



Biotechnology & Microbiology

POSTGRADUATE COURSE

“Green Technology Foresight about challenges from biotechnology and ICT”

Digit-BioTech

2019-1-BG01-KA203-062371

DIGIT~Bio~TECH



L01 SYSTEMS BIOLOGY AND OMICS TECHNOLOGY

Advanced level

AUTHOR:

VENTSISLAVA PETROVA



Contents

Contents	2
OMICS technologies towards improving the quality of life	3
OMICS Technologies and Molecular Medicine	3
OMICs technologies in diagnostics.....	4
OMICs analysis for revealing genetic architecture of common disease	9
Application of OMICs approach in disease treatment	10
OMICs technologies for disease prevention - future prospects.....	11
Application of OMICS Technologies in Toxicology	11
Toxicological studies via OMICs analysis.....	12
OMICs applications for toxicology risk assessment and regulatory submissions	16
Environmental OMICS and Biodiversity	17
Single-cell genomics	18
Omics Approaches in Industrial Biotechnology and Bioprocess Engineering	23
Omics-Guided Biotechnology.....	23
Nutri-omics Research.....	28
Genomics and nutrition	28
Transcriptomics and nutrition	30
Proteomics and nutrition	30
Metabolomics and nutrition	31
Nutrition and other OMICs	31
References	32



OMICS technologies towards improving the quality of life

OMICS TECHNOLOGIES AND MOLECULAR MEDICINE

Advances in omics technologies - such as genomics, transcriptomics, proteomics and metabolomics - have begun to enable medicine at an extraordinarily detailed molecular level. The rapidly decreasing costs of high-throughput sequencing and other massively parallel technologies, such as mass spectrometry, are enabling their use in clinical research and clinical practice. Exome and genome sequencing are already being used to aid diagnoses, particularly of rare diseases, to inform cancer treatment and progression and, in early efforts, to create predictive models of disease in healthy individuals. Numerous research efforts and companies are focusing on genome-wide profiles of genetic, gene expression and other omics data, such as the microbiome, as biomarkers for disease. These techniques have also been applied in identifying risk loci for disease, defining precise pathophysiology of the disease, as well as performing research in Occupational Environmental Health (OEH).

Ideally, different technologies would be combined both to help diagnose disease and to create a holistic picture of human phenotypes and disease. However, implementation of multi-omics data introduces new informatics and interpretation challenges. What are the ways in which integrative omics can impact medicine by helping to manage health, as well as diagnose and treat disease? In fact, to improve the treatment outcomes and reduce adverse events that matter to both the clinician and patient one of the perspective solutions is to develop the personalized medicine - customized medical treatment created to fit individuals' characteristics, needs and unique molecular and genetic profile (Fig. 1).

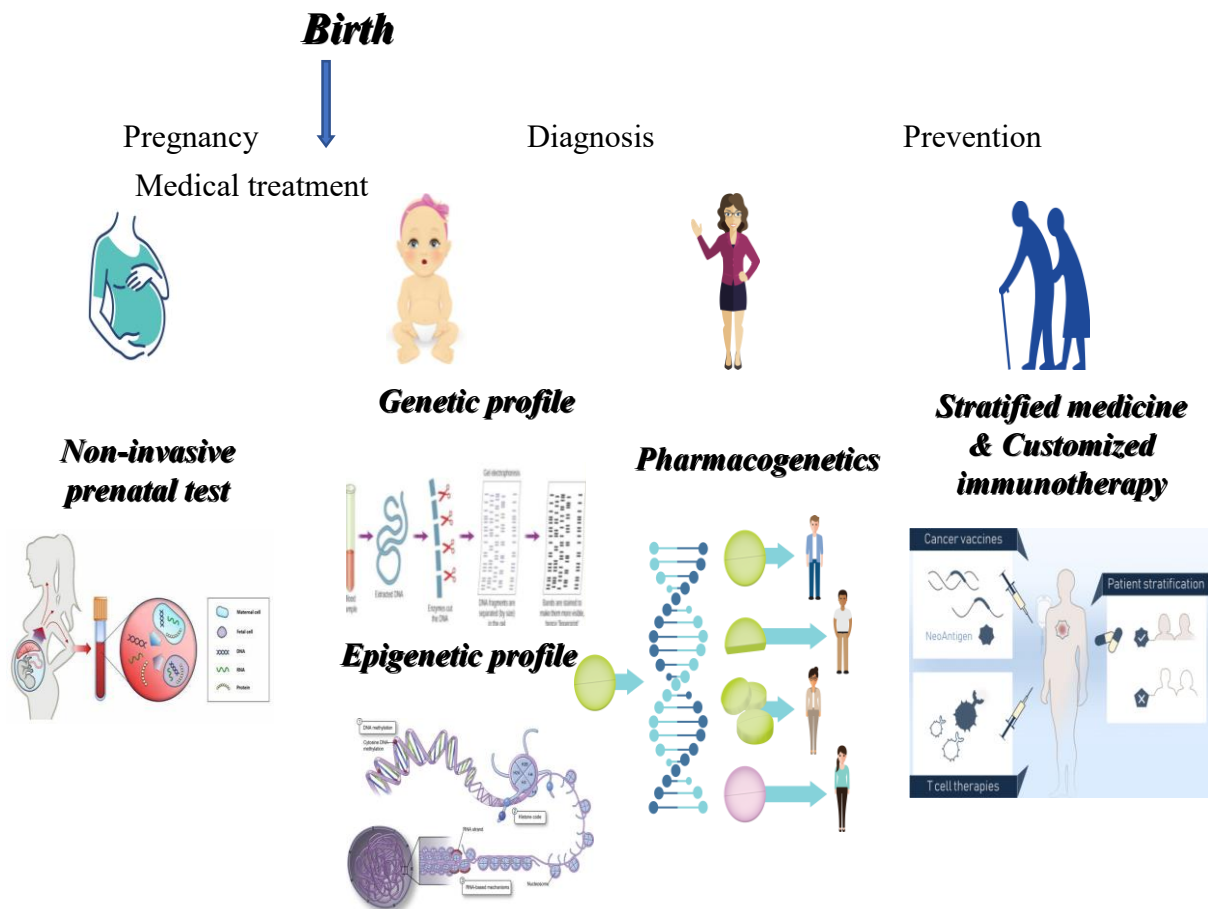


Fig. 1. Personalised medicine throughout the lifespan

1. OMICs technologies in diagnostics

PRENATAL DIAGNOSTICS

Nearly 10% of pediatric diseases and 20% of infant deaths are due to Mendelian diseases. Genetic defects can be a major threat to the baby's health. To overcome this issue, last developments in the omics technologies are employed in a person's life even before birth. Prenatal screening has become widespread



2019-1-BG01-KA203-062371

in many countries for examining not only abnormalities of a fetus but also the possible risks of diseases or abnormal conditions after birth. When a genetic predisposition exists, such as cystic fibrosis or muscular dystrophy, consulting a medical specialist or genetic diagnosis company become of great importance. The latest innovations in molecular and cellular biology allowed various genetic diagnoses for fetuses. Till recently, different invasive technologies have been applied such as amniocentesis and chorionic villus sampling. Although the procedures have been improved, there is still risk of miscarriage and serious side effects. One of the innovative attempts to avoid the side effects of conventional invasive technologies is to isolate and analyze fetal nucleated cells from maternal blood. Unfortunately, the number of fetal cells that can be isolated is about 3-5 cells per 30 ml of maternal blood. The huge jump in the field of non-invasive prenatal testing (NIPT) has been made after the advancement of omics technologies and their combination with various experimental approaches. The use of DNA amplification technology combined with NGS allowed the direct detection of the fetal DNA in maternal blood. Currently, such approach is used to determine the sex and Rh blood type, and to analyze chromosomal aneuploidies. The cell-free fetal DNA is typically released in maternal blood during the process of apoptosis of placental cells. The fragmented cell-free fetal DNA is ~200 bp long and has a half-life of ~16 min and can be detected in maternal plasma from early pregnancy (5-7 weeks). In addition to fetal DNA, fetal RNAs can also be used for NIPT. For example, fetal trisomy 21 can be diagnosed by checking the allelic ratio of *PLAC4* mRNA, produced from chromosome 21 in the placenta. If a baby receives heterogeneous alleles from the parents, an unequal allelic ratio (2 : 1) suggests trisomy.

Integrating fetal genomic, transcriptomic, proteomic, metabolomic, and epigenetic data will help pushing forward the prenatal genetic diagnostics. In this way, any genetic disorder could be identified early in the pregnancy by applying non-invasive approaches. However, some concerns occur regarding raise of serious ethical issues possibilities when advanced genomic technology are used. Therefore, further studies on the impact of new technologies on society, such as possible discrimination based on genetic information, undetermined harmful effects of nanotechnologies on humans and nature, and copyright infringements due to the development of information technology should be considered (see LO8).

LAB-ON-A-CHIP TECHNOLOGY IN DIAGNOSTICS

Rapidly after the development of micro-electro-mechanical systems (MEMS), the potential use of these miniaturized platforms for various applications has been revealed. During the past few decades, interest in biological or biomedical MEMS (BioMEMS) has been significantly increased and it has found widespread applications in a various area life science including diagnostics, therapeutics, drug delivery, biosensors and tissue engineering. These integrated systems are also known as “lab-on-a-chip” (LOC) or “micro-total analysis systems” (μ TAS) and provide a solid way for miniaturization, integration, automation, and parallelization of chemical processes.



Lab-on-a-chip devices integrate multiple laboratory functions on a single chip (from few square millimeters to a few square centimeters in size). These platforms provide complex chemical and/or biological analyses which can offer cheaper, faster, controllable and higher performance of bio-chemical assays at a very small scale when compared with conventional laboratory tests. They are capable of handling extremely small fluid volumes down to less than a few picoliters (few micro-droplets of whole blood, plasma, saliva, tear, urine or sweat). This fact is very important in many clinical trials where usually very small volumes of patient samples are available. Moreover, the automation with eliminating the human interfering parameters can increase the reliability in the analysis.

LOC has been used in different applications, such as: DNA extraction and purification, PCR, qPCR, molecular detection, electrophoresis, etc.

LOC for DNA Extraction and Purification: Often for diagnostic assays nucleic acid have to be purified from cells. The process of DNA extraction generally includes two steps: cell lysis by disrupting cellular and nucleus membrane, and DNA purification from membrane lipids, proteins, and RNA. Cell lysis on LOC devices can be achieved through several types of lysis methods: chemical, thermal, ultrasonic, electrical, mechanical or some other type of lysis. The well known DNA purification methods through extraction columns, using the DNA adsorption on silica beads under certain buffer conditions, are also adaptable to microfluidic techniques.

LOC Devices for PCR, qPCR, and Molecular Detection: After DNA extraction, the most frequent analysis is the PCR. The importance of PCR in genomic analysis led to the development of numerous LOC devices. Miniaturization of volume and the high surface to volume ratio defines rapid thermal transfer and more accurate PCR. This analysis requires a post-analysis and amplicons' size detection, which generally is carried out by electrophoresis. With the introduction of LOCs, DNA electrophoresis was one of the first molecular processes that was integrated on a chip. This miniaturization diminishes even more the overall process time and reduces the reagent consumption.

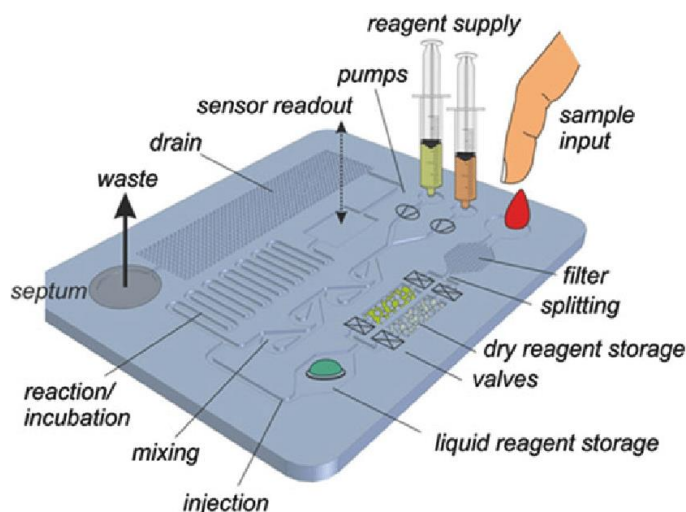


Fig. 2. Lab-on-a-Chip device

qPCR is another technique that was modified to LOC devices. Using ultra-fast pressure controller and fluorescence reader, ultra-fast qPCR microfluidic system had been developed for the molecular detection of diseases like Anthrax and Ebola. The detection efficiency is identical to commercial systems and results are achievable in less than 8 minutes being 7-15 times faster than traditionally used ones. Other advanced field is digital microfluidics that deals with emulsion and droplets within LOC devices. Ultra-low quantity of DNA can be captured within droplets, and detection limits can be enhanced with one copy number detection within LOC droplet qPCR.

Genomic Application: Due to the numerous sequencing projects, such as the Human Genome Project, the field of genomics gained a tremendous attention in recent years. Next-generation sequencing techniques of DNA which are currently under development include sequencing-by-hybridization (SBH), nanopore sequencing, and sequencing-by synthesis (SBS), the latter of which includes many different DNA polymerase dependent strategies. Using LOC approach allows the drastic reduction in cost and duration of high-throughput sequencing.

When studying the complex DNA samples expression an integration of multiple biosensors in connection with DNA microarrays is required. The most essential characteristics of these LOCs are miniaturization, speed, and precision. This technology offers vast possibilities for rapid multiplex analysis of nucleic acid samples, including the diagnosis of genetic diseases, detection of infectious agents, measurements of differential gene expression, drug screening, or forensic analysis.

Biochemical Applications: LOC strategy is also applicable in coupling enzymatic and immunological assays on a single-channel microfluidic device. Typical example is the simultaneous measurement of glucose and insulin. Key for the effective realization of such glucose/insulin monitoring is the integration of relevant sample pretreatment/cleanup procedures essential for whole-blood analysis.



2019-1-BG01-KA203-062371

This innovative LOC approach can be applicable to the incorporation of other analytes and additional sample handling steps, as is desired for creating miniaturized clinical instruments.

Proteomics: Proteomics is one of the great scientific challenges in the post-genome era. LOC devices are useful for creating advanced techniques to answer complex analytical problems, such as proteome profiling. LOC devices could be applied in four key areas of proteomics: chemical processing, sample preconcentration and cleanup, chemical separations, and interfaces with mass MS. LOC being a miniaturized device could be used for separation and detection of proteins. This process is characterized with less reagent consumption, easy operation, and very fast analysis. Another benefit is that the data from the LOC devices can easily be compared to those obtained using 2D-PAGE. In order to quantitate proteins a LOC MS protein profiling device is developed. Some authors describe LOC devices that are used for direct infusion into the mass spectrometer, including Capillary Electrophoresis (CE) separation, on-chip sample digestion, and infusion or CE before MS analysis. Researchers use electrospray ionization MS and focus on the interface design between LOC and the mass spectrometer.

Biosensors: For detecting target analytes LOC biosensors are also developed. Such devices intricately link a biological recognition element with a physical transducer converting the biorecognition event into an electrical signal. There are two types of LOC biosensors: bioaffinity and biocatalytic devices. Bioaffinity devices are based on selective binding of the target analyte to a surface-enclosed ligand partner (e.g., antibody, oligonucleotide). For example, in hybridization biosensors, there are immobilized single-stranded (ss) DNA probes onto the transducer surface. When a duplex is formed it can be detected following the association of an appropriate hybridization indicator. On the other hand, in biocatalytic devices, an immobilized enzyme is used for recognizing the target substrate. For example, sensor strips with immobilized glucose oxidase for monitoring of diabetes (for more information see LO3).

Cell Research: Using LOC devices a small-scale experiment with cells could also be performed. The single cell is positioned within the microchannel in a predetermined location, making its handling and manipulation easy. The microsystem allows handling small volumes of fluid containing small quantities of analyte. LOC systems also provide precise control of the local environment around the cell, monitoring the physical or chemical nature of the surface to which the cell adheres, the local pH, and temperature. When the cell is to be exposed to a different type of stimuli, numerous and often complex fluid-handling components can be used. If the cell is a subject of further analysis, lysis and necessary analytical methods could also be introduced onto the same platform. Thus, sample loss or dilution is avoided, which would inevitably occur if multiple devices were used.

Drug Development: The industry needs for development of new drugs and to predict their potential behavior in animals and cells triggers other analytical applications of LOC systems: e.g., analytical systems to monitor and optimize the production of protein drugs such as therapeutic antibodies; assays based on primary human cells that could predict performance in human clinical trials and others. These LOC systems are proven to be highly reproducible, easily manipulated, and could be used routinely.



2. OMICs analysis for revealing genetic architecture of common disease

Identifying genetic markers that are associated with a specific disease feature is key issue when estimating the disease risk. As the genetic information of an individual remains mostly unchanged, the mutated sequences can be reliable markers for certain diseases. In this regard one of the important applications of omics technology is the genome-wide association studies (GWASs). They can identify common genetic variants, generally single nucleotide polymorphisms (SNPs), associated with certain disease features. In general, SNP information is collected from two groups, a patient group and a healthy control group. Then it is statistically analyzed, and the SNPs associated with the specific disease characteristics of the corresponding patient group re identified.

For example, variations in the apolipoprotein E (APOE) gene are a well-known genetic factor that is correlated with the late-onset Alzheimer's disease (LOAD). After the assessment of 502,627 SNPs in 1086 AD patients, it was demonstrated that rs4420638 on chromosome 19 is the strongest marker differentiating AD patients from the control group. Other important success is achieved during investigation of hereditary cancers. BRCA1 and BRCA2 genes are one of the most important genetic markers for hereditary breast and ovarian cancers. These genes encode tumor suppressor proteins, which help repairing damaged DNA through homology directed repair. Some mutations in BRCA1 and/or BRCA2 significantly raise the risk of developing breast and/or ovarian cancers. For BRCA1, two mutations were detected: 185delAG and 5382insC. Approximately 55-65% of women who inherit one of the harmful BRCA1 mutations and around 45% of women who carry the BRCA2 mutation have developed breast cancer by the age of 70. Numerous other disease markers such as autosomal-dominant colorectal cancer, Li-Fraumeni multi-cancer syndrome, mucopolysaccharidosis IV, and hypertrophic cardiomyopathy, have also been successfully identified using NGS.

However, for the most common diseases such as diabetes, obesity, schizophrenia and autism, a combination of multiple genetic and environmental factors is responsible. A growing number of evidence indicates that genetic markers commonly explain only a part of the total trait variation. One of the missing crucial factors are the epigenetic variations. Various epigenetic modifications have been reported to be associated with the early stage of different cancers (abnormal growth of a tumor). Moreover, it is well known that colorectal cancer is more frequently identified in men than in women. That is why it has been suggested that estrogen may help suppress colorectal cancer. In addition, the epigenetic profiles of adjacent mucosa, and non-adjacent mucosa samples of 95 colorectal cancer patients have revealed that the promoter of the MGMT, a DNA repair gene, was frequently methylated. Up to now, thousands of genomic loci have been identified and associated with various human diseases. However, the difficulties arise when these genes have to be characterized in the context of the molecular pathophysiology and the corresponding interacting genes and pathways.



3. Application of OMICs approach in disease treatment

The main aim of personalized medicine is to decide on the most effective treatment option for an individual. In this regard, the simultaneous quantification of multiple protein markers can play an important role in subtyping tumors and selecting optimized therapies for each subtype. High-throughput, single cell-based protein quantification techniques, such as CyTOF, can be excellent alternatives for immunohistochemistry-based methods.

For example, in the last decade a new approach for treating cancers has emerged - allowing the patient's immune system combat his or her cancer, called adoptive cell transfer (ACT). Glioblastoma multiforme (GBM) is the most frequent and aggressive type of malignant brain tumor. Although GBM has been treated with surgical removal, radiation treatment, and chemotherapy methods, there have been no noticeable effects. Over the past 20 years, numerous clinical trials have been performed to use the immune system to cure cancers such as GBM. Among numerous ACT approaches, particular attention has been paid on the dendritic cell (DC)- and T cell-based vaccines. They can penetrate and induce anti-tumoral response in the brain and show great potential for effectively killing various cancers including the malignant and advanced types. Autologous DCs, obtained from tumor tissue or in the form of a tumor lysate of a patient, can be grown *in vitro* in the presence of tumor-associated antigens (TAAs), a method called 'pulsing'. The pulsed, autologous DCs are reintroduced into the patient to stimulate specific, cell-mediated, anti-tumoral cytotoxicity against GBM and other malignant cancers. More revolutionary approaches, e.g. engineering specific immune cells using zinc-finger nucleases (ZFNs) and CRISPR technology, have also been attempted.

Every year, thousands of patients undergo a transplantation of organ or hematopoietic stem cells. However, mortality among such patients remains very high. A standard procedure for matching donors with recipients involves human leukocyte antigen (HLA) typing. However, it becomes clear that non-HLA factors can also considerably affect prognosis and development of graft-versus-host disease (GVHD). In order to overcome this, many omics applications can be used to determine optimal donor–recipient matches, as well as to examine different markers of rejection. For example, free DNA sequencing can be used to assess the severity of organ rejection as well as to simultaneously detect the eventual presence of viral DNA as a marker of infection. Additional omics data, such as RNA or protein expression, may also be utilized to evaluate the donor–recipient compatibility. Thus, integration between several omics' technologies may be used as a useful tool for transplant biology.

Furthermore, the microbiome has been associated with many common human diseases. Whereas causality is simple in genomic data, where DNA influences phenotypes, it is more difficult to unravel whether microbiome composition is a cause or effect of disease. Nevertheless, it is evident that patients with inflammatory bowel disease, type 2 diabetes and obesity, have different microbiome profiles from those of healthy controls. In addition, the microbiome has a great influence on immune function, which in some cases has been putatively causally related to disease in animal models. As our understanding of the microbiome advances, integrative assessment of this and other omics technologies data will help for our better understanding of human disease. Recently, it has been shown that human genetic profiles



2019-1-BG01-KA203-062371

influence overall gut microbiota composition. Additionally, interactions between human genetics and microbiomes have been shown to influence disease manifestation. Similarly, metabolic signaling between hosts and their microbiomes has become an area of active research, and there is increasing evidence that a number of metabolites secreted from gut bacteria may play a role in human disease. Therefore, it is obvious that integrated analysis across genome, metabolome, microbiome and other omics profiles will provide means for managing health and combating disease.

4. OMICs technologies for disease prevention - future prospects

Other important goal of personalized medicine is to prevent the development of various diseases before their onset. Current preventive treatments are limited to surgery - removing a tissue or an organ if a person possesses relevant cancer markers. However, this approach is not applicable to diseases like Alzheimer, Huntington's disease, and congenital muscular dystrophy. One of the recent approaches to address such issues is the nutrigenomics-based diet treatment. Omics profiling can be an effective approach in detecting large-scale or pathway-level changes, can show patient-specific trends and add statistical support through repeated measurements.

APPLICATION OF OMICS TECHNOLOGIES IN TOXICOLOGY

The use of omics tools to evaluate environmental health status is becoming more and more important in environmental management and risk assessment. Omics permit connecting molecular events and adverse effects with biological levels of organization relevant for risk assessment schemes. The analysis of the RNA transcripts level or cellular proteomics have already gained particular interest. These studies helped in better understanding the interactions between the stress factors and the organism, but also to unravel the modes of action (MoA) of various toxic substances. Therefore, from a toxicological point of view, omics techniques allow to effectively and accurately generate important information on substance-induced molecular perturbations in cells and tissues that are associated with adverse outcomes.

Cellular regulatory pathways involve a range of different (bio)molecules, which are characterized with different physicochemical properties and display complex non-linear interactions. A single-omics technique can measure biomolecules of a specific type, RNAs or proteins, and frequently these are not even the entirety of biomolecules of one type but a smaller subset of them. For example, the detection of short and long RNAs, or short peptides and proteins requires different transcriptomic or proteomic approaches. Only the use of multi-omics, also known as cross-omics approaches, permits to directly identify a significant fraction of a pathways' response to chemical exposure. Thereof, utilizing multi-omics strategies for toxicological research questions has greater importance.



2019-1-BG01-KA203-062371

Such system toxicology approach, integrating classical toxicology with quantitative analysis of large molecular networks and functional changes occurring across multiple levels of biological organization, is promising in analyzing the full spectrum of xenobiotics toxicity. A systems toxicology approach includes the following elements:

- Absorption and distribution of the xenobiotic within the biological system.
- Transformation of the xenobiotic and generation of toxic and/or non-toxic intermediates.
- Interaction of the xenobiotic and/or its products with the cellular targets.
- Modifications in the cellular molecules including genes, proteins, and lipids.
- Structural manifestations of target organ toxicity.
- Functional manifestations of toxicity, and
- Elimination of the xenobiotic and/or its intermediates from the biological system.

1. Toxicological studies via OMICs analysis

For better understanding the cellular response to chemical exposure, studies should be focused on different levels: transcriptomics, proteomics, phosphoproteomics, and metabolomics. Epigenomics should be considered as well for specific cases.

Transcriptomics. Transcriptomics main aim is the comprehensive detection of RNAa in the cell. A pathway response in transcriptomics is primarily detected via a known set of target genes that are found differentially expressed. The techniques applied to determine changes in gene expression in response to exposure to a xenobiotic include: northern hybridization, quantitative real time PCR (QRT-PCR), subtractive hybridization, serial analysis of gene expression, microarray analysis, and next generation sequencing. Depending on the purpose of the study, one or more of these techniques may be applied to define either partial or global expression profiles in the tested biological sample. However, the most commonly used gene expression profiling techniques are QRT-PCR, microarray analysis, and NGS. If the aim is to investigate the expression profile of a single or a limited number of transcripts, QRT-PCR analysis is the method of choice. QRT-PCR analysis is a relatively simple technique that includes the reverse transcription of mRNAs followed by PCR amplification of the resulting cDNAs using primers specific for the targeted gene as well as for one or more house-keeping genes. The amount of the PCR amplified gene products are then quantified. As QRT-PCR can accurately determine the expression of an individual or limited number of transcripts, it has some limitations from a system toxicology perspective. In order to study the systems-level toxicity of a given xenobiotic, it is necessary to employ a high throughput techniques that is capable of determining expression levels of the entire transcriptome. That is why the microarray and next generation sequencing are the most popular global transcriptome analysis techniques applied in toxicology research.

The microarray technique represents a major breakthrough analyses in the post-human genome sequencing era. Like Southern hybridization technique, the microarray approach is grounded on the



2019-1-BG01-KA203-062371

ability of DNA molecules to bind or hybridize in a nucleotide sequence-specific manner. Microarray can be considered as a high throughput northern blot hybridization in which almost all the genes that are expressed in a biological sample at a given time can be detected simultaneously. Therefore, it allows to detect small differences in the global gene expression profiles in a given biological sample in response to xenobiotic(s) exposure. The microarray technique involves the isolation of high quality RNA, reverse transcription of the mRNA to synthesize cDNA, cRNA synthesis and labeling to generate “probes”, hybridization of the “probes” with “targets” present on the microarray, detection of the signal intensity of the hybridized probe-target complex, and analysis of the resulting data to determine the levels of expression of the genes of interest. Microarray-based analysis has been used in different areas such as environmental toxicology, drug development, and occupational toxicology. The data obtained provides a snapshot of all the genes whose expression levels are affected in response to xenobiotic exposure.

Next generation sequencing (NGS) is the most recently developed global transcriptome analysis. This technique can be applied to determine a global gene expression profile in RNA obtained from different biological samples. The key steps involved in NGS are RNA isolation, removal of abundant RNA molecules such as ribosomal RNA and globin RNA, preparation of sequencing libraries, PCR amplification, and sequencing of the libraries. The resulting data is analyzed, and the quantity of the individual transcripts is determined.

Proteomics: Transcriptomics is often the first choice for systems-level investigations as numerous well-established measurement methods have been developed during the years. However, protein alterations can be considered to be closer to the relevant functional impact of a studied stimulus. Although mRNA and protein expression are tightly related through translation, their correlation is limited. The mRNA transcript levels only explain about 50% of the variation of protein levels. This is due to the additional levels of protein regulation including their rate of translation and degradation. Moreover, the regulation of protein activity does not stop at its expression level but is often further controlled through posttranslational modifications. From toxicology point of view such examples for important post-transcriptional regulation include: the tight regulation of p53 and hypoxia-inducible factor (HIF) protein-levels and their rapid post-transcriptional stabilization, e.g., upon DNA damage and hypoxic conditions; the regulation of several cellular stress responses (e.g., oxidative stress) at the level of protein translation; and the regulation of cellular stress response through protein phosphorylation cascades. The most often applied proteomics analyzes being of key importance for toxicology include: 2D polyacrylamide gel electrophoresis, gel-free liquid chromatography, mass spectrometry and analysis of the posttranslational modifications.

2D polyacrylamide gel electrophoresis (2DGE) is used to study perturbations on the proteome based on changes in protein expression. The 2DGE technique relies on the separation of proteins based on their pH (charge) as well as their size. It has the capability to separate and visualize up to 2000 proteins in one gel. The first dimension, which is known as isoelectric focusing (IEF) separates the proteins by their isoelectric point (pI). The second dimension further separates the proteins by their mass. State-of-the-art image acquisition and analysis software allow the comparison of control and treated samples helping to identify the differentially regulated proteins by their relative intensity. A variant of 2DGE is



2019-1-BG01-KA203-062371

difference gel electrophoresis (DIGE) which is based on labeling of proteins with fluorescent cyanine dyes (Cy2, Cy3 and Cy5). Although 2DGE is a powerful tool to identify many proteins, the approach has its limitations. The main constraint is that not all proteins can be separated by IEF, such as membrane, basic, small (< 10 kDa) and large (> 100 kDa) proteins. The second limitation is that less abundant proteins are often masked by the abundant proteins in the mixture.

Other frequently used techniques are the gel-free liquid chromatography mass spectrometry (LC MS/MS). The LC approach employs the differences in the physiochemical properties of proteins and peptides, i.e., size, charge, and hydrophobicity. 2D-LC can be applied to fractionate protein mixtures on two columns with different physiochemical properties and maximize the separation of proteins and peptides in complex mixtures. Mass spectrometry is widely believed to be the key technology platform for toxicoproteomics. Its main advantages include unsurpassed sensitivity, improved speed and the ability to produce high throughput datasets. Due to the high precision of MS, in tissues and biological matrices can be detected peptides in the femtomolar (10^{-15}) to attomolar (10^{-18}) range. This is extremely beneficial in comparative analysis where simultaneous analysis of control and treated samples are carried out. Thus, comprehensive understanding of how stimuli affect the proteome, and the subsequent identification of potential biomarkers can be achieved.

As system biology requires accurate quantification of a particular set of peptides/proteins across multiple samples, targeted approaches for biomarker quantification have been also developed. Such technique is the selected reaction monitoring (SRM). SRM measures peptides produced by the enzymatic digestion of the proteome in triple quadrupole MS. An SRM-based proteomic experiment starts with the selection of a list of target proteins, derived from preceding experiments or data search such as a pathway map or literature. This step is followed by 1) selection of the proteotypic target peptides that optimally and uniquely represent the protein target, 2) selection of a set of suitable SRM transitions for each target peptide, 3) detection of the selected peptide transitions in a sample, 4) optimization of SRM assay parameters, and 5) application of the assays to the detection and quantification of the proteins/peptides. The major advantages of the SRM technique are: 1) possibilities to monitor tens to hundreds of proteins during the same run, 2) possibilities for absolute and relative quantification, 3) highly data reproducibility, and 4) yielding absolute molecular specificity. However, some limitations of this technique also exist: 1) only a limited number of measurable proteins can be included in the same run, and 2) even with its high sensitivity it cannot detect all the proteins present in an organism.

Another MS-based targeted approach known as parallel reaction monitoring (PRM) has also been employed. It is centered on the use of next-generation, quadrupole-equipped high-resolution and accurate mass instruments (mainly the Orbitrap MS system). This technique is closely related to SRM but allows for the measurement of all fragmentation products of a given peptide in parallel.

Other key approach is the analysis of the posttranslational modifications (PTMs) occurred due to the toxic exposure. They represent an essential mechanism for regulating the cellular proteome. PTMs are chemical modifications that have an important role in regulating activity, localization and interactions of cellular biomolecules. The identification and characterization of proteins and their PTM sites are very



2019-1-BG01-KA203-062371

important to the biochemical understanding of the PTM pathways. This knowledge provides deeper insights into the possible regulation of the cellular physiology triggered by PTM. Some examples of PTMs include phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis. During the past decade, MS-based proteomics has proven to be a powerful tool for the identification and mapping of PTMs. Moreover, the MS-based approaches took great advantage from the progress in MS instrumentation making the analysis more sensitive, accurate and with higher resolution for the detection of less abundant proteins.

Metabolomics: The metabolome is different to the other omics analysis, as it is dealing with a collection of chemically highly heterogeneous molecules. The metabolome is typically defined as the complete balance of all small molecule metabolites (< 1500Da) found in a specific cell, organ or organism. Thanks to initiatives like the Human Metabolome Project, the endogenous metabolome of human and animal model organisms has been mapped to a large extent. Metabolome measurements employ either nuclear magnetic resonance (NMR)-based detection or gas or liquid chromatography hyphenated to MS. NMR based approaches have the capacity to quantify the studied metabolites and constitute the gold standard for the structural elucidation of unknown metabolites. MS-based approaches surpass NMR in terms of sensitivity and can be run in an untargeted or targeted manner.

Metabolomics differs from proteomics and transcriptomics in several aspects:

- (i) detects the response at very different levels of a pathway, and
- (ii) simultaneously capturing molecules of a pathway that respond on extremely different time scales.

The progression of metabolomics has made it routinely and highly useful in many toxicology-related studies. It has remarkable potential in screening drug-induced cellular or organ toxicity. Rapid assessment of biological samples, together with mathematical and statistical analysis (chemometrics) of obtained data serves as a powerful tool for drug safety assessment. For example, if a drug or chemical substance can cause hepatic toxicity by inducing cellular oxidative stress, metabolomics analysis of the drug-exposed biological sample can provide direct information about metabolites that are markers of the oxidative stress. It can also identify metabolites that are considered as known biomarkers of hepatic injury, thus suggesting hepatic pathology. In addition, metabolomics analysis can also provide information about metabolites that are indirectly associated with toxicity.

Metabolomics has several advantages over conventional clinical pathology assessment. For instance, a study examining an unknown compound has shown that this compound increases serum cholesterol levels. Upon the utilization of metabolomics, it has been observed that, apart from cholesterol, the same compound also raises several phytosterols, indicating that the higher cholesterol is a result of the sterol absorption in the gut. Such findings clearly indicate that metabolomics has a much higher potential in terms of identifying exact toxicological endpoints.



2019-1-BG01-KA203-062371

Epigenomics. Main focus of epigenomics is to study the chemical modifications of the DNA and proteins organizing the three-dimensional structure of genomic DNA. Nowadays, DNA methylation is assessed by a modified genome sequencing approach (Methylome-seq). However, this approach still requires high coverage associated with still significant sequencing costs. That is why the targeted microarray-based approaches are most frequently used. DNA methylation includes two different chemical modifications with distinct functional consequences - 5-methylcytosine and 5-hydroxymethylcytosine. Histone modifications are analyzed using Chromatin-immuno-precipitation method applying antibodies specific for the targeted modification followed by sequencing (ChIP-seq). Typically, to get a comprehensive picture, several modifications need to be detected. That requires performance of several ChIP-seq analysis simultaneously. As ChIP-seq needs more sample material than the other epigenomic approaches, such a strategy may be hard to implement in toxicological studies. ATAC-seq is an alternative approach that does not investigate chemical modifications directly, but rather one of the modifications' main consequence, DNA accessibility.

However, knowledge on the involvement of regulatory pathways and specific epigenetic modifications is still scarce. Epigenomics is thus of limited value when a pathway-based data integration approach is used. However, epigenomic data have proven highly valuable for trans-generational studies or when trying to predict long-term effects of chemical exposure based on omics data of short-term exposure.

2. OMICs applications for toxicology risk assessment and regulatory submissions

In the last decades, 'omics technologies have been extensively used in research and it was demonstrated that they are capable to provide a profound insight into the biochemistry and physiology of the cell and any harmful effect of xenobiotics. This has led to an enthusiastic approval by research toxicologists. Hopes were expressed that 'omics technologies would provide the tools to identify a set of biomarkers of adverse effects and their modes-of-action. Thus, the prediction of human effects during substance hazard assessment will be improved and a contribution will be made to the development of alternative methods to animal testing. Despite this, the translation of 'omics into the regulatory domain remains at best cautious.

The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) has, therefore, organized a series of workshops to evaluate the possibilities and challenges of omics techniques in chemical risk assessment. The most recent workshop concluded that omics approaches contribute to answering relevant questions in risk assessment including:

- (i) the classification of substances and definition of similarity,
- (ii) the elucidation of the mode of action of substances, and
- (iii) the identification of species-specific effects and the demonstration of human health relevance.



2019-1-BG01-KA203-062371

However, a need was identified to increase the reproducibility of omics data acquisition and data analysis, as well as to define the best practices. Furthermore, the regulatory use of any test method is closely related to the specific legal framework that is relevant for the substance under investigation. Considering the REACH Regulation (EP and Council of the EU, 2006), “omics-based data could potentially be used for regulatory submissions”. To use omics in REACH dossiers, the REACH Regulation lists different requirements: specific standard information that cover specific physico-chemical, toxicological and ecotoxicological limits. These data are evaluated during hazard and risk assessment, and they are analyzed to determine risk management alternatives, respectively. The application and integration of omics tools may be useful in different levels of regulatory hazard identification and assessment contributing to:

1. Classification and labelling (C&L) of substances.
2. Weight-of-evidence (WoE) approaches to elucidate the MoA of the substance under investigation.
3. Substantiation of chemical similarity.
4. Determination of points-of-departure (PoDs) for hazard assessment.
5. Demonstration of species-specific effects and human health relevance.

The fast development of omics technologies pose several challenges for facilitating their use for hazard assessment, particularly from the regulatory submission standpoint. Sets of ‘big data’ must be condensed applying complex approaches and applying specific knowledge to obtain information that is pertinent for hazard assessment. Moreover, knowledge in this rapidly evolving area is not necessarily familiar to many researchers working in the regulatory settings. Accordingly, the lack of regulatory approval of omics technologies is not only linked to a lack of best practices frameworks and quality criteria, but also to a lack of confidence in the analysis and the level of uncertainty with respect to what constitutes enough data. Scientists and regulators have to first gain experience with the data obtained from this new technology and then to build confidence in its applicability.

ENVIRONMENTAL OMICS AND BIODIVERSITY

Generally, most of the environmental studies comprise culturing of isolated microorganisms or amplifying and sequencing conserved genes. However, some difficulties exist in understanding the complexity of large numbers of various microorganisms in an environment. This led to the development of new techniques allowing to enrich specific microorganisms for upstream analysis, such as single-cell isolation and analyses. Advanced tools in metagenomics and single-cell genomics (SCG) are paving the way to new methods of studying and understanding our environment.

Early environmental studies of organism diversity required samples cultivation to increase DNA quantity before genomic libraries to be constructed. At the time, sequencing of 16S and 18S rRNA was



2019-1-BG01-KA203-062371

the best option available to define environmental diversity. However, it became clear that majority of sample communities represents uncultivable species. Fortunately, polymerase chain reaction (PCR) of rRNA allows access to both cultivable and uncultivable sample diversity. Because PCR can be performed directly on environmental samples without cloning, it can be utilized to amplify targeted genes directly from the environment.

Importance of sequencing both noncoding rRNA and protein coding genes is growing, with specific interest in reconstructing bacterial genomes. A comprehensive sampling of all genes from all organisms existing in a complex sample is achievable via shotgun metagenomic sequencing. Using this method make possible to evaluate bacterial diversity and to detect the abundance of microorganisms in a specific environment. Moreover, it allows the reconstitution of complete genomes of uncultivable microorganisms. In addition to the discovery of new archaeal, bacterial, protozoan, algal, fungal, and eukaryotic species, this technique is crucial for the reconstruction of viral genomes. This was a critical advancement taking into account that viruses lack a shared universal phylogenetic marker such as bacterial 16S rRNA or eukaryotic 18S rRNA.

Another approach is the metatranscriptome approach which rather than using genomic DNA (gDNA) involves the collection of the entire RNA community. The latest is used to construct cDNA libraries that are sequenced. Although RNA is less stable than DNA, it gives a snapshot of the current state of the community by revealing up and down regulated genes. To date, one of the most important metatranscriptome studies are performed for the marine water communities.

1. Single-cell genomics

Metagenomic sequencing is a helpful tool for understanding environmental communities. However, assembling gene catalogs and composite genomes can be challenging. Furthermore, it is difficult to distinguish between genes that originate from the same or different organisms in genomes that are already assembled (Fig. 3). Therefore, there is considerable interest in developing a system to understand the organization of discovered genes and pathways within genomes.

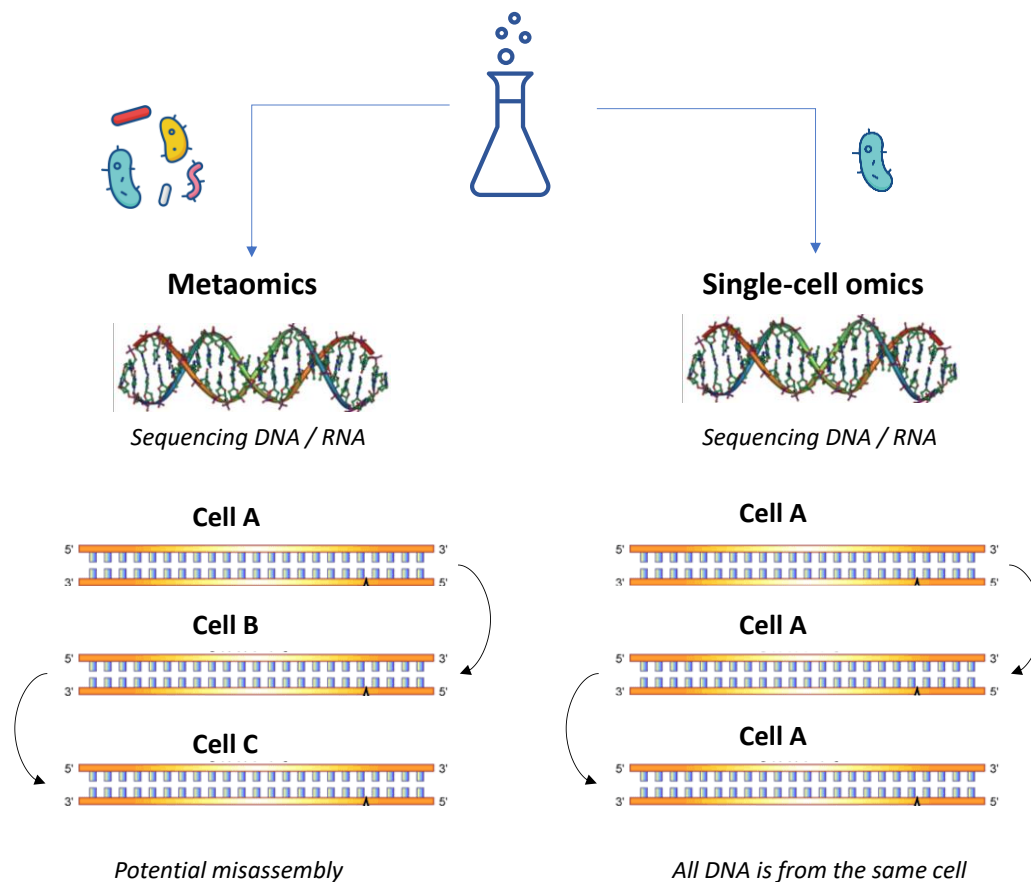


Fig.3. Metagenomics vs single cell genomics

The challenge of sample heterogeneity could be partially solved via using two techniques - scaling-up and deep-sequencing. Different approaches are required to compare metagenomic samples collected from different environments and to reveal the composition and organization of microorganism genomes in the environment. One of the approaches is to enrich the culture of a specific organisms based on their specific environmental functions. This type of targeted-metagenomic approach permits for the development of composite microbial genomes. Additionally, to the uncultivability of many organisms, other problems occur when faster growing microorganisms outgrow slower ones, subsequently introducing additional biases.

Hence, instead of culturing, a complex sample can be enriched for a target cell population using fluorescence-activated cell sorting (FACS). Using this approach helps sort cells on the basis of their shape, size, and density. Alternatively, specific organisms of interest can be enriched by handling a small



number of cells that possess specific function in the environment. With the development of whole genome amplification (WGA) techniques (e.g., multiple displacement amplification (MDA) or rolling circle amplification (RCA), the gDNA of a restricted number of cells can be amplified and sequenced. However, WGA techniques create potential deviations because the abundance of specific sequences vary and therefore are not equally amplified. Hence, gDNA amplification of defined populations with a restricted number of cells compromises the quantitative analysis of metagenomics.

Latest advancements in single-genome amplification techniques enable SCG using physical cell separation (usually by FACS on 96-well plates), cell lysis, and WGA. An amplified genome from a single cell can then be sequenced. While metagenomics is accomplished by collecting the microorganism community and extracting and sequencing total DNA, SCG separates and studies individual cells from the microorganism community. Obtained sequencing data provides quantitative information on genomic variability in microorganism populations. Gene insertions, deletions, duplications, and genome rearrangements can be studied on a single-cell level. In this way the complex metabolic pathways can be analyzed for an individual cell.

SINGLE-CELL ISOLATION AND PROCESSING

Studying the genome and its regulation at the population level is done by analyzing millions of cells at once. However, such investigations do not allow for a view of the heterogeneity in biological systems, nor do they characterize the current state of the population. That is why the SCG is gaining more and more popularity for studying both cultivable and especially uncultivable microorganisms. The process to obtain, multiply, and sequence genetic material from either live or dead cells is relatively simple. First the cell needs to be isolated and lysed and then the released genetic material is amplified and the genomic library can be constructed.

One of the approaches for single cell isolation is the micromanipulation. Cells of interest are identified, isolated, and examined microscopically. Suitable cells are isolated using a micropipette, laser source, optical tweezers, or by real-time microfluidic flow. Using microfluidics, a desired phenotype correlating to a specific genome can be selected. In addition, cells can be easily observed before capture.

Other approach is the random encapsulation. Cells are randomly selected by serial dilution. The diluted sample can then be transferred into microwells, or the cells can be encapsulated by microdroplets. Using this technique, individual cells can be separated, processed, and subjected to genomic analysis.

Flow cytometry is the most popular procedure of random encapsulation. FACS allows the isolation of cells possessing specific criteria, such as the size, shape, color, or even presence of specific nucleic acids or activities, using fluorescent dyes (Fig. 4). One of the main advantages of FACS is its high throughput, high sorting speed, and ability to distinguish live cells. Moreover, FACS can detect the

presence of cells in a droplet, reject empty droplets, and transferred the cells with desired properties into multiwell / microtiter plates. Then cells can be lysed, and genomes can be amplified in single wells.

Other possibility for single cell isolation is the microdroplet technology (Fig. 4). Microdroplets are the emulsion formed when cells in the water phase are vigorously mixed with oil. Oil droplets encapsulate cells in the water phase. Originally, PCR performed on an emulsion was thought to be much more specific than analysis of cells in the water phase alone. It was suggested that the yields of reaction products are greater and fewer chimeric byproducts are formed. Moreover, the droplets formed enclosed a limited number of template molecules (ideally only one). Latest developments in microfluidics permit to control the droplet size (nanoliter or even the picoliter), providing aqueous uniform (monodisperse) droplets in oil. One of the main advantages of the two-phase system of microdroplets over FACS is the possibility to handle single cells as separated units. Each droplet act as an independent well where cells can be grown and later screened for the products expressed. Released gDNA can be effectively amplified in microdroplets and used for sequencing-library preparation. Bacterial, yeast, plant, insect, and mammalian cells can be easily studied using microdroplet technology on agarose or other microgels. Even multicellular organisms can be encapsulated and grown in a droplet. Hypothetically, it is possible to measure any secreted molecule for which a fluorescently labeled ligand exists. Furthermore, the activities of various cell proteins can be monitored.

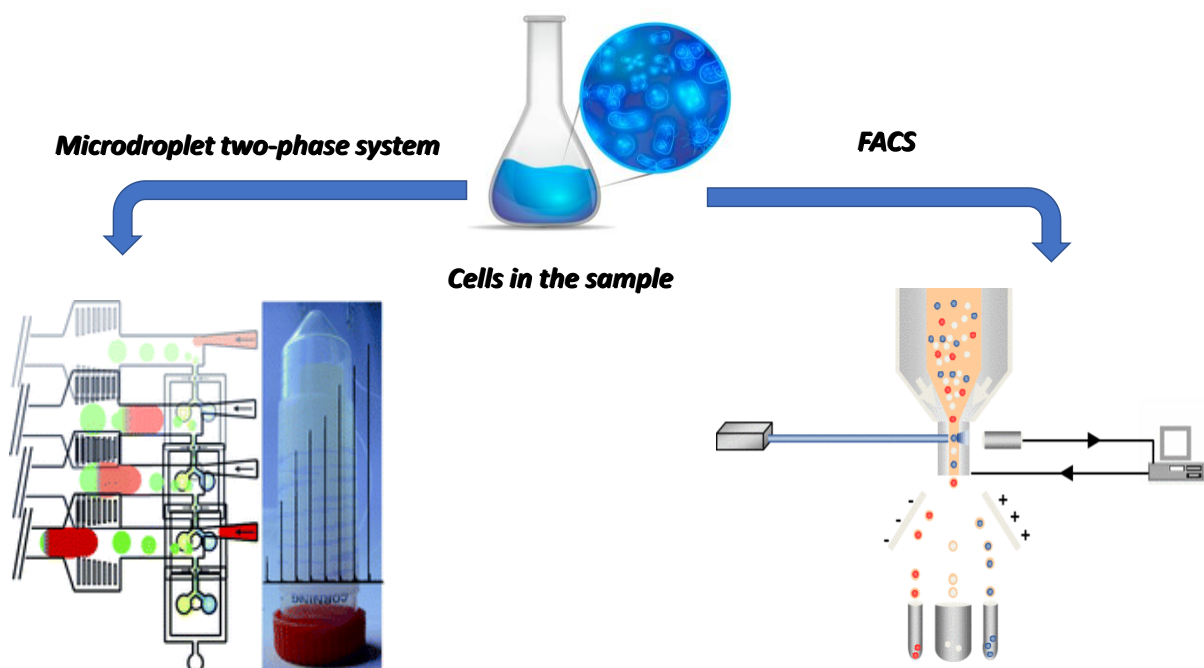


Fig. 4. Single cell isolation techniques



APPROACHES FOR STUDYING VARIOUS CELLS

Living cells can generally be classified into bacteria, archaea, and eukaryotes. Although viruses are not considered “living,” they constitute an important part of life science studies. Various single-cell analysis exists that are suitable for different organisms.

Viruses are present in practically every environment. They are the most abundant and diverse biological entities. However, to isolate single virus and sequence its genome, a suitable cultivable virus-host system must first be established. For example, eukaryotic algae are a host to an incredible diversity of viruses. However, using both FACS and single-droplet systems, viruses can theoretically be isolated without dependence on the viral-host system.

Archaea and bacteria have a similar cell structure. Only differences in cell composition and organization set these two domains apart. Bacterial cells are ordinarily 0.5 - 5.0 μm in diameter, while archaea can be larger and reach 15 μm in diameter. Basically, all methods described above can be used to separate and process single archaeal and bacterial cells. However, during FACS high currents may slow the growth of these organisms.

Eukaryotes differentiate from other life domains by their membrane bound nucleus and membrane-bound organelles. They may be multicellular or single-celled organisms. Cell size typically range from about 1 - 100 μm or they are ten-times larger than bacteria. Dilution, FACS, and droplet sorting are all appropriate methods for isolation of eukaryotic cells.

Multicellular organisms can theoretically also be isolated and processed by various methods. FACS for example can easily sort multicellular organisms. If incubated in droplets for a sufficient amount of time, single-celled organisms may synthesize a protein or ligand and have the opportunity to multiply, potentially creating a multicellular structure.

Molecular studies of our environment and ecology gain more and more knowledge for living communities. With the development of molecular tools, especially sequencing machines, it became possible to sequence whole communities from selected environments. Recent advancements in single-cell technologies facilitate the isolation and amplification of genomic material from single cells and allow structuring of numerous sequence-based libraries. Furthermore, data obtained from SCG supplements data obtained by metagenomics.



OMICS APPROACHES IN INDUSTRIAL BIOTECHNOLOGY AND BIOPROCESS ENGINEERING

Biotechnology utilizes biological processes, organisms, or systems to generate products and technologies that are improving human lives. The use of biological systems to produce bioproducts of commercial importance is a key component of the biotechnology industry. This biotechnological approach has found application in different areas: energy, material, pharmaceutical, food, agriculture, and cosmetic industries. The bioproducts made from biomanufacturing processes are usually metabolites and proteins, which are obtained from cells, tissues, and organs. The biological systems synthesizing these bioproducts can be natural or modified by genetic engineering, metabolic engineering, synthetic biology, and protein engineering. “Omics” technologies have great importance for biotechnology and metabolic engineering helping in characterizing and understanding metabolic networks. The significant amount of knowledge acquired from omics-driven experiments can be applied in the development of biotechnological tools and the advancement of metabolic engineering. This enables the manipulation of complex biological systems toward creating robust industrial biomanufacturing strategies.

1. Omics-Guided Biotechnology

“Omics” tools are used more and more in the development of biotechnological processes and the production of many vital products. Application of “omics” technologies in characterizing and understanding biological systems has permitted to select and predict phenotypes, which facilitates the optimization of biotechnological processes toward enhanced production (in quality and quantity) of commercially relevant products (Figure 5).

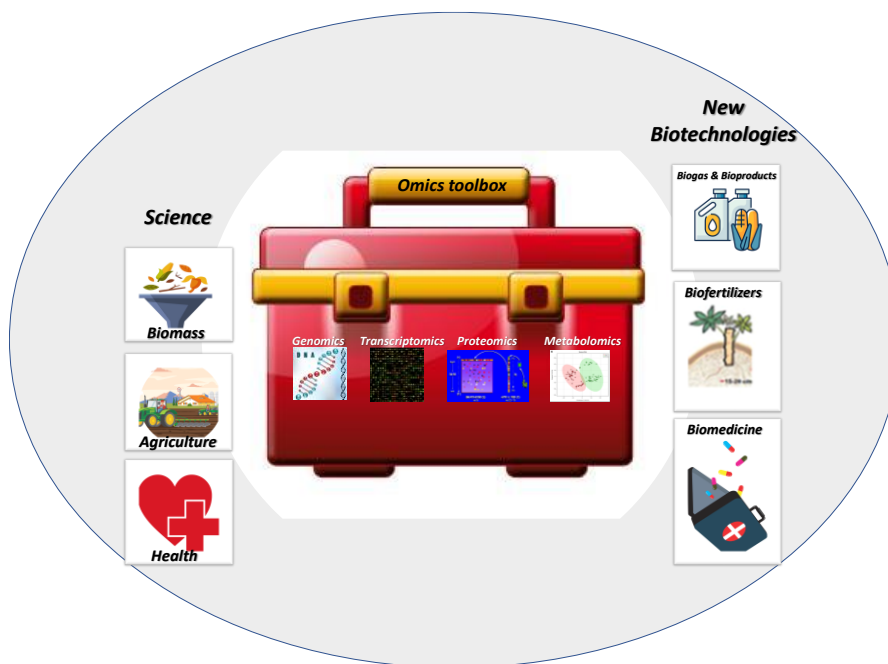


Fig. 5. Omics tools in Biotechnology

PRODUCTION OF BIOFUELS AND BIOPRODUCTS

Microbial production of natural compounds represents an attractive and more sustainable alternative to traditional petrochemicals. Its implementation has led to a growing catalog of natural products and high-value chemicals. For instant, the use of lignocellulosic biomass represents an economical approach to generate biofuels and bioproducts. However, to achieve stable conversion of low-cost raw material into value-added products at industrial levels needs systematic engineering plans. The Design-Build-Test-Learn (DBTL) cycle is becoming an increasingly implemented approach for metabolic engineering experiments. It represents a systematic and effective tool to force development efforts in biofuels and bio-based products. The DBTL cycle uses *in silico* approach to Design and Build genetic constructs in microbial hosts. Next, the information obtained from “omics” technologies, during the Test phase of the cycle, is transferred on to Learning processes (Figure 6). What is Learned is then fed back to new design cycles to achieve further strain development and optimization. Thus, a rapid optimization of microbial strains producers of any chemical compound of interest is facilitating. The weakest point in the DBTL cycle workflow is the Learning process since mathematical models are only as good as their assumptions. Therefore, both high quality and large “omics” data sets are required to improve training models, ensuring increased accuracy and reliability of the Learning process.

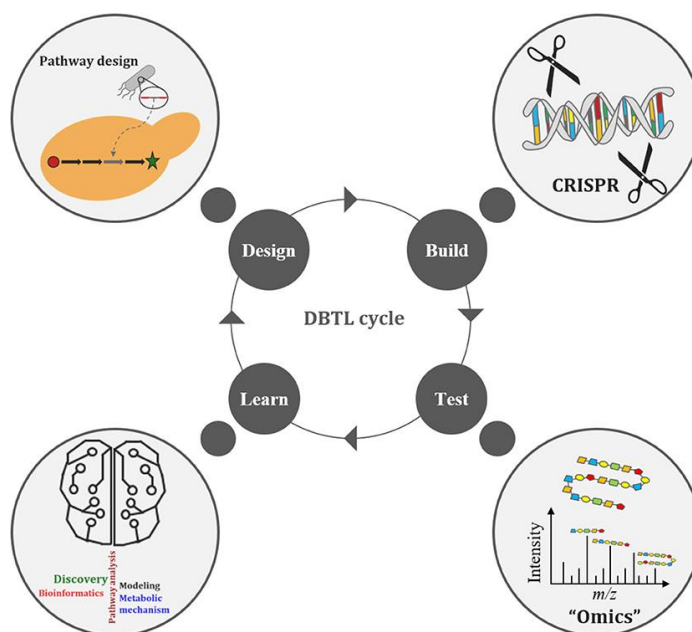


Fig. 6. DBTL cycle for optimization of biofuels and bio-based products

Often in the DBTL cycle genomic sequence information is the traditional approach utilized in the initial stages of a study. For example, barcode sequencing (Bar-seq) (a method using a short section of DNA from a specific gene or genes) can be used to study population dynamics of *Saccharomyces cerevisiae* deletion libraries during bioreactor cultivation, allowing the identification of factors that impact the diversity of a mutant pool. Single cell sequencing guided approach can be used to identify key mutated promoters for adjusting expression levels, thereby facilitating the dynamic regulation of microbial growth. Proteomics-guided approaches have been employed to engineer polyketide biosynthesis platforms for *in vitro* production of adipic acid in yeast. Additionally, metabolomics permits the assessment of pathway flux, carbon source utilization, and cofactor imbalance, which all contribute to the identification of pathway bottlenecks. Such metabolomics guided approach is already used for the characterization of the cannabinoid production in engineered *S. cerevisiae*. Cannabinoid analogs have been identified produced by several promiscuous pathway genes. Furthermore, application of metabolomics helped the design and optimization of a novel isopentenyl diphosphate-bypass mevalonate pathway in *E. coli* for C5 alcohol production. Recently, some authors have utilized the DBTL cycle to engineer *Rhodospiridium toruloides*, an oleaginous yeast species growing on lignocellulosic materials and producing the diterpene ent-kaurene, a potential therapeutic, exhibiting antimicrobial, anti-inflammatory, cardiovascular, diuretic, anti-HIV, and cytotoxic effects.



2019-1-BG01-KA203-062371

AGRICULTURAL AND FOