotation System (DAS), which is a client-server system, is an indicative example.

- **Linked data:** A network of interconnected data accessible online. Graphical user interfaces (GUI) that consist of hyperlinks, which connect associated data from numerous data providers and, hence, form a large system of Linked Data. BIO2RDF is an indicative example of a database that uses this approach as the basis for data integration.

Data centralisation, data warehousing and dataset integration are based on the “eager” theoretical framework, whereas hyperlinks, federated databases and linked data are based on the “lazy” theoretical framework regarding the way that is selected for data integration.

Data formats are described as an organised way for the demonstration of data and metadata in a file. Scientists began to store biological data in formatted files because the exponential growth of data created the need to analyse them using computer systems and databases. A problem that has arisen in relation to file formatting is the emergence of various formats, even for the representation of the same type of data. In some cases, it has been observed that more than one format classes may be used to represent the data and metadata in a single file. Moreover, research has demonstrated that the most widely used format classes are: i) tables, ii) FASTA-like, iii) tag-structured, and iv) GenBank-like. The ideal solution to this matter would be for scientists to agree upon using a limited number of specific formats so as to simplify the process of data integration. The design of converters that have the ability to translate all the different classes of formats would also provide a helpful solution.

Currently, over 1,700 databases including data of biological interest are in use, according to the non-exhaustive list curated by the Nucleic Acids Research journal. In order to be deemed valuable for a specific purpose, all datasets that are present in a database have to be integrated and structured. The existing biological databases are comprised of information about a vast range of biology research topics, such as non-vertebrate genomics, protein sequence, human genes and diseases, DNA nucleotide sequence, cell biology, immunology, metabolic and signalling pathways, proteomics, etc.

As previously mentioned in the Basic Level, the classification of biological databases is dependent upon several factors, including the scope of data coverage and the level of biocuration. Nevertheless, their classification according to the type of data is one of the simplest and most comprehensive ways of categorising biological databases. Therefore, in the following section, these will be described as DNA, RNA, protein, disease, expression and pathway databases.

DNA databases

DNA databases focus on handling DNA data from numerous or a few particular species. The main purpose of human DNA databases is to establish the reference genome, to conduct profiling of human genetic variation, to associate genotype with phenotype and to identify human microbiome metagenomes. A DNA database example is GenBank, a publicly available collection of all the studied DNA sequences. As of February 2021, over 776 billion nucleotide bases in over 226 million sequences are available in GenBank (<http://www.ncbi.nlm.nih.gov/genbank/statistics>).



2019-1-BG01-KA203-062371

RNA databases

These databases include information about non-coding RNAs (ncRNAs), such as microRNAs and long non-coding RNAs (lncRNAs), which do not encode proteins. The purpose of RNA databases is to decode ncRNAs, of which lncRNAs are the most commonly studied, and to describe their functions and interactions. An RNA database example is RNAcentral, which consists of a unified view of ncRNA sequence data derived from a number of databases, some of which are Rfam, miRBase and lncRNAdb.

Protein databases

Protein databases were developed in order to create a vast compilation of universal proteins, identify protein families and domains, reconstruct phylogenetic trees and conduct profiling of protein structures. PDB, which consists of thousands of structures of biological macromolecules, is an indicative example of protein databases.

Disease databases

By definition disease databases include information about different types of diseases, but are mostly focused on providing data concerning various types of cancer. One of the most important cancer projects that has been developed is The Cancer Genome Atlas (TCGA), the objective of which is to gather a broad range of omics data, such as mRNA, SNP and methylation, for over twenty different forms of human cancer.

Expression databases

Expression databases may be utilised for a number of tasks, such as studying tissue-specific gene expression and regulation, storing expression data, detecting differential and baseline expression, and examining and reviewing expression information obtained from RNA and protein data. As an expression database, the Human Protein Atlas incorporates expression profiles for a significant percentage of human protein-coding genes derived from RNA and protein data.

Pathway databases

Pathway databases include data about biological pathways that can be utilised by researchers for the analysis of metabolic, regulatory and signalling pathways. A characteristic example of pathway databases is KEGG PATHWAY, which contains information regarding molecular interaction and reaction networks.

The National Center for Biotechnology Information (NCBI), part of the US National Library of Medicine at the National Institute of Health, has developed an integrated database retrieval system, which offers access to 34 different databases collectively containing 3.0 billion records, named Entrez. The global search page of Entrez (<https://www.ncbi.nlm.nih.gov/search/>) provides links to the web portal



2019-1-BG01-KA203-062371

for each of the 34 databases. The Entrez system is easy to use because it allows users to download data in a variety of formats and to perform text searching using simple Boolean queries. Records are linked between databases on the basis of asserted relationships; these records can be represented in various formats. Moreover, users of Entrez have the option to download single records or batches of records. Some of the 34 databases that are part of Entrez are the following: PubMed (<https://pubmed.ncbi.nlm.nih.gov>), which contains scientific and medical abstracts/citations; BioSample (<https://www.ncbi.nlm.nih.gov/biosample>), which comprises descriptions of biological source materials; GEO Profiles (<https://www.ncbi.nlm.nih.gov/geoprofiles>), which includes gene expression and molecular abundance profiles; and, dbVar (<https://www.ncbi.nlm.nih.gov/dbvar>), which contains data from genome structural variation studies.

The data submitted to NCBI are derived from three sources: i) directly from researchers, ii) national and international partnerships or agreements with data providers and research consortia, and iii) internal curation efforts. Of note, NCBI is responsible for the management of the GenBank database and partakes in the International Nucleotide Sequence Database Collaboration (INSDC) in cooperation with the EMBL-EBI European Nucleotide Archive (ENA) and the DNA Data Bank of Japan (DDBJ).

As databases have proven to be a useful tool in many scientific fields, their use is steadily gaining ground in the healthcare sector. Nowadays, technological advancements in the area of data science have enabled healthcare professionals to compile, process and analyse health-related data, leading to the improvement of not only the delivery of care, but also the safety of patients and consumers. In order for these improvements to take place, relevant data must be gathered, stored and analysed in an efficient and secure manner, and exchanged across the different service levels present in a healthcare system. This has led to the development of Electronic Health Records (EHRs), databases which store patient data that can be accessed and utilised by healthcare professionals.

EHRs can be defined as medical databases that offer users, which in this case are healthcare professionals and administrative staff, access to health records. The most distinct types of EHRs are the Electronic Medical Record (EMR) and the Personal Health Record (PHR). EMRs consist of information, which is submitted by a single hospital department, an entire hospital or parts of the hospital. They can also contain information from a number of hospitals. Information to this type of EHR is normally added only by hospital staff. On the contrary, PHRs are managed by the patients, which are able to enter information. PHRs are described as electronic applications that provide a secure platform for patients to control and share their health data. The main difference between the two types of EHR systems is that, in PHRs, health records have to be presented in such a way that is understandable by the patient, whereas, in EMRs, the way health records are presented resembles health records on papers, since they are only accessed by healthcare providers.

The first EHR system became available in the 1960s mainly due to the build-up of unstructured and unused patient information over a period of several decades. Large organisations started to set up database systems in order to store and structure data into central repositories. These databases allowed the organisation and collection of data from many different sources, including pharmacies, laboratories, clinical studies, and constituents of clinical care, such as medication administration records. Currently,



2019-1-BG01-KA203-062371

the implementation of EHR systems is mostly observed in high income countries. For example, the Health Information Technology for Economic and Clinical Health Act (HITECH Act of 2009) prompted the digitization of the healthcare delivery system in the US and the subsequent development of the Medicare and Medicaid EHR Incentive Programs.

The primary purpose for the creation of EHRs was the need to archive and structure patients' records. They were later designated for billing and quality improvement reasons. As technological advancements took place, over the years EHRs became more inclusive, dynamic and interconnected. Nonetheless, compared to other industries, big data have not been used to best advantage in the medical industry. This has happened predominantly due to the poor quality of collected data and poorly structured datasets. Prior to the development of EHRs, medical research was based on disease registries or chronic disease management systems (CDMS). These repositories have significant limitations, since they consist of collections of data that are often related to only one particular disease. Furthermore, they cannot perform translation of the data or conclusions to other diseases and may include information from a group of patients in a specific geographic area. On the other hand, EHR data is largely varied, thus facilitating the analysis of complex clinical interactions and decisions.

The components of EHRs are different types of medical data, ranging from health records to raw sensory data. Medical data can be categorised into sensitive data or non-sensitive data. Sensitive data include patient information or can be associated with a patient. Non-sensitive data include sensory data, which are also called measurement data due to the fact that they are only comprised of samples of sensors, such as samples of an EEG measurement. Data stored in a medical database are referred to as metadata. The most common type of database used for storing medical data is the relational database, which presents data in the form of tables comprised of rows and a set number of columns. Some databases may include patient information, such as the medical history of a patient, or anonymised data that can be utilised in studies.

Medical data can be divided into several categories as described below:

- Medical and laboratory data: Healthcare workers can submit orders for medication or laboratory studies into a physician order entry system, which are subsequently carried out by laboratory or nursing staff. Examples of this category of data are prescriptions for medication and microbiology results.
- Billing data: This category of medical data is comprised of codes used by hospitals to file claims with their insurance providers. The International Classification of Diseases, constructed by the WHO, and the Current Procedural Terminology, sustained by the American Medical Association, are the most popular coding systems.
- Images: These may be radiographic images resulting from x-rays, echocardiograms and computed tomography (CT) scans.
- Notes and reports: These may be associated with the progress of patients. Discharge summaries also belong to this category. Findings from imaging studies are usually described in operative reports. Notes have to be partially structured using a templating system.



2019-1-BG01-KA203-062371

- Physiological data: This category of medical data contains vital signs, such as heart rate and blood pressure, as well as ECG and EEG waveforms.

Relational databases are most frequently used for medical data management and storage. They can be referred to as a collection of tables that are connected by shared keys. A database schema determines how the tables will be structured and their relationships. A simple medical database may contain four tables:

- Table 1: a patient list
- Table 2: a hospital admissions log
- Table 3: a list with vital sign measurements
- Table 4: a dictionary of vital sign codes and associated labels

Primary and foreign keys can be used in order to link the four tables.

The preponderance of healthcare databases provides limited access to data for various reasons, including privacy concerns and plans to monetise the data. Nonetheless, a number of open access health databases are available for public use, some of which are described below.

The Medical Information Mart for Intensive Care (MIMIC) database

The MIMIC database (<http://mimic.physionet.org>) was created in 2003 as the result of a collaboration between MIT, Philips Medical Systems and the Beth Israel Deaconess Medical Center (BIDMC). The data entered into this database was sourced from medical and surgical patients admitted to all Intensive Care Units at BIDMC. It consists of information from over forty thousand patients, detailed physiological and clinical data, and is de-identified and openly accessible to researchers. Two types of data are present in this database: clinical data derived from EHRs, which are stored in a relational database comprised of approximately 50 tables, and bedside monitor waveforms stored in flat binary files. The goal of this collaboration is to produce and assess advanced ICU patient monitoring and decision support systems that will ultimately make the process of decision-making in critical care more efficient, quicker and accurate.

PCORnet

PCORnet, the National Patient-Centered Clinical Research Network, is an initiative that began in 2013 with the objective to integrate data from several Clinical Data Research Networks and Patient-Powered Research Networks. It contains 29 networks that will facilitate access to vast amounts of research. It collects data from routine patient visits and data that are shared by individual patients via personal health records or community networks with other patients.



2019-1-BG01-KA203-062371

Open NHS

The National Health Services (NHS England) maintains one of the largest repositories in the world containing data related to peoples' health. Open NHS¹⁰ is an open source database that provides access to information made available to the public by the government or other public bodies. This project was established in order to increase transparency and monitor the efficiency of the British healthcare sector. Patients, healthcare workers and commissioners are given the opportunity to compare the quality of care in various locations of the country by simply accessing the available data in the specially designed database.

Database de-identification

One of the primary steps to building an EHR database is de-identification. Before a database becomes available for use by researchers and applications, it is essential that measures are taken in order to ensure that privacy policies and regulations are followed. For structured data, such as columns of a table, de-identification is based on the categorisation of data and the subsequent deletion or cryptography of those that are flagged as protected. For unstructured data, such as discharge summaries, different techniques of natural language processing are used, from simple regular expressions to complex neural networks, which try to find all information that is protected throughout free text in order to perform deletion or cryptography.

THE APPLICATION OF BLOCKCHAIN IN DIGITAL HEALTH

Blockchain technology is based on the concept of having a decentralised system for data storage, where a copy of the ledger of the performed transactions will be provided to each participant/node. This will make it unfeasible for someone to modify the data without the other participants being informed. Strong centralised entities would benefit from the application of blockchain. The applications of Digital health greatly depend on centralised systems. Therefore, blockchain has the potential to transform Digital health by altering the way data are stored and secured. Various areas have been proposed for its application, including supply chains, drug verification, claims reimbursement, access control and clinical trials.

Medical data has been found to be the most highly valued data by hackers, as recent studies have estimated that a single health record may cost up to 400 USD. This means that keeping the data secure in medical databases is of utmost importance. Blockchain can provide a solution to this issue by ensuring data privacy, integrity, authentication and authorisation. Blockchain data are encrypted and if someone has to delete or make their data useless, they are given that ability by applying a key destruction

¹⁰ Open data at the NHS. Available from: <http://www.england.nhs.uk/ourwork/tsd/data-info/open-data/>



2019-1-BG01-KA203-062371

mechanism, where the key that was originally used for the encryption of the message will be destroyed, or made useless. Afterwards, the data stored in the blockchain will not be accessible to read.

Blockchain is able to fulfil two essential needs regarding data sharing: integrity and non-repudiation. Integrity means that the query and retrieved data cannot be altered, once the retrieval operation has been performed. Non-repudiation means that the knowledge retrieval service does not possess the ability to deny that the specific data have been delivered by the service as a response to a given query at a particular time. Blockchain can be defined as a distributed transaction management system that cannot be corrupted. It can be applied for EHR integration, sharing and access control, preservation and management.

A theoretical blockchain-based query notary service may be comprised of three computational layers:

- a data consumer front-end
- an interface for communicating with biomedical database interfaces, and
- the contract engine, which organises the query and returns the retrieved results to the consumer, performs and prepares transactions, and manages contracts and their metadata

Two different schemes may be employed to apply the notary service: the basic scheme and the versioning scheme. The basic scheme applies a query-response ledger by which the user receives a sealed proof verifying that at a specific time a particular query has been placed in a biomedical database, which returned specific results. This scheme may be employed to ensure the integrity and non-repudiation of a query, when a vital biomedical task relies on the specific query. The versioning scheme permits the non-reputable versioning of data retrieved from a dynamically evolving biomedical database at numerous occasions in time, always using the same query. This scheme may be applied to confirm different versions of changing medical evidence as retrieved from a biomedical database with content that is frequently updated.

The incorporation of blockchain technology in pharmaceutical or life science applications has the capacity to decentralise the interface and data sharing, leading to increased efficiency, higher speeds and unlimited scalability. Blockchain renders data immutable, which would be useful in clinical trials for ensuring that clinical data cannot be manipulated by researchers at a later time. It can also be utilised in the process of drug identification, tracing and verification. There are certain risks associated with the implementation of blockchain, such as privacy concerns, off-chain transactions and doubts about this technology due to lack of adoption. Nonetheless, the benefits of blockchain technology far outweigh possible drawbacks and could have a significant role in limiting the methods used for illegal activities.



References

- Agha-Mir-Salim L, Sarmiento RF. 2020. Health information technology as premise for data science in global health: A discussion of opportunities and challenges. In: *Leveraging Data Science for Global Health*. Cham: Springer International Publishing, 3–15.
- Amid C, Alako BTF, Balavenkataraman Kadhivelu V, Burdett T, Burgin J, Fan J, Harrison PW, Holt S, Hussein A, Ivanov E et al. 2020. The European nucleotide archive in 2019. *Nucleic Acids Res.*, 48:D70–76.
- Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. 2004. Uniprot: the universal protein knowledgebase. *Nucleic Acids Res.*, 32 (Suppl 1):115–9. doi: 10.1093/nar/gkh131.
- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.*, 40(Web Server issue):597–603. doi: 10.1093/nar/gks400.
- Belleau F, Nolin MA, Tourigny N, Rigault P, Morissette J. 2008. Bio2RDF: towards a mashup to build bioinformatics knowledge systems. *J Biomed Inform.*, 41(5):706–16.
- Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2014. GenBank. *Nucleic Acids Res.*, 42:D32–D37.
- Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2014. GenBank. *Nucleic Acids Res.*, 42:D32–D37.
- Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2014. GenBank. *Nucleic Acids Res.*, 42:D32–D37.
- Bornberg-Bauer E, Paton NW. 2002. Conceptual data modelling for bioinformatics. *Brief Bioinform.*, 3(2):166–80.
- Bulgarelli L, Núñez-Reiz A, Deliberato RO. 2020. Building electronic health record databases for research. In: *Leveraging Data Science for Global Health*. Cham: Springer International Publishing, 55–64.
- Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, et al. 2013. Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.*, 41: D226-232.
- Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. 2013. The Cancer Genome Atlas Pan-Cancer analysis project, *Nat Genet.*, 45: 1113-1120.
- Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur O, Anwar N, et al. 2011. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.*, 39(Database issue): 685–90.



2019-1-BG01-KA203-062371

Chavali LN, Prashanti NL, Sujatha K, Rajasheker G, Kavi Kishor PB. 2018. The Emergence of Blockchain Technology and its Impact in Biotechnology, Pharmacy and Life Sciences. *Current Trends in Biotechnology and Pharmacy.*, 12(3):304–10.

Courtney JF, Paradise DB, Brewer KL, Graham JC. 2010. *Database Systems for Management*. 3rd edition. The Global Text Project.

Dowell RD, Jokerst RM, Day A, Eddy SR, Stein L. 2001. The distributed annotation system. *BMC Bioinformatics.*, 2:7.

Edgar F. Codd https://en.wikipedia.org/wiki/Edgar_F._Codd

Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. 2014. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc JAMIA.*, 21(4):578–582.

Fortier PJ, Michel HE. 2003. *Computer Data Processing Hardware Architecture*. In: *Computer Systems Performance Evaluation and Prediction*. Elsevier, p. 39–106.

Hellerstein JM, Stonebraker M, Hamilton J. 2007. Architecture of a database system. *Found Tren Databases.*, 1(2):141–259.

Johnson A, Pollard T, Shen L et al. 2016. MIMIC-III, a freely accessible critical care database. *Sci Data* 3., 160035.

Karsch-Mizrachi I, Takagi T, Cochrane G. 2018. International Nucleotide Sequence Database, C The international nucleotide sequence database collaboration. *Nucleic Acids Res.*, 46:D48–51.

Kleinaki A-S, Mytis-Gkometh P, Drosatos G, Efraimidis PS, Kaldoudi E. 2018. A blockchain-based notarization service for biomedical knowledge retrieval. *Comput Struct Biotechnol J.*, 16:288–97.

Kozomara A, Griffiths-Jones S. 2014. MiRBase: annotating high confidence microRNAs using deep sequencing data, *Nucleic Acids Res.*, 42: D68-73.

Lapatas V, Stefanidakis M, Jimenez RC, Via A, Schneider MV. 2015. Data integration in biological research: an overview. *J Biol Res (Thessalon).*, 22(1):9.

Lastdrager E. 2011. *Securing Patient Information in Medical Databases* [Internet]. University of Twente;. Available from: https://essay.utwente.nl/61035/1/MSc_E_Lastdrager_DIES_CTIT.pdf

Marshall J, Chahin A, Rush B. 2016. Review of clinical databases. In: *Secondary Analysis of Electronic Health Records*. Cham: Springer International Publishing;, 9–16.

Nguyen KA. *Database System Concepts*. OpenStax CNX; 2009 [cited 2021 Jan 29]. Available from: <http://cnx.org/contents/b57b8760-6898-469d-a0f7-06e0537f6817@1>

Ogasawara O, Kodama Y, Mashima J, Kosuge T, Fujisawa T. 2020. DDBJ database updates and computational infrastructure enhancement. *Nucleic Acids Res.*, 48:D45–50.

Okuda S, Yamada T, Hamajima M, Itoh M, Katayama T, Bork P, et al. 2008. KEGG Atlas mapping for global analysis of metabolic pathways, *Nucleic Acids Res.*, 36: W423-426.



2019-1-BG01-KA203-062371

- Oliveira AL. 2019. Biotechnology, big data and artificial intelligence. *Biotechnol J.*, 14(8):e1800613.
- Pollard T, Dernoncourt F, Finlayson S, Velasquez A. 2016. Data Preparation. In: *Secondary Analysis of Electronic Health Records*. Cham: Springer International Publishing;, 101–14.
- Ponten F, Schwenk JM, Asplund A, Edqvist PH. 2011. The Human Protein Atlas as a proteomic resource for biomarker discovery, *J Intern Med.*, 270: 428-446.
- Quek XC, Thomson DW, Maag JL, Bartonicek N, Signal B, Clark MB, et al. 2015. IncRNADB v2.0: expanding the reference database for functional long noncoding RNAs, *Nucleic Acids Res.*, 43, D168-173.
- Rose PW, Beran B, Bi C, Bluhm WF, Dimitropoulos D, Goodsell DS, et al. 2011. The RCSB Protein Data Bank: redesigned web site and web services, *Nucleic Acids Res.*, 39: D392-401.
- Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, et al. 2021. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, 49(D1):D10–7.
- Schuler G.D., Epstein J.A., Ohkawa H., Kans J.A. 1996. Entrez: molecular biology database and retrieval system. *Methods Enzymol.*, 266:141–162.
- The RNACentral Consortium, RNACentral: an international database of ncRNA sequences. 2015. *Nucleic Acids Res.*, 43: D123-129.
- Watt A, Eng N. Types of Data Models. In: Watt A, Eng N, editors. *Database Design - 2nd edition*. BCcampus; 2014 [cited 2021 Jan 29]. Available from: <https://opentextbc.ca/dbdesign01>
- Watt A. Characteristics and Benefits of a Database. In: Watt A, Eng N, editors. *Database Design - 2nd edition*. BCcampus; 2014 [cited 2021 Jan 29]. Available from: <https://opentextbc.ca/dbdesign01/>
- Watt A. Data Modelling. In: Watt A, Eng N, editors. *Database Design - 2nd edition*. BCcampus; 2014 [cited 2021 Jan 29]. Available from: <https://opentextbc.ca/dbdesign01>
- Watt A. The Entity Relationship Data Model. In: Watt A, Eng N, editors. *Database Design - 2nd edition*. BCcampus; 2014 [cited 2021 Jan 29]. Available from: <https://opentextbc.ca/dbdesign01>
- Watt A. The Relational Data Model. In: Watt A, Nelson E, editors. *Database Design - 2nd edition*. BCcampus; 2014 [cited 2021 Jan 29]. Available from: <https://opentextbc.ca/dbdesign01>
- Zou D, Ma L, Yu J, Zhang Z. 2015. Biological databases for human research. *Genomics Proteomics Bioinformatics.*, 13(1):55–63.
- Zuniga PCC, Zuniga RAC, Mendoza MJ-A, Cariaga AA, Sarmiento RF, Marcelo AB. 2020. Workshop on Blockchain Use Cases in Digital Health. In: *Leveraging Data Science for Global Health*. Cham: Springer International Publishing;, 99–107.



Project website: www.digit-biotech.eu

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



**LO7 ENVIRONMENTAL
BENEFIT FROM MODERN
BIOTECHNOLOGY AND ICT
APPLICATIONS**

Advanced level

AUTHOR:
DIANA DI GIOIA



Contents

Renewable energy: biotechnology for biogas and bioethanol production	3
Biogas.....	3
Bioethanol	8
Biotechnology for bioplastic production (level B).....	14
Main steps towards modern BIOPLASTICS	14
1. Starch-based plastics	16
2. Polylactic acid plastics	18
3. Polyhydroxyalcanoates.....	19
Biotechnology for the remediation of contaminated sites	21
Microbial technologies for honeybee's health	24
References	33



Renewable energy: biotechnology for biogas and bioethanol production

BIOGAS

The biogas producing process has already been described in the basic level section. To sum up what already described, in developing countries there has been an increased interest in the development of technologies to produce renewable energy sources. Anaerobic digestion has received a new attention in recent years since the energy crisis of the early 1970s, and especially following the Gulf war. The process involves the treatment of agricultural and industrial waste of varying types in the production of biogas. Interest in the anaerobic treatment of agro-industry waste is increasing because it is economical, has lower energy requirements and is ecologically sound, among several other advantages, compared with aerobic treatment processes. The process produces digested sludge, which is mainly used as fertilizer for crop production since the nutrients in the raw material remain in the mineralized sludge as accessible compounds. Treating waste to yield fuel while recycling nutrients constitutes a sustainable cycle.

Anaerobic digestion is a complex, natural, two-stage process of degradation of organic compounds through a variety of intermediates into methane and carbon dioxide, by the action of a consortium of microorganisms. The interdependence of the bacteria is a key factor in the anaerobic digestion process. In the first stage, the volatile solids in manure are converted into fatty acids by anaerobic bacteria known as “acid formers.” In the second stage, these acids are further converted into biogas by more specialized bacteria known as “methane formers.” The anaerobic digestion process, which has been at work in nature for millions of years, can be managed to convert a farmer’s often problematic waste-stream into an asset. Instability during both the start-up and operation of the anaerobic degradation process can be problematic due to the low specific growth rate of the methanogenic microorganisms involved.

Here we give some extra details about the reactors used for biogas production. Several parameters are known to be important for the development and management of a biogas producing plant. In particular:

the working temperature. The process can be performed in:

- psychrophilic conditions (20°C) (not much used in conventional plants)
- mesophilic conditions (35 – 42°C)
- thermophilic conditions (> 50°C);

Mesophilic digestion. The digester is heated to 30–35°C and the feedstock remains in the digester typically for 15-30 days. Mesophilic digestion tends to be more robust and tolerant than the thermophilic process, but gas production is less, larger digestion tanks are required and sanitisation, if required, is a separate process stage.

Thermophilic digestion. The digester is heated to 55°C and the residence time is typically 12-14 days. Thermophilic digestion systems offer higher methane production, faster throughput, better pathogen and virus ‘kill’, but require more expensive technology, greater energy input and a higher degree of operation and monitoring. During this process 30–60% of the digestible solids are converted into biogas.

Therefore, the process in thermophilic condition is faster, but mesophilic conditions are used when the characteristics of the feeding substrate(s) change with time, season, etc.

→ **the solid content in the reactor.** We may distinguish:

- wet/humid processes (5 - 8% dry matter in the reactor)
- semi dry processes (dry matter = 8 - 20%)
- dry processes (dry matter >20%)

→ **the metabolic phases in the reactor.**

- ONE PHASE: the entire microbial chain is kept in a single reactor;
- TWO PHASES: the hydrolytic fermentative phase is separated from the methanogenic one.

How does a biogas plant work? Please check the website:

<https://www.youtube.com/watch?v=3UafRz3QeO8>

The following picture (Fig. 1) shows the different bioreactor configurations that can be developed to perform biogas production. They can differ for two parameters: the hydraulic scheme and the way microorganisms are working in the reactor (free or immobilized cells).

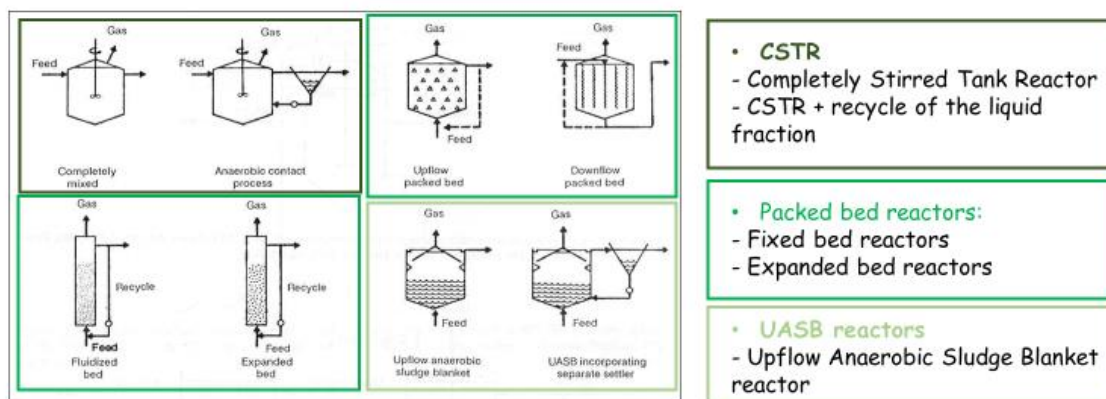


Fig. 1. Reactors for biogas production



2019-1-BG01-KA203-062371

The **continuous stirred-tank reactor (CSTR)** is a common model for a chemical reactor in environmental engineering. It is a batch reactor equipped with an impeller or other mixing device to provide efficient mixing. An ideal CSTR assumes perfect mixing. In a perfectly mixed reactor, the feeding is instantaneously and uniformly mixed throughout the reactor upon entry. Consequently, the performance is a function of residence time and reaction rate. The contact with the solid phase of the bioreactor can be improved by a sedimentation tank that separates the liquid medium from the solid part, which is then sent back to the bioreactor (referred to anaerobic contact process in the figure).

CSTRs consist of: a tank reactor (usually of constant volume), a stirring system to mix reactants (impeller or fast flowing introduction of reactants), feed and exit pipes to introduce reactants and remove products CSTR are commonly used in industrial processing. Biodigestors for biogas production are continuous agitated-tank reactors made of concrete or steel.

The **anaerobic packed-bed reactor** is filled with an inert support that provides a very large surface area for microbial growth. The influent passes through the media and anaerobic microbes attach themselves to the support creating a thin layer of anaerobic bacteria called biofilm—this film gives the digester its name, fixed film reactor or packed bed reactor. These microbes then continue to grow by removing material from the wastewater as it flows by. In most digesters the microbes are floating in the liquid and a portion of these active growing microorganisms are continuously discharged with the effluent. In a packed-bed digester the bacteria remain attached to the plastic support when effluent is discharged. Microorganisms are already “at work” when the new influent is added. Packed-bed digesters have smaller reactor vessels, shorter retention times and must be loaded with a feedstock that will readily flow through the media without clogging. Three to five day retention times are typical and digesters can be run at ambient temperatures in hot climates but are usually heated to mesophilic or thermophilic temperatures.

Which are the advantages of anaerobic packed-bed reactor? Increased stability and performance in anaerobic reactors can be achieved if the microbial consortium is retained in the reactor. Two means of achieving this are to use dense bacterial granula as in **UASB reactors** or a microbial biofilm attached to inert carriers in the above described packed-bed reactors. **Upflow anaerobic sludge blanket (UASB)** technology, normally referred to as UASB reactor, is indeed a form of anaerobic digester that is used for wastewater treatment and as a methanogenic (methane-producing) digester. A similar but variant technology to UASB is the expanded granular sludge bed (EGSB) digester (Fig. 2). An **expanded granular sludge bed (EGSB) reactor** is a variant of the UASB concept. The distinguishing feature is that a faster rate of upward-flow velocity is designed for the wastewater passing through the sludge bed. The increased flux permits partial expansion (and from this the name of the reactor is derived) of the granular sludge bed, improving wastewater-sludge contact as well as enhancing segregation of small inactive suspended particle from the sludge bed. The increased flow velocity is either accomplished by utilizing tall reactors, or by incorporating an effluent recycle (or both).

2019-1-BG01-KA203-062371

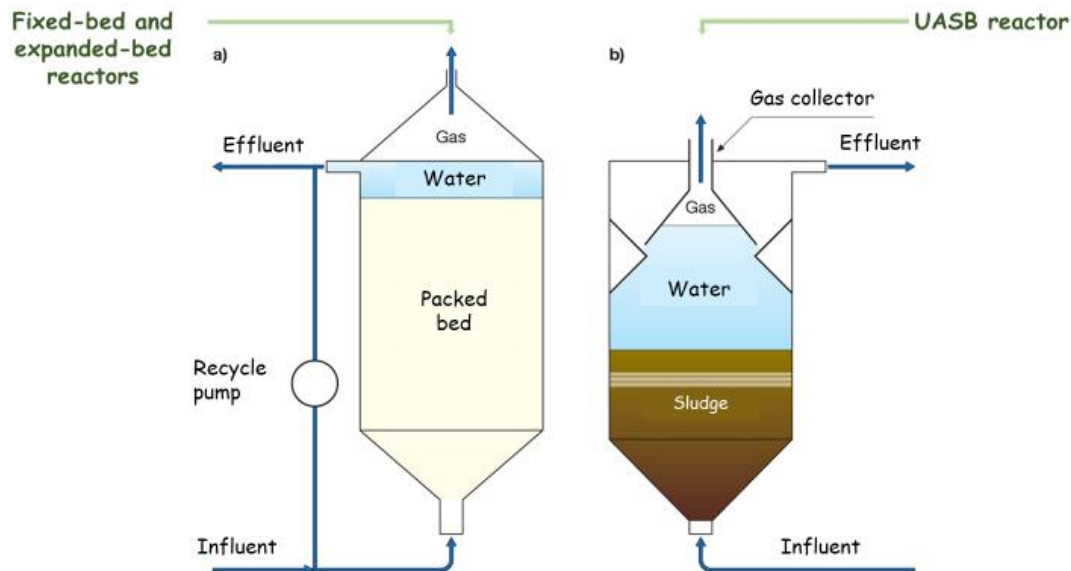


Fig. 2. Fixed bed/Expanded bed reactors (left) and UASB reactor

UASB is an anaerobic process that forms a blanket of granular sludge which suspends in the tank. Wastewater flows upwards through the blanket and is processed by the anaerobic microorganisms. The upward flow combined with the settling action of gravity suspends the blanket with the aid of flocculants. The blanket begins to reach maturity at around three months. Small sludge granules begin to form and they contain organic matter and bacteria without any support matrix, the flow conditions create a selective environment in which only those microorganisms capable of attaching to each other survive and proliferate. Eventually the aggregates form dense compact structures referred to as "granules". Biogas with a high concentration of methane is produced, and this may be captured and used as an energy source, to generate electricity for export and to cover its own running power. The technology needs constant monitoring when put into use to ensure that the sludge blanket is maintained, and not washed out (thereby losing the effect). The heat produced as a by-product of electricity generation can be reused to heat the digestion tanks. The packing medium in the packed-bed reactor and the granular sludge in the UASB reactor serve as a filter preventing bacterial washout and also providing a larger surface area for faster biofilm development and improved methanogenesis. Specific surface area, porosity, surface roughness, pore size, and orientation of the packing material were found to play an important role in anaerobic reactor performance. Biofilm or fixed-film reactors depend on the natural tendency of mixed microbial populations to adsorb onto surfaces and to form a biofilm. Many carrier materials have been investigated regarding their suitability as supports for biofilm, including cheap, readily available materials like sand, clay, glass, quartz and a number of plastics. In nature, microorganisms inhabit the



2019-1-BG01-KA203-062371

outer and inner surfaces of stone, gravel or sand. This biofilm formation becomes an important factor for water self-cleaning ability. The growth of microorganism in a biofilm is the basis for biological water treatment such as denitrification and for intensification of aerobic and anaerobic wastewater treatment. The use of packed-bed reactors to treat different kinds of wastewater has also been reported, for example, dairy and brewery wastewater. The biofilm formation on carrier materials improves the conversion rates by reducing its sensitivity toward concentration variations and inhibiting substance. The efficiency of removing organic matter in fixed-bed reactors is directly related to the characteristics of the support material used for immobilization of anaerobes. Reticular polyurethane foam has a high specific surface area. It is an excellent colonization matrix for an anaerobic filter reactor. Pore size was one of the most important parameter for microbiological and engineering requirements in high-efficiency beds. Many kinds of bedding model have been considered for degrading a variety of organic wastes in anaerobic digestion reactors.

The development of fixed biomass reactors has ensured that significant advances in the knowledge and application of anaerobic processes for waste treatment have taken place. Compared to conventional units, fixed film bioreactors perform efficiently at higher organic loading rates, due to more effective biomass retention in the reaction zone resulting in higher cellular retention times. Immobilized biomass anaerobic reactors also show better responses to organic shock loads and toxic inputs. In many cases, immobilized biomass reactors completely recover their performance after such troubles

How does a UASB reactor work? Please have a look at this video
<https://www.youtube.com/watch?v=0QsEdlJgIII>

A useful digestion output of the anaerobic digestion process is digestate. Digestate is the remaining part of the degraded biomass after biogas production: it is stable organic matter rich in various nutrients (N, P, K). Depending on the feedstock used for biogas production, digestate can be directly usable as organic fertiliser in the same way raw animal slurries are spread on fields in agriculture. It can also be further upgraded to recover high quality mineral nutrients. Digestate use as organic fertiliser displays multiple advantages: it allows reuse of nutrients and substitutes mineral fertiliser of fossil origin. Compared to raw manure, digestate is also sanitised thanks to the biogas production process neutralising most of the pathogens of the original feedstock such as bacteria and crop diseases. Digestate homogeneity and density also allow for faster penetration in the soil compared to raw manure, making nutrients more easily accessible to plants in the soil. If unfit for agricultural purposes, digestate can be further processed and used as a raw material for industrial processes.

An overall scheme of a biogas producing plant is presented below (Fig. 3).

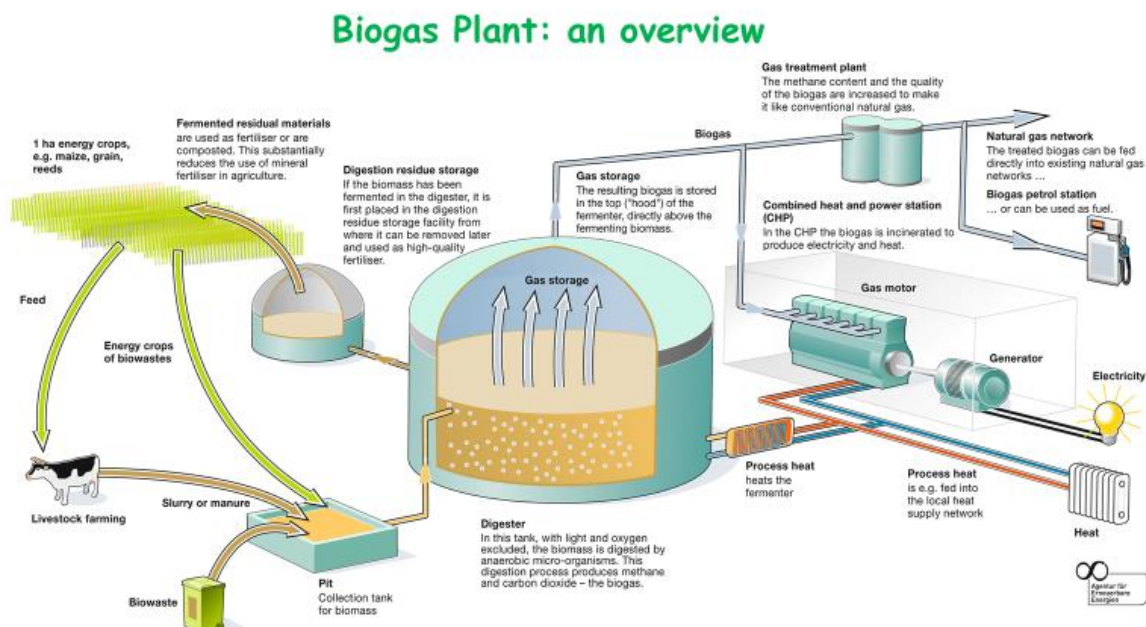


Fig. 3. A biogas producing plant

BIOETHANOL

Countries worldwide have considered and directed policies toward the increased and economic utilization of biomass for meeting their future energy demands in order to meet carbon dioxide reduction targets as specified in the Kyoto Protocol as well as to decrease reliance and dependence on the supply of fossil fuels. Although biomass can be a huge source of transport fuels such as bioethanol, biomass is commonly used to generate both power and heat, generally through combustion. Ethanol is at present the most widely used liquid biofuel for motor vehicles. The importance of ethanol is increasing due to a number of reasons such as global warming and climate change.

The global market for bioethanol has entered a phase of rapid, transitional growth. Many countries around the world are shifting their focus toward renewable sources for power production because of depleting crude oil reserves. The trend is extending to transport fuel as well. Ethanol has potential as a valuable replacement of gasoline in the transport fuel market. However, the cost of bioethanol production is more compared to fossil fuels. Brazil and the USA are the two major ethanol producers accounting for 62% of the world production. Large scale production of fuel ethanol is mainly based on **sucrose from sugarcane in Brazil or starch, mainly from corn, in the USA**. Please see a schematic figure below (Fig. 4)

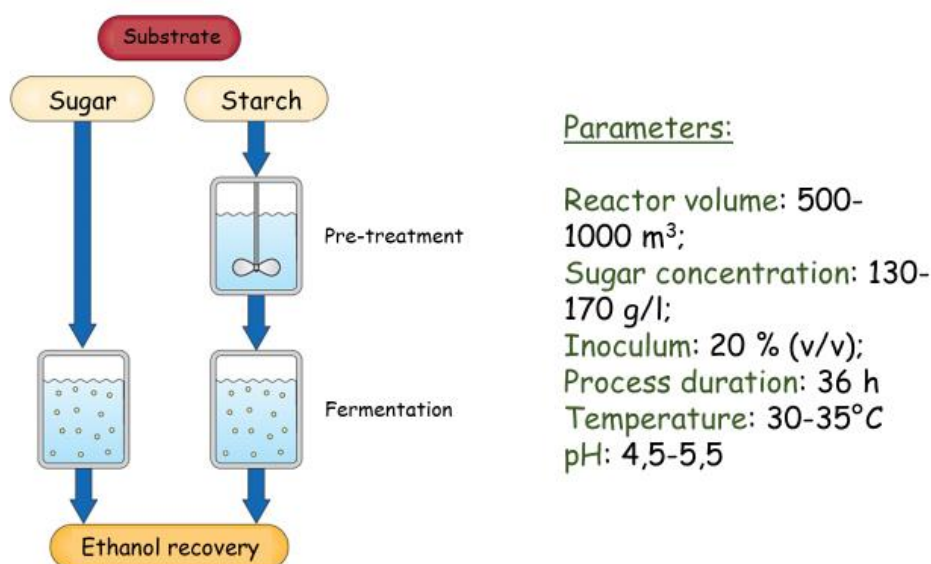


Fig. 4. Schematic production of bioethanol from sugar crops

Current ethanol production based on corn, starch and sugar substances may not be desirable due to their food and feed value. Cost is an important factor for large scale expansion of bioethanol production. The green gold fuel from lignocellulosic wastes avoids the existing competition of food versus fuel caused by grain-based bioethanol production. Hence bioethanol production could be the route to the effective utilization of agricultural wastes. Rice straw, wheat straw, corn straw, and sugarcane bagasse are the major agricultural wastes in terms of quantity of biomass available.

Lignocellulosic materials are renewable, low cost and are abundantly available. It includes crop residues, grasses, sawdust, wood chips, etc. Extensive research has been carried out on ethanol production from lignocellulosics. Lignocellulosics are processed for bioethanol production through three major operations:

- 1) pretreatment for delignification is necessary to liberate cellulose and hemicellulose before hydrolysis;
- 2) hydrolysis of cellulose and hemicellulose to produce fermentable sugars including glucose, xylose, arabinose, galactose, mannose and fermentation of reducing sugars.
- 3) The non-carbohydrate components of lignin also have value added applications

The most important processing challenge in the production of biofuel is pretreatment of the biomass. Lignocellulosic biomass is composed of three main constituents namely hemicellulose, lignin



and cellulose. Pre-treatment methods refer to the solubilization and separation of one or more of these components of biomass. It makes the remaining solid biomass more accessible to further chemical or biological treatment. The lignocellulosic complex is made up of a matrix of cellulose and lignin bound by hemicellulose chains. The pretreatment is done to break the matrix in order to reduce the degree of crystallinity of the cellulose and increase the fraction of amorphous cellulose, the most suitable form for enzymatic attack. Pretreatment is undertaken to bring about a change in the macroscopic and microscopic size and structure of biomass as well as submicroscopic structure and chemical composition. It makes the lignocellulosic biomass susceptible to quick hydrolysis with increased yields of monomeric sugars.

The goals of an effective pretreatment process are:

- (i) formation of sugars directly or subsequently by hydrolysis to avoid loss and/or degradation of sugars formed
- (ii) to limit formation of inhibitory products
- (iii) to reduce energy demands and minimize costs.

Physical, chemical, physicochemical and biological treatments are the four fundamental types of pretreatment techniques employed. In general, a combination of these processes is used in the pretreatment step.

Among **physical pretreatment**, the first step for ethanol production from agricultural solid wastes is the *mechanical size reduction* through milling, grinding, or chipping. This reduces cellulose crystallinity and improves the efficiency of downstream processing. *Pyrolysis* is a physical treatment: the materials are treated at a temperature higher than 300°C, whereby cellulose rapidly decomposes to produce gaseous products and residual char. The residual char is further treated by leaching with water or with mild acid. The water leachate contains enough carbon source to support microbial growth for bioethanol production. Glucose is the main component of water leachate. Pretreatment of lignocellulosic biomass in a *microwave* oven is also a feasible method which uses the high heating efficiency of a microwave oven. Microwave treatment utilizes thermal and non-thermal effects generated by microwaves in aqueous environments. Heat is generated in the biomass by microwave radiation, resulting from the vibrations of the polar bonds in the biomass and the surrounding aqueous medium. This unique heating feature results in an explosion effect among the particles and improves the disruption of recalcitrant structures of lignocellulose. In the non-thermal method, i.e., the electron beam irradiation method, polar bonds vibrate, as they are aligned with a continuously changing magnetic field and the disruption and shock to the polar bonds accelerates chemical, biological and physical processes.

Among **physicochemical treatments**, steam explosion is a promising one making biomass more accessible to cellulase attack. This method of pretreatment does not use any catalyst and the biomass fractionates to yield levulinic acid, xylitol and alcohols. In this method the biomass is heated using high-pressure steam (20-50 bar, 160-290°C) for a few minutes; the reaction is then stopped by sudden decompression to atmospheric pressure. When steam is allowed to expand within the lignocellulosic



matrix it separates the individual fibers. The high recovery of xylose (45-65%) makes steam-explosion pretreatment economically attractive.

Chemical pretreatment methods involve the usage of dilute acid, alkali, ammonia, organic solvent, CO₂ or other chemicals. These methods are easy in operation and have good conversion yields in short span of time. *Acid pretreatment* is considered as one of the most important techniques and aims for high yields of sugars from lignocellulosics. It is usually carried out by concentrated or diluted acids (usually between 0.2% and 2.5% w/w) at temperatures between 130 °C and 210 °C. The acid medium attacks the polysaccharides, especially hemicelluloses which are easier to hydrolyze than cellulose. However, acid pretreatment results in the production of various inhibitors like acetic acid, furfural and 5- hydroxymethylfurfural. These products are growth inhibitors of microorganisms. Hydrolysates to be used for fermentation therefore need to be detoxified. *Alkaline pretreatment* of lignocellulosics digests the lignin matrix and makes cellulose and hemicellulose available for enzymatic degradation. Alkali treatment of lignocellulose disrupts the cell wall by dissolving hemicelluloses, lignin, and silica, by hydrolyzing uronic and acetic esters, and by swelling cellulose. Crystallinity of cellulose is decreased due to swelling. By this process, the substrates can be fractionated into alkali-soluble lignin, hemicelluloses, and residue, which makes it easy to utilize them for more valuable products. The end residue (mainly cellulose) can be used to produce either paper or cellulose derivatives. *Organic solvent* are alternative methods for the delignification of lignocellulosic materials. The utilization of organic solvent/water mixtures eliminates the need to burn the liquor and allows the isolation of the lignins (by distillation of the organic solvent). Examples of such pretreatments include the use of 90% formic acid and that of pressurized carbon dioxide in combination (50% alcohol/water mixture and 50% carbon dioxide). Other various organic solvents which can be used for delignification are methanol, ethanol, acetic acid, performic acid and peracetic acid, acetone, etc.

Biological treatments. Enzymatic hydrolysis is the preferred saccharification method because of its higher yields, higher selectivity, lower energy cost and milder operating condition than chemical processes.

Different mode of fermentation. Fermentation of bioethanol can be carried out in batch, fed-batch, repeated batch, or continuous mode. In batch process, substrate is provided at the beginning of the process without addition or removal of the medium. It is known as the simplest system of bioreactor with flexible and easy control process. The fermentation process is carried out in a closed-loop system with high sugars concentration at the beginning and ends with high product concentration. There are several benefits of batch system including complete sterilization, does not require labour skills, it is easy to manage the feedstocks, and can be controlled easily. However, the productivity is low and needs intensive and high labour costs. The presence of high sugar concentration in the fermentation medium may lead to substrate inhibition of cell growth and ethanol production. Cells recycle batch fermentation is a strategic method for effective ethanol production as it reduces time and cost for inoculum preparation. The other advantages of repeated-batch process are easy cell collection, stable operation, and long-term productivity. Sugar materials and immobilized yeast cells are used to facilitate cell separation for cell recycling. However, its application in the process of lignocellulosic materials is



2019-1-BG01-KA203-062371

extremely difficult because lignocellulosic residue remain in the fermentation medium together with yeast cells. The use of free cells in this system reduces yeast cell concentration and results in lower ethanol production in the subsequent batches. Repeated-batch fermentation can be performed by replacing free cells with the immobilized cells. Fed-batch fermentation is a combination of batch and continuous mode which involves the addition of substrate into the fermenter without removing the medium. It has been used to overcome the problem of substrate inhibition in batch operation. Volume of culture in fed-batch processes can vary widely but it must be fed properly at certain rate with the right component composition. Productivity of fed-batch fermentation can be increased by maintaining substrate at low concentration which allows the conversion of sufficient amount of fermentable sugars to ethanol. This process has higher productivity, higher dissolved oxygen in medium, shorter fermentation time and lower toxic effect of the medium components compared to other types of fermentation. However, ethanol productivity in fed-batch is limited by feed rate and cell mass concentration.

Continuous operation is carried out by constantly adding substrates, culture medium and nutrients into a bioreactor containing active microorganisms. Culture volume in continuous operation must be constant and the fermentation products are taken continuously from the media. Various type of products can be obtained from the top of the bioreactor such as ethanol, cells and residual sugar. The advantages of continuous system over batch and fed-batch system are higher productivity, smaller bioreactor volumes and less investment and operational costs. At high dilution rate, ethanol productivity is increased while ethanol yield is decreased due to incompletely substrate consumption by yeasts. However, the possibility for contamination to occur is higher than other types of fermentation. Moreover, the ability of yeasts to produce ethanol in continuous process are reduced due to long cultivation time.

Factors affecting bioethanol production

Several factors influence the production of bioethanol: temperature, sugar concentration, pH, fermentation time, agitation rate, and inoculum amount. The growth rate of the microorganisms is directly affected by the temperature. High temperature which is unfavorable for cells growth becomes a stress factor for microorganisms. The ideal temperature range for fermentation is between 20 and 35°C for *Saccharomyces cerevisiae*. Free cells of *S. cerevisiae* have an optimum temperature near 30°C whereas immobilized cells have slightly higher optimum temperature due to its ability to transfer heat from particle surface to inside the cells. Moreover, enzymes which regulate microbial activity and fermentation process are sensitive to high temperature which can denature its tertiary structure and inactivates the enzymes. Thus, temperature is carefully regulated throughout the fermentation process.

The increase in sugar concentration up to a certain level caused fermentation rate to increase. However, the use of excessive sugar concentration will cause steady fermentation rate. This is because the concentration of sugar use is beyond the uptake capacity of the microbial cells. Generally, the maximum rate of ethanol production is achieved when using sugars at the concentration of 150 g/L. The initial sugar concentration also has been considered as an important factor in ethanol production. High



2019-1-BG01-KA203-062371

ethanol productivity and yield in batch fermentation can be obtained by using higher initial sugar concentration. However, it needs longer fermentation time and higher recovery cost.

Ethanol production is influenced by pH of the broth as it affects bacterial contamination, yeast growth, fermentation rate and by-product formation. The permeability of some essential nutrients into the cells is influenced by the concentration of H^+ in the fermentation broth. Moreover, the survival and growth of yeasts is influenced by the pH in the range of 2.75–4.25. In fermentation for ethanol production, the optimum pH range of *S. cerevisiae* is 4.0–5.0 [34]. When pH is lower than 4.0, a longer incubation period is required but the ethanol concentration is not reduced significantly. However, when then pH was above 5.0, the concentration of ethanol reduces substantially.

Fermentation time affects the growth of microorganisms. Shorter fermentation time causes inefficient fermentation due to inadequate growth of microorganisms. On the other hand, longer fermentation time gives toxic effect on microbial growth especially in batch mode due to the high concentration of ethanol in the fermented broth. Complete fermentation can be achieved at lower temperature by using longer fermentation time which results in lowest ethanol yield.

Agitation rate controls the permeability of nutrients from the fermentation broth to inside the cells and removal of ethanol from the cell to the fermentation broth. The greater the agitation rate, the higher the amount of ethanol produced. Besides, it increases the amount of sugar consumption and reduces the inhibition of ethanol on cells. The common agitation rate for fermentation by yeast cells is 150–200 rpm. Excess agitation rate is not suitable for smooth ethanol production as it causes limitation to the metabolic activities of the cells.

Inoculum concentration does not give significant effects on the final ethanol concentration, but it affects the consumption rate of sugar and ethanol productivity.

Biotechnology for bioplastic production (level B)

MAIN STEPS TOWARDS MODERN BIOPLASTICS

- ✓ Bioplastics are not a real innovation: natural resins were used since ancient times (for example amber, shellac, etc.)
- ✓ Starting from 1860, the first plastics deriving from cellulose were released (eg. celluloid, cellophane)
- ✓ In the 1940s, Henry Ford made car parts with plastics obtained from soy
- ✓ In the '50s plastics derived from oil spread
- ✓ Oil crisis in the 70s: the interest in bioplastics was rediscovered.

Currently there is an increase in the demand for bioplastics, mainly due to the pressing environmental problems (depletion of resources, greenhouse effect, waste disposal, etc.). Fig. 5, 6 and 7 give some basic info about bioplastics.

Bioplastics: state of art

Packaging waste accounts for more than a third of municipal solid waste.

The EU directive 2004/12/CE (regarding packaging and packaging waste) defined that, by 2009, 60% of packaging waste must be recycled. According to this rule, composting is considered as a way of recycling waste.

In compliance with this directive, the individual member states took various measures. Italy and France decided that since 2010 all plastic bags had to be made of biodegradable plastics. In Italy this measure was implemented from January 2011. This leads to a potential saving of 300 000 tons of plastic bags per year, which corresponds to 430 000 tons of oil and 200 000 tons of CO₂.

In December 2015, the EC adopted an EU action plan for the circular economy, in which plastics were identified as a key priority and committed itself to elaborating "one strategy to address the challenges posed by plastics at all stages of the value chain and take into account their whole life cycle."

In 2017, it confirmed its intention to work towards achieving the objective of the recyclability of all plastic packaging by 2030.



Fig. 5. Bioplastics and European regulations

Plastic and rubber are polymeric materials consisting of monomers. These are mainly produced from petroleum and the originated material is therefore non-renewable. Around 4% of the world's oil consumption is used as raw material in plastic production, and a similar amount is used as energy in the

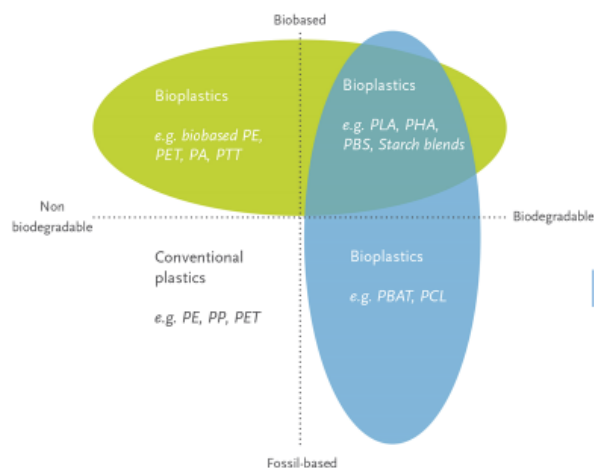
2019-1-BG01-KA203-062371

production process. In addition to petroleum, plastic production requires the use of chemical additives such as plasticizers, flame retardants, heat and UV stabilizers, biocides, pigments, and extenders. Several additives are classified as hazardous according to the EU regulations (carcinogenic, mutagenic, harmful for reproductive health or for aquatic life, or having persistent negative impacts on the environment).

In the '60s plastic was considered for the first-time as a concern in sea and ocean pollution and negative health impacts on humans and the environment were starting to be described. Indeed, plastics release toxic chemicals throughout the life cycle of the product.

Plastic **recycling** emerged as a possible solution. Recycling, however, is not the only solution needed to solve the plastic waste crises that is polluting the environment. Plastic can range from being unrecyclable, recyclable only once or twice, or at a defined number of times but not forever. After this limit, the plastic will end up in a landfill. Furthermore, a lot of plastic consumers do not even allow their plastic to have this long of a life. **Renewable plastics**, meaning plastics derived from renewable sources and easily biodegradable in the environment, may offer a solution to the problem of bioplastic poisoning.

What are bioplastics?



Bioplastics are not just one single material. They comprise of a whole family of materials with different properties and applications. According to European Bioplastics, a plastic material is defined as a bioplastic if it is either biobased, biodegradable, or features both properties.

Bioplastics are **biobased, biodegradable, or both**.
'Biobased' does not equal 'biodegradable'

Fig. 6. Biobased and biodegradable plastics

Global production bioplastics capacities by material type (2019)

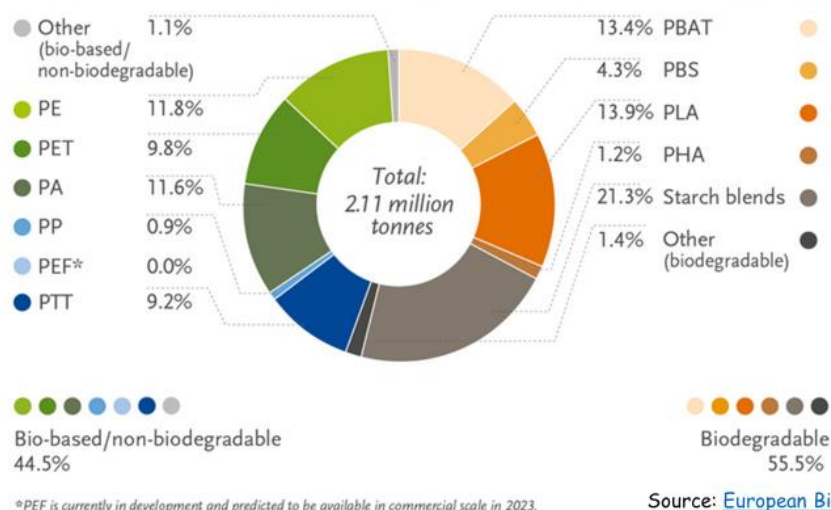


Fig. 7. Main bioplastics produced worldwide

Bio-based is defined in European standard EN 16575 as “derived from biomass”. Biodegradable materials are materials that can be broken down by microorganisms like bacteria or fungi into water, carbon dioxide or methane and biomass. However, biodegradability depends on the environmental conditions: presence of microorganisms, temperature, and availability of oxygen and water. Compostable materials are materials that break down at composting conditions. Industrial composting conditions require elevated temperature (55°C - 60°C) combined with a high relative humidity and the presence of oxygen, and they are in fact optimal when compared against other degradation conditions like in soil, surface water and marine water. Compliance with EN 13432 is considered a good measure for compostability of packaging materials. According to this standard, plastic packaging can be called compostable. Some details on the 3 main categories of bioplastics are given below.

1. Starch-based plastics

75% of all organic material on earth is present in the form of polysaccharides. An important polysaccharide is starch. Plants synthesize and store starch in their structure as an energy reserve. Starch is found in seeds, tubers, or roots of the plants. Sources of starch are corn, wheat, rice, potato, tapioca, pea, and many other plant resources. Most of the starch produced worldwide is derived from corn. Starch is generally extracted from plant resource by wet milling processes. Starch consists of two types of glucose polymers: amylose and amylopectin. Amylose is essentially a linear polymer in which glucose units are predominantly connected through α -D-(1, 4) glucosidic bonds. Amylopectin is a branched

polymer, containing periodic branches linked with the backbones through α -D-(1, 6) glucosidic bonds. The content of amylose and amylopectine in starch varies and depends on the starch source.

An important class of plastics is represented by starch-based plastics. Beginning in the early 1990s, research and technology developments have permitted to complex natural polymers like starch (from maize, potato etc.) with biodegradable macromolecules (polymeric complexing agents) in order to obtain thermoplastic and biodegradable innovative materials on an industrial scale. In particular, Novamont's starch-based technology (Fig. 8) employs processing conditions able to almost completely destroy the crystallinity of amylose and amylopectin, in the presence of macromolecules, which are able to form a complex with amylose. They can be of natural or synthetic origin and are biodegradable. The complex formed by amylose with the complexing agent is generally crystalline and it is characterised by a single helix of amylose formed around the complexing agent. Unlike amylose, amylopectin does not interact with the complexing agent and remains in its amorphous state. The source of the starch, i.e. its ratio between amylose and amylopectin, the processing conditions and the nature of the complexing agents allow engineering of various supramolecular structures with very different properties. Over the last few years many successful efforts have been made to increase the amount of renewable raw materials for producing biodegradable polyesters. Novamont is therefore one of the most important players in starch-based bioplastics. The company is currently working in the development of a biorefinery project consisting of an innovative development model capable of synthesising various chemical intermediates using renewable raw materials cultivated with low input and in marginal areas instead of fossil raw materials.

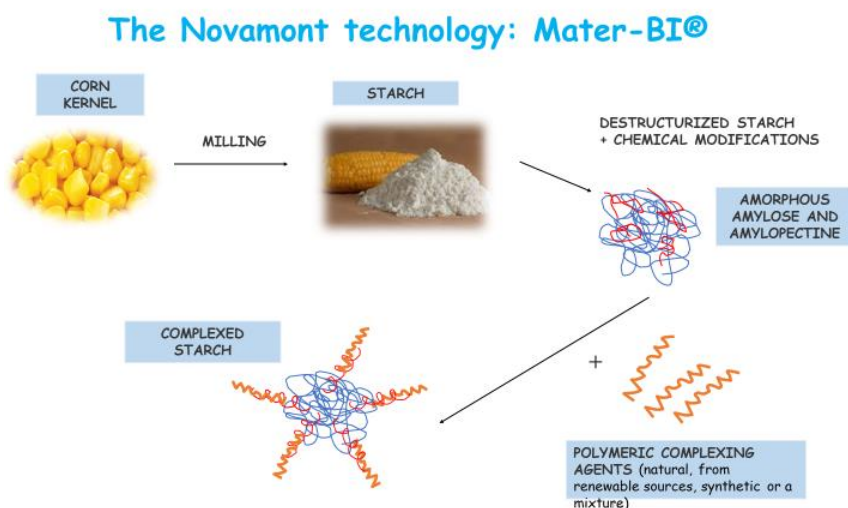


Fig. 8. Novamont's starch-based technology

2. Polylactic acid plastics

Synthetic biodegradable polylactones such as poly-lactic acid (PLA), poly-glycolic acid (PGA), and polycaprolactone (PCL) are polymers that are degraded by simple hydrolysis of the ester bonds. The hydrolytic products from such degradation process are then transformed into non-toxic subproducts (Fig. 9).

PLA plastics are derived from the fermentation of agricultural by-products such as starch-rich substances like maize, wheat or sugar and corn starch. The process involves conversion of corn, or other carbohydrate sources

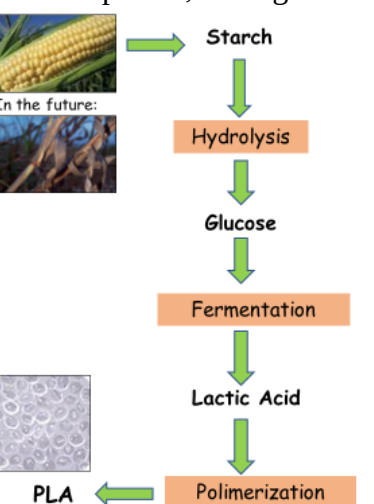


Fig. 10. PLA production from starch

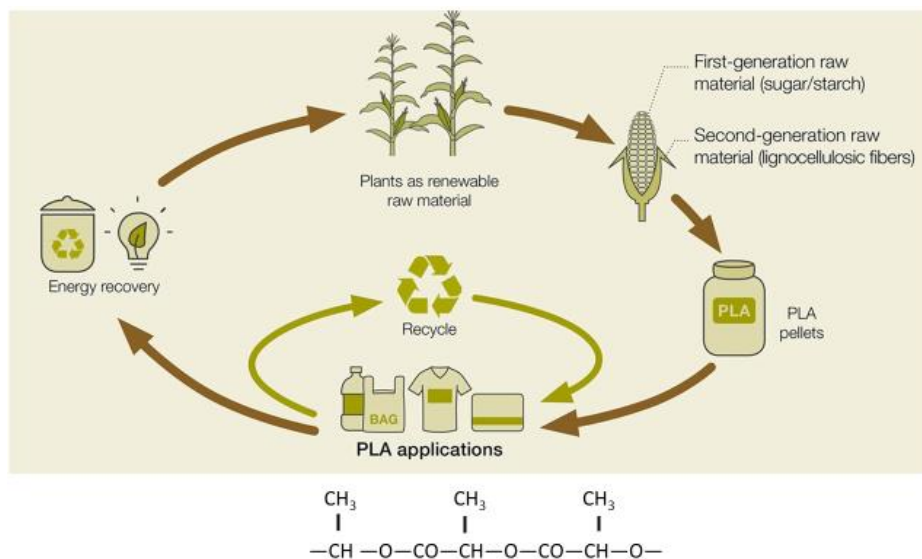


Fig. 9. PLA life cycle

into glucose followed by fermentation into lactic acid (Fig. 9 and 10). PLA derived from lactic acid is thermoplastic, biodegradable aliphatic polyester having ample potential for packaging applications. The

lactic acid monomers are either directly polycondensed or undergo ring opening polymerization of lactide resulting in formation of PLA pellets. The properties of PLA as packaging material depend on the ratio between the two optical isomers of the lactic acid monomer. When 100% L-PLA monomers are used it results in very high crystallinity and melting point, whereas 90/10% D/L copolymers fulfils the requirements of bulk packaging. PLA is the first



2019-1-BG01-KA203-062371

biobased polymer commercialized on a large scale and can be shaped into injection moulded objects, films and coatings. PLA has replaced high-density polyethylene, low-density polyethylene (LDPE), polyethylene terephthalate and PS as packaging material.

The main properties of PLA are: i) the mechanical resistance and heat sensitivity are similar to traditional plastics; ii) hardness, stiffness and degree of elasticity are similar to PET, iii) it can contain fats, oils, alcohol and aliphatic molecules, iv) scarce resistance to acids and bases is but good resistance to UV radiation, v) it can be printed and dyed, vi) it can be transformed into goods through standard machines used for traditional plastics, vii) the post-use phase may involve composting in industrial plants.

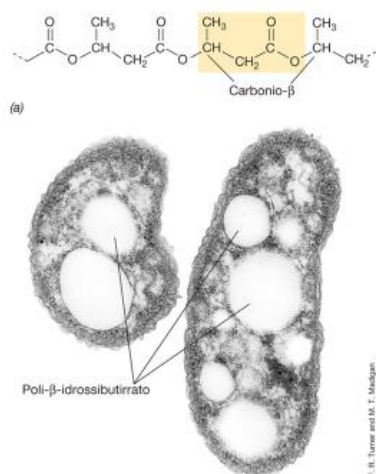
3. Polyhydroxyalkanoates

Polyhydroxyalkanoates (PHAs) are bio-degradable polymers that are accumulated by some bacteria as storage compound in form of intracellular granules. PHA is one of the biopolymers that can effectively replace the conventional petrochemical plastics with their material properties that parallel them. Even then, their production at large scale is still limited by its high production cost compared with conventional fossil-fuel based plastics as the PHA price, depending on polymer composition, ranges from 2.2 to 5.0 €/kg that is at least three times higher than the major petrochemical based polymers which cost less than 1.0 €/kg (calculations made in 2016).

In the majority of companies producing PHAs, mostly pure cultures are used. The problem with the use of pure cultures is the requisites for sterility, refined substrates if plant-based feedstocks are not used, thus limiting the process of commercialisation. All these issues shall be overcome using Mixed Microbial Cultures (MMCs): this combines the transformation of waste into value added product production. Biological treatment of wastewater and sludge management for recovering carbon from wastewater as PHAs is a route to transform end-of-pipe environmental protection infrastructure into bio-refineries. Integration strategies for MMC PHA production within wastewater treatment processes have been proposed for industrial process wastewater and municipal wastewater treatment.

2019-1-BG01-KA203-062371

One of the best characterized members of the PHA family is polyhydroxybutyrate (PHB), produced by microorganisms that store it inside the cell cytoplasm. In 1926, a microbial production of linear polyester of D (-)-3-hydroxybutyric acid as intracellular granules, which occurred in both gram-



- Discovered for the first time in 1926 in *Bacillus megaterium*
- It was then found in many other bacterial genera (e.g. *Pseudomonas* and *Alcaligenes*)
- Stored inside the cell in big granules until 90% of its dry weight
- Produced when the cell is under a nutritional stress

Fig. 11. Polyhydroxybutyrate

difficulties are undoubtedly related to production costs, both upstream and downstream processes. Approximately, 40% and 50% of overall production cost of PHB have been assigned to crude material and separation/purification systems, respectively. In bioextraction techniques, genetic engineering is the most commonly used to introduce microorganisms, and they are capable of effectively extracting PHB from PHB accumulating cells. There are several approaches that have been investigated including bacteriophage-mediated lysis system and predatory bacteria, which are better than the conventional extraction approaches that produce environmentally harmful solvents, with higher cost of degradation. Thus, due to the non-environmental and non-economical friendly properties of conventional extraction methods, more attention is given to bioextraction systems.

positive and gram-negative bacteria under a starvation conditions, was first discovered (Fig. 11). Cost is the major drawback of PHB production during industrialization. Industrial production of PHB is costly than that of petroplastics. Large quantities PHB production is estimated about 4.4 USD/kg, i.e. far more expensive than polypropylene production cost, which is close to 1 USD/kg. The financial

Biotechnology for the remediation of contaminated sites

This unit is basically focused on the technologies that allow the study of the microbiota in soil or complex matrixes. The main technologies are depicted below (Fig. 12).

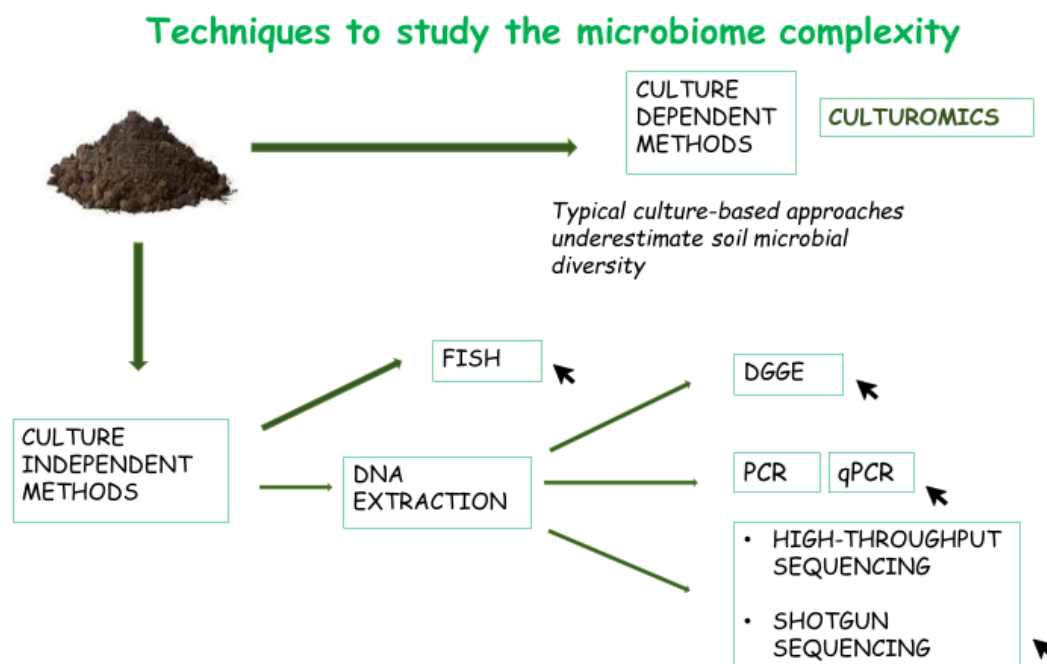


Fig. 12. Techniques to study the microbiome complexity

The culturomic approach is well described in the video that has been produced within Digit-Biotech. This "classical method", provide for the identification of microorganisms through the isolation of pure cultures, followed by tests that analyze some morpho-physiological and biochemical characteristics. These analyzes are often not sufficient for the identification of most species of microorganisms and moreover are limited to cultivable species which represent a very small percentage of all species found in nature. These tests also have the serious limitation of requiring considerable time consuming. However, they have the great advantage of the obtainment through the isolation approach of target microorganisms that can be used in the bioremediation approach.

Over the past few decades, research in the field of microbiology environmental have shown that microbial communities play a functional role of control of ecosystems that is not attributable to



2019-1-BG01-KA203-062371

individual species but to the communities themselves as "functional units". This functional activity of microbial communities is, in many cases, responsible for important processes for humans, including the biodegradation of original waste in wastewater treatment plants and landfills, composting and, in general, all the processes in which chemical transformations of the substances produced by the activities take place. Denaturing Gradient Gel Electrophoresis (DGGE), real time PCR (or quantitative PCR), and whole genome approach (high-throughput sequencing and shotgun sequencing) are the main technologies used for the study of microbial populations in the environment.

DGGE: it is an electrophoretic separation technique used for the separation and analysis of DNA fragments that differ in the nucleotide sequence also of a single base pair. In classical electrophoresis conducted on agarose or acrylamide gel, DNA fragments are separated on the basis of molecular weight; the running speed decreases parallel to the increase in length of the fragment. On the contrary, in the DGGE, fragments of DNA of equal molecular weight are separated according to the denaturation pattern. The presence of heat or chemical denaturants allows the denaturation of the two filaments constituents of a double-stranded DNA (dsDNA) molecule. Temperature and concentration of denaturant to which the separation of the two filaments occurs strongly depend on the sequence of the fragment itself. In particular, the determining factors are: quantity of bonds hydrogen that are established between complementary bases and type of interactions that are established between bases adjacent on the same strand (stacking interaction). A DNA molecule therefore has domains with characteristic melting temperatures or T_m , determined by nucleotide sequence. DNA fragments almost identical in molecular weight, but that they also differ in a single nucleotide, they can be characterized by T_m and melting domains different from each other. The DGGE analysis is conducted on polyacrylamide gel containing a gradient denaturing in such a way that the dsDNA is subject, during the run, to an increase in denaturation conditions with consequent separation at the melting domains. In the upper part of the gel, where there are mild denaturation conditions, the melting domains at lower T_m begin to partially denature, creating branched molecules with less mobility. The increase in denaturation conditions along the polyacrylamide gel can determine the total dissociation of partially denatured fragments in single-stranded DNA (ssDNA). Experimentally, the complete dissociation of the two dsDNA strands is hindered by introducing, at the end of each filament, domains characterized by high contained in G + C and high T_m . G + C-rich regions are artificially created at one end of the dsDNA by means of incorporation of a GC-clamp during amplification reactions. The incorporation of the GC-clamp is made possible by the use of primers characterized by a sequence of about 30-40 GC at the 5' end. The presence of the GC-clamp of the same sequence at the extremity of each molecule causes the differences between the stroke profiles of the analyzed fragments to be mainly determined by variations in the sequence of low melting domains. Since T_m is determined by the nucleotide sequence, the presence of a single mutation is capable of generating a different denaturation profile and, consequently, a different electrophoretic run. Hence the recurrence of polymorphisms in highly conserved genes can be analyzed by DGGE and can provide useful information to characterize the structure of microbial communities. In fact, with denaturing gradient gel electrophoresis an electrophoretic profile formed by a series of bands is obtained in which, as a first approximation, the number of bands is proportional to the number of species present and the position of each band is different for each species. The DGGE technique therefore



2019-1-BG01-KA203-062371

provides a simple approach to obtaining microbial community profiles that can be used to identify spatial and temporal differences in the community structure or to monitor changes in structure that occur in response to environmental disturbances.

Real time PCR: It is a technique that allows to amplify and at the same time quantify a target DNA sequence. It involves the use of fluorescent dyes, such as Sybr Green that intercalate in the minor sulcus of the DNA double strand, or probes with specific sequences, consisting of oligonucleotides labeled with fluorescent agents. The emitted fluorescence is constantly measured and provides "real time" information on the amount of amplicon produced. From an amplification reaction a graph with a sigmoidal curve is obtained; this will start as soon as possible the greater the quantity of starting DNA and will continue to grow with an exponential trend until it reaches a maximum value (plateau), in which the reaction will slow down due to the exhaustion of the substrates.

In studying a Real Time PCR graph, three parameters are established:

- the fluorescence baseline or baseline region;
- the threshold line, parallel to the base line;
- the threshold cycle or CT, specific for each sample, identifies the value of the PCR cycle in which the exponential phase curve intersects the threshold line.

Most Real-Time PCR instruments are programmed to read the wavelengths of the SYBR Green emission and excitation spectrum (respectively 495nm and 537nm). This dye is very sensitive to light, it binds only to double-stranded DNA and therefore only to the newly synthesized amplicon. The samples are quantified on the basis of calibration curves obtained through the use of known quantities of 16S rDNA gene copies. The comparison between the signal emitted by the unknown sample with the fluorescence values used for the construction of the calibration curve allows the quantification of a specific microbial species. In addition to the quantitative measurement of target bacteria, intercalators such as SYBR Green allow to distinguish amplicons of different lengths and to detect non-specific amplifications that may be present.

Next Generation Sequencing: The peculiarity of this technology introduced in 2006 consists not only in the ability to sequence a single DNA fragment at a time, extending this process to millions of fragments at the same time, but also in the ability to sequence DNA fragments in both directions.

The first step involves single-stranded DNA fragments, at the ends of which univocal sequences, called "index", are loaded onto a flow of cells where they are captured on a surface containing "oligonucleotides still" complementary to the indexes, on which they are immobilized for the preparation of the libraries. The hybridization between the latter and the DNA fragments occurs through heating and cooling processes, followed by incubation with specific reagents and an isothermal polymerase. Through a "bridged" amplification each fragment is amplified distinctly from the others, creating a cluster of clones. When the cluster generation is complete, the generated models, after appropriate denaturation, are ready for actual sequencing.

Illumina uses a technology based on chain terminating fluorescent nucleotides with an OH at 3'; this ensures that a single base per cycle is incorporated. An imaging step follows to identify the



2019-1-BG01-KA203-062371

nucleotide incorporated in each cluster and a chemical step to remove the fluorescent group and terminal OH to allow the incorporation of another base in the next cycle.

At the end of the sequencing, which takes about 4 days, the sequence of each cluster is subjected to selection processes (trimming) to eliminate the low-quality products. During the data analysis the fragments of various lengths are aligned and superimposed, in this way it is possible to identify the sequence of the starting filament. In a standard procedure, at least 40-50 million sequences are analysed.

The shotgun sequencing is the sequence of all the genomes present in a complex matrix, such as a soil sample. Shotgun sequencing is therefore the most efficient way to sequence a large piece of DNA. For this, the starting DNA is broken up randomly into many smaller pieces, sort of in a shotgun fashion, with each of those pieces then sequenced individually. The resulting sequence reads generated from the different pieces are then analyzed by a computer program, looking for stretches of sequence from different reads that are identical with one another. When identical regions are identified, they are overlapped with one another, allowing the two sequence reads to be stitched together. This computer process is repeated over and over and over again, eventually yielding the complete sequence of the starting piece of DNA. The initial random fragmenting and reading of the DNA gave this approach the name "shotgun sequencing".

Microbial technologies for honeybee's health

Numerous biotic and abiotic stresses, such as the massive use of pesticides in agriculture and climate change, are compromising the survival of pollinating insects, with potentially harmful consequences on both agroecosystems and natural systems. In fact, bees are responsible for the pollination of 84% of cultivated plant species, 35% of which are of global importance and 78% of wild ones. Suffice it to say that 70% of seed crops alone (such as carrots, onions, garlic, etc.) are strictly dependent on insect pollination, as well as 80% of the 264 crop species of interest in Europe. From this, it follows that the activity of pollinating insects, including bees, plays an essential role at an economic level whose monetary estimate is about € 15 billion / year in Europe alone, while at worldwide the estimate grows to 153 billion €/year.

In addition to an incalculable value for the maintenance of biodiversity and balances present in the various ecosystems, then, bees supply honey, beeswax, propolis, pollen and royal jelly: in Europe, the data collected in 2010 showed a production of about 220 000 tons of honey with prices ranging from 1.50 to 40 €/kg depending on the area of origin. Or, in Australia the production of honey and beeswax annually hovers around a commercial value of \$ 90 million, underlining again the importance that the beekeeping sector plays in the panorama. world economy. Honeybee colonies have declined rapidly from 6 million in the 1940s to about 2.6 million today. High annual honeybee colony loss is still observed and has become the norm for beekeepers. Gut health plays a significant role in innate host immune response and adaptability to the multitude of stressors honeybees face today.

2019-1-BG01-KA203-062371

Broods affected by "Colony Collapse Disorder" (CCD), or "Hive Depopulation Syndrome" also showed significant signs of imbalance. The causes of this syndrome are not yet clear, but it is thought that they may be attributable to changes in environmental factors, malnutrition, the presence of pathogens and the massive use of insecticides. The symptomatology sees the presence of broods that abandon their larvae despite the presence of the queen, and lack of appetite for pollen and nectar stocks that are not consumed immediately (Fig. 13).

One of the possible causes of this die-off may be related to gut microbiota dysbiosis, as microbial alteration in terms of quantity and composition. With this term we indicate the phenomenon that negatively affects the beneficial functions of the microbiota and that are associated with specific metabolic imbalances. In fact, these deficiencies could create serious problems on the development of young adults by affecting their ability to develop resistance genes, including those for the synthesis of vitellogenin, and by inhibiting the functions of the immune system, given the evidence that the same microbiota promotes its effectiveness (Fig. 14). Few years ago, we tried to understand which factors are able to destabilize the microbiota, arriving at the conclusion that dysbiosis is caused by both biotic and abiotic factors. Considering biotic stresses, it has been seen that diet, the presence of specific pathogens and

Honeybee decline (CCD): a serious problem for honeybee health

Possible causes

- 🐝 Pests (e.g. Varroa, Vespa velutina...) and pathogens (e.g. Bacteria, microsporidia...) and viruses (e.g. Deformed Wing Virus, Chronic Bee Paralysis Virus...);
- 🐝 Poor nutrition (e.g., due to loss of foraging habitat and increased reliance on supplemental diets);
- 🐝 Pesticide and agrochemicals exposure;
- 🐝 Bee management practices (e.g., long migratory routes to support pollination services);

What is CCD?

➡ It is the acronym of "Colony Collapse Disorder". CCD is the phenomenon that occurs when the majority of worker bees in a colony disappear and leave behind a queen, plenty of food and a few nurse bees to care for the remaining immature bees and the queen.



Click for watching

Fig. 13. Insights on Colony Collapse Disorder

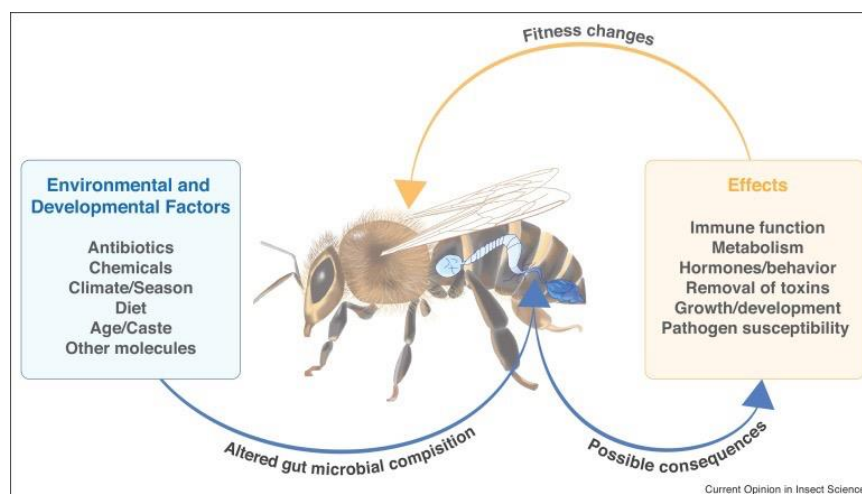


Fig. 14. Consequences of altered gut microbial compositions in bees



2019-1-BG01-KA203-062371

disorders (e.g., CCD) and adverse environmental conditions play a fundamental role. The lack of nutrients has a destructive impact on the normal development of the intestinal microbial flora, the consequence of which is to increase honeybee's mortality, as well as increase the susceptibility to diseases and pathogens. Furthermore, the anomalous temperatures induce a state of stress in the hosts such as to have dramatic repercussions on the symbionts.

Considering abiotic stressors, the damage is almost entirely attributable to the use of insecticides, fungicides, acaricides and antibiotics. The bees, in fact, during foraging operations risk to ingest indirectly and to encounter the active ingredients, both on the main crops treated and on the neighboring ones subjected to drift. This could cause serious problems and imbalances in metabolism and immune defenses: in fact, there is the possibility that exposure to certain substances interferes with the ability of bees to regulate their microbial gut population.

For those reasons, one of the most innovative future prospect aim to understand the in-depth relationship between microorganisms and honeybees, in order to improve their dramatic lifespan and conditions.

Gut disbiosis: an example

Given the growing interest that public opinion is showing towards this product, the first objective analysis falls on the effect of Glyphosate (N-phosphonomethyl-glycine). It is a non-selective post-emergence systemic herbicide, therefore a total herbicide. Its mechanism of action interrupts the metabolic pathway responsible for the synthesis of phenylalanine, tyrosine and tryptophan, inhibiting the synthesis of 3-phosphoshikimate-1-carboxyvinyltransferase (EPSP synthase). This herbicide has always been seen as one of the least toxic products for animals, as they lack this metabolic pathway. Despite this, it has been shown that it can affect non-target organisms showing highly toxic effects towards earthworms, microalgae, aquatic bacteria, rhizosphere, and endophytes. Affecting the bacteria, then, it should be emphasized that the effects were also detected in intestinal microorganisms and symbionts of the fauna adjacent to agricultural areas, including bees.

Specifically, Motta et al. (2018) conducted a study aimed at characterizing the microbiota of bees exposed to Glyphosate, concluding that the absolute abundance of *S. alvi*, *G. apicola*, *Lactobacillus* sp. and *Bifidobacterium* sp. (Fig. 15) has undergone a significant decrease.

2019-1-BG01-KA203-062371

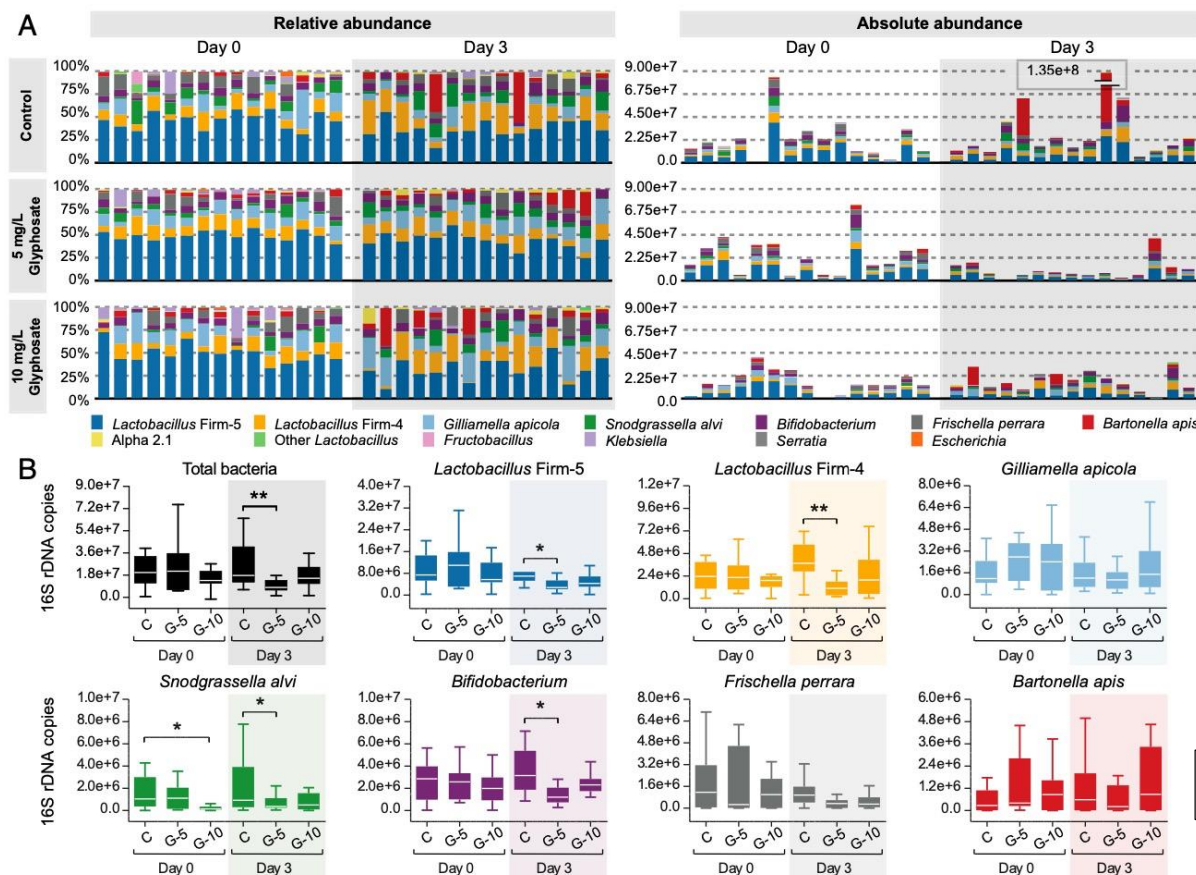


Fig. 15. Analyses on the bee gut microbiota

The product compromised the bacterial flora by stopping their growth without, however, directly killing them; it was therefore hypothesized that the effect fell on cell division during the early days of colonization. The bees that encountered the herbicide in the field, in fact, would have carried the active ingredient inside the hive which, being very stable and insoluble in water, would have been able to remain on the surfaces for a long time. Similarly, even in the field, persistence means that contamination can last for a long time. Inside the hive, therefore, the diffusion by trophallaxis and contact with other bees means that the product reaches the young larvae fed by adults, irreparably altering the development of beneficial symbiont species.

2019-1-BG01-KA203-062371

How is the microbiota acquired?

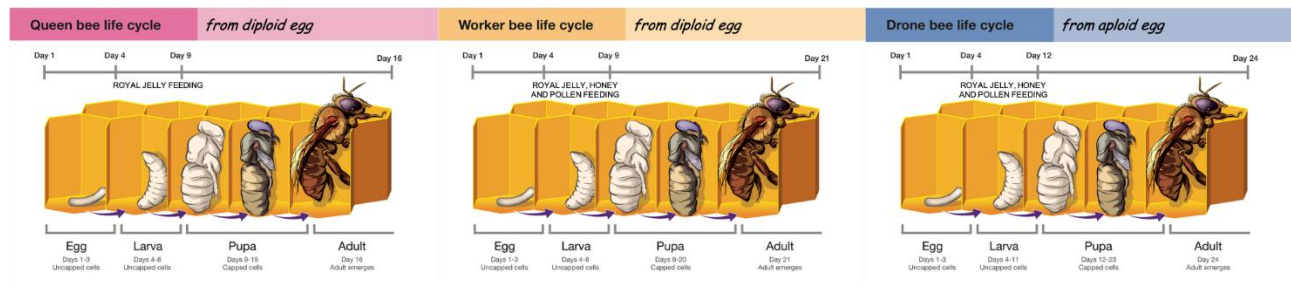


Fig. 16. The growth cycle of a bee

Natural bacteria picked up by honeybees from flowers while collecting nectar and pollen reside predominantly in honeybee midgut and hindgut. Gut bacteria naturally found in honeybees are dynamic. During development of the larvae (Fig. 16), bacterial population fluctuates. Larvae receive some bacteria from the nurse bees feeding them. During pupation, the gut lining is shed, and the gut of a newly emerging adult honeybee is sterile. The gut is quickly repopulated with characteristic microbiota. How does this happen? Main pathways are oral trophallaxis, interaction with hive material, and fecal-oral transmission. In particular, the characteristic microbiota of adult bees begins to develop about four days after the flicker.

Although the factors that allowed the evolution of the microbiota for every living being are still unknown, it is proven that social bees possess a distinctive microflora depending on the family they belong to. For example, it has been seen that analyzing the microbiota of different genera of eusocial corbiculate Apoidea such as *Bombus* spp., *Megachile* spp. and *Apis* spp., the main bacterial genera were recurrent (*Snodgrassella* spp., *Gilliamella* spp., *Bifidobacterium* spp. and *Lactobacillus* spp.) but the species varied in relation to the insect species. Hence the hypothesis that sees the microbiota as the result of a dynamic co-evolution between microorganisms and hosts, dependent on the environment and on the genotypic variations to which the species have been subjected over the centuries, whose richness is also correlated with the size of individual bees and entire colonies. In fact, the establishment of a species-specific microbial flora is the result of a long selection in which the optimal beneficial ratios have been established both for the microorganisms and for the hosts.

Who and where are they?

It has been estimated that within the intestines of adult worker bees there are about 1 billion bacterial cells, 95% of which are located, specifically, in the hindgut (Fig. 17). Here, a specific differentiation between ileus and rectum was noted; in the first three species of Proteobacteria such as *G. apicola*, *F. perrara* and *S. alvi* which form a dense biofilm in correspondence with the Malpighian tubes and which continues along the length of the ileum wall. In the rectum, however, a dense bacterial

2019-1-BG01-KA203-062371

community prevails formed by three classes of Gram positive, such as Firmicutes (Firm-4, Firm- 5) and *Bifidobacteria*. As for the midgut, it has been seen that there are mainly *Lactobacillus* spp. and *Acetobacteraceae*, that is, those taxa that are also found in pollen, nectar and more generally within the hive. It can therefore be said that a pre-adapted microbiota does not exist in the midgut and that it varies in relation to the environment and the individual's eating habits. In quantitative terms, then, here there is a much less abundant flora than in the rectum. The midgut also contains few bacteria, and those presents are more concentrated in the pro-ventricular area adjacent to the hindgut.

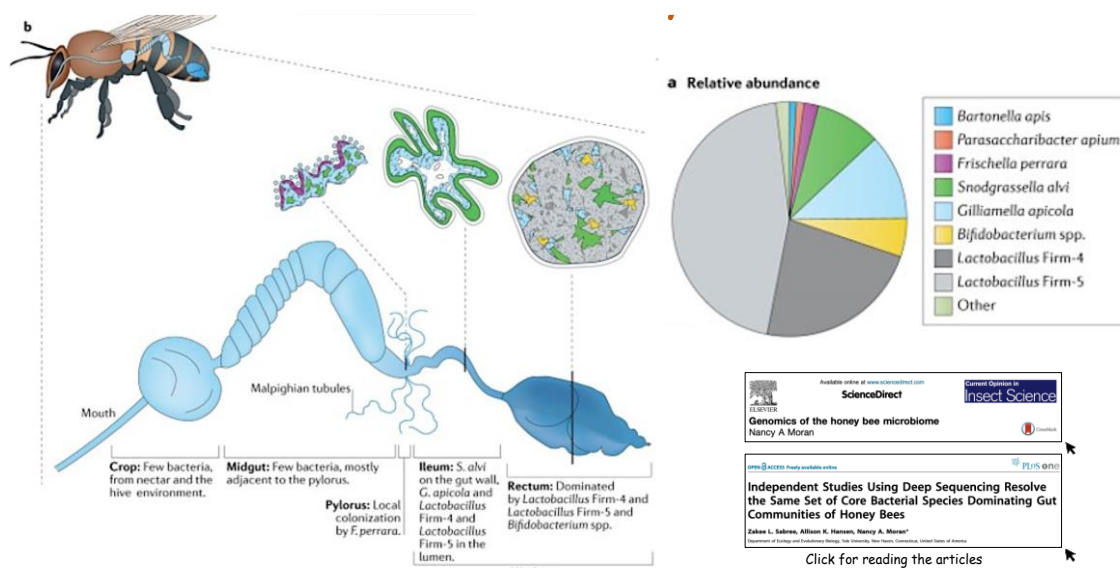


Fig. 17. The core bee microbiota

What are the gut microbiota functions?

In the last few years, the scientific community has begun to take an increasing interest in the role that the microbiota plays in the wellbeing of honeybees (Fig. 18). Numerous studies and research have shown that interactions with the host have supportive effects both at a metabolic and nutritional level and in terms of immune response to pathogens. As for food support, a balanced bacterial flora is necessary for a correct assimilation of nutrients as, thanks to its enzymatic activity, it participates in the degradation of complex sugars. In addition to being responsible for the presence of cellulases, hemicellulases and ligninolytic enzymes in the intestine useful for the digestion process of pollen grains, the richness of species is such as to allow the coexistence of different sugar catalysis pathways (especially for *Gammaproteobacteria*, *Firmicutes* and *Bifidobacteriaceae*). Indeed, it has been estimated that 91% of the protein transcripts linked to the digestion of plant macromolecules and to the fermentation phenomena of monomeric subunits are produced by bacteria. Another concrete example



2019-1-BG01-KA203-062371

concerns pectin-lyases capable of degrading the pectins present in the cells of the wall of pollen grains. The latter is also an excellent indicator of the high genetic variability and adaptability within the same species of microorganisms. In fact, it has been seen that only some strains of *G. apicola* possess them while others are completely devoid of them. The importance of a greater digestive capacity and, consequently, the ability to metabolize nutrients that could not be demolished, has been demonstrated in various studies. Zheng et al. (2017), for example, by comparing bees with a normal microbiota and others without any type of intestinal flora, they highlighted appreciable physiological differences. In the former, the symbionts positively influenced the size of the intestine, the weight of individuals, the values of vitellogenin and insulin and the sensitivity to sugars. These findings then suggested that the microbiota could influence the appetite and growth of the bees' body through the increase of signals related to the presence of insulin.

In addition to the hydrolysis of complex carbohydrates, intestinal microorganisms produce useful metabolic substrates, such as vitamin B and other and short chains of fatty acids. For example, it has been seen that the genera *Lactobacillus* sp. and *Bifidobacterium* sp. are involved in the processes of fermentation of pollen and nectar so that they can be considered responsible for the vitaminic value of honey. The phenomenon of symbiosis with the host goes beyond nutritional and metabolic support: the microbiota plays an important role in supporting the immune system. In fact, in the first place, the bacteria could directly stimulate the production of the bee's own defense molecules. Following the contact between the epithelial surface and the peptidoglycan (major component of the cell wall of Gram-positive bacteria), the immune system could activate the genes to produce 6 antimicrobial peptides such as: abaecin, hymenoptaecin, apidicin, defensin-1 and defensin-2. The production of these compounds, then, is accentuated by the alterations of the microbial membranes themselves and can also be induced by exposure to some pathogenic and non-pathogenic microorganisms. For example, *Frischella perrara*, a symbiont that colonizes the ileum region in the hindgut, above all stimulates the production of apidicin. Secondly, the microbiota may be directly responsible to produce antimicrobial compounds, which, among other things, is confirmed by numerous studies. Saraiva et al. (2015), for example, documented the presence of numerous genes involved in the biosynthesis of streptomycin and of secondary metabolites expressed by symbionts and which may play a role in maintaining the microbiota.

2019-1-BG01-KA203-062371

Generic function of the honey bee gut microbiota	Specific function	Target microorganisms (where available)	Generic function of the honey bee gut microbiota	Specific function	Target microorganisms (where available)
Nutritional support	Source of vitamins, fatty acids, amino acids	-	Host protection: other strategies	Antimicrobial activity against <i>Paenibacillus larvae</i> , <i>Melissococcus plutonius</i> and <i>Ascosphaera apis</i>	<i>Bacillus</i> spp., <i>Lactobacillus</i> spp., <i>Bifidobacterium</i> spp.
	Lignin degradation	<i>Fructobacillus</i> spp.			
	Sugar uptake systems	<i>Gammaproteobacteria</i> , <i>Firmicutes</i> , <i>Bifidobacteriaceae</i>			
	Breakdown of plant-derived macromolecules	<i>Gammaproteobacteria</i> , <i>Firmicutes</i> , <i>Bifidobacteriaceae</i>		Biofilm formation and structures resembling extracellular polymeric substances	LAB symbionts from honey crop
	Pectin degradation activity (strain specific)	<i>G. apicola</i>		Biosynthesis of cell wall exopolysaccharides	"Firm4" and "Bifido" groups
	Aerobic oxidation of the end-products of the fermentation process	<i>S. alvi</i>		Genes encoding a relevant number of functions related to biofilm formation and host interaction (Type IV pili, outer membrane proteins, and secretion)	<i>G. apicola</i> and <i>S. alvi</i>
	Glycoside hydrolase activities	Lactic acid bacteria			
	Trehalose degradation IV pathway	<i>Bifidobacterium</i> spp.			
	Increased expression level of antimicrobial peptides (AMPs) under pathogen exposure in bee larvae	-			
	Increased expression level of selected AMPs in bee larvae upon feeding with probiotic bacteria	-			
Direct stimulation of the bee's immune system	Strong positive correlation between the amount of total honey bee gut bacteria and transcript levels AMPs	-			

Fig. 18. Functions of the bee gut microbiota

What can we do?

From our and general experience in humans and animals, biotic and abiotic stresses could negatively affect the composition of the gut microbiota and therefore induce specific changes in the microorganism activities at gut level.

We must ask ourselves if any kind of microbiota modulation, by the administration of selected strains, could restore this perturbation, reduce bee mortality and/or improve honeybee health.

Probiotics are "live microorganisms that, when administered in adequate quantities, bring a benefit to the health of the host, excluding references to biotherapeutic agents and beneficial microorganisms not used in food" (FAO / WHO, 2001). Their administration, then, must not be associated with negative effects on organisms and the environment. The mode of action of these microorganisms can be summarized in the following functions:

- PROTECTIVE FUNCTION: dislocation of pathogens, competition for nutrients, competition with receptors and production of antimicrobial molecules (eg bacteriocins, organic acids ...);
- STRUCTURAL FUNCTION: barrier effect, biofilm on the intestinal hair wall, development of the immune system;

2019-1-BG01-KA203-062371

- **METABOLIC FUNCTION:** differentiation and proliferation of intestinal epithelial cells, catalysis of carcinogenic substances present in the diet, synthesis of vitamins, fermentation of non-digestible sugars, ionic absorption, and energy saving.

However, it seems that the main mechanism of action of probiotics is the stimulation of the immune system: following cohesion with the intestinal wall, they are able to stimulate a series of cascade signals that activate the synthesis of antimicrobial peptides. In practice, they are able to carry out those tasks described previously that the bee's microbiota naturally performs, proving to be a hypothetical optimal aid in safeguarding and optimizing it, as well as improving the bee's life prospects.

As with human and animal nutrition, the main bacteria considered capable of these benefits are *Lactobacillus* spp., *Bifidobacterium* spp., *Bacillus* spp.

In conclusion, the use of probiotics has recently begun to be evaluated also within the hive itself. In fact, although bees prefer to consume fresh pollen, under certain conditions, such as seasonality, they need to supply themselves with the stored reserves. The humid environment (50-60% RH) that is created following the collection of pollen increases the risk of uncontrolled bacterial and, above all, fungal growth. There is a clear need to preserve the stocks inside the hive to avoid infections that could lead to fatal outcomes, such as calcified larvae due to *Ascosphaera apis*, or intestinal infections due to *Nosema* spp.

Overview of beneficial microorganism applications for the treatment of the main honey bee microbial infections

Honey bee disease	Infection dose	Microorganisms/metabolites	Source	Reported effect(s)	References
<i>P. larvae</i> - AFB	10 ³ and 10 ⁶ spores/ml	<i>L. kunkei</i> , <i>L. mellis</i> , <i>L. kimbladii</i> , <i>L. kullabergensis</i> , <i>L. helsingborgensis</i> , <i>L. melliventris</i> , <i>L. apis</i> , <i>L. mellifer</i> , <i>B. asteroides</i> and <i>B. coryneforme</i> (10 ⁷ bacteria/ml)	Honey crop	Reduced larvae mortality	Forsgren et al. 2010
	Not described	<i>B. thuringiensis</i> HD110, <i>B. laterosporus</i> BMG65.	Honey bee gut	Reduced larvae mortality	Hamdi and Daffonchio 2011
<i>M. plutonius</i> - EFB	10 ² -10 ⁶ -10 ⁸ bacteria/ml	<i>L. kunkei</i> , <i>L. mellis</i> , <i>L. kimbladii</i> , <i>L. kullabergensis</i> , <i>L. helsingborgensis</i> , <i>L. melliventris</i> , <i>L. apis</i> , <i>L. mellifer</i> , <i>B. asteroides</i> and <i>B. coryneforme</i> (10 ⁷ bacteria/ml)	Honey crop	Reduced larvae mortality	Vásquez et al. 2012
<i>N. ceranae</i>	1st trial: 10 ⁴ spores/μl 2nd trial: natural infection	<i>L. kunkei</i> Dan39, <i>L. plantarum</i> Dan91 and <i>L. johnsonii</i> Dan92, <i>B. asteroides</i> DSM 20431, <i>B. coryneforme</i> C155, <i>B. indicum</i> C449. (10 ⁶ -10 ⁷ cfu/ml of sugar syrup)	Honey bee gut	Reduced spore detection	Baffoni et al. 2016
<i>Nosema</i> spp.	10 ³ spores/μl	<i>P. apium</i> C6 (10 ⁶ cfu/500 μl)	2nd instar larvae	Reduced spore detection	Corby-Harris et al. 2014
Diseased bees		<i>L. johnsonii</i> CRL1647 (10 ⁵ cfu/ml)	Honey bee gut	Reduced spore detection	Audisio et al. 2015
Diseased bees		10 ⁵ spores/mL of <i>Bacillus subtilis</i> Mor12 spores	Honey	Reduced spore detection	Sabaté et al. 2012

Adapted from Alberoni et al. 2016



References

- EFSA (2009). Bee mortality and bee surveillance in Europe. CFP/EFSA/AMU/2008/02
- Motta EVS, Raymann K, Moran NA. 2018. Glyphosate perturbs the gut microbiota of honeybees. *Proc Natl Acad Sci USA*, pp. 1-6
- Zheng H, Powell JE, Steele MI, Dietrich C, Moran NA. 2017. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc Natl Acad Sci USA*, 114: 4775-4780
- Saraiva MA, Zemolin APP, Franco JL, Boldo JT, et al. 2015. Relationship between honeybee nutrition and their microbial communities. *Antoine Van Leeuwenhoek*, 107: 921-933
- FAO/WHO (2001). Health and nutritional properties of probiotics in food including powder milk with live lactic acid bacteria. Food and Agriculture Organization of the United States, World Health Organization
- Kainthola J, Kalamdhad AS, Goud VV. 2019. A review on enhanced biogas production from anaerobic digestion of lignocellulosic biomass by different enhancement techniques. *Process Biochemistry*, 84: 81-90
- Shamurad B, Sallis P, Petropoulos E, Tabraiz S, Ospina C, Leary P, et al. 2020. Stable biogas production from single-stage anaerobic digestion of food waste. *Applied Energy*, 263: 114609
- Kumar M, Dutta S, You S, Luo G, Zhang S, Show PL, et al. 2021. A critical review on biochar for enhancing biogas production from anaerobic digestion of food waste and sludge. *Journal of Cleaner Production*, 127143
- Azhar, S. H. M., Abdulla, R., Jambo, S. A., Marbawi, H., Gansau, J. A., Faik, A. A. M., & Rodrigues, K. F. (2017). Yeasts in sustainable bioethanol production: A review. *Biochemistry and Biophysics Reports*, 10, 52-61.
- Galbe M, Zacchi G. 2007. Pretreatment of lignocellulosic materials for efficient bioethanol production. *Biofuels*, 41-65.
- Kim S, Dale BE. 2004. Global potential bioethanol production from wasted crops and crop residues. *Biomass and bioenergy*, 26(4): 361-375.



2019-1-BG01-KA203-062371

Vilpoux O, Averous L. 2004. Starch-based plastics. Technology, use and potentialities of Latin American starchy tubers, 521-553.

Sin LT. 2012. Polylactic acid: PLA biopolymer technology and applications. William Andrew.

Koh JJ, Zhang X, He C. 2018. Fully biodegradable Poly (lactic acid)/Starch blends: A review of toughening strategies. International journal of biological macromolecules, 109: 99-113.

Poltronieri P, Kumar P. 2017. Polyhydroxyalkanoates (PHAs) in industrial applications. Handbook of Ecomaterials. Cham: Springer International Publishing, 1-30.



Project website: www.digit-biotech.eu

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.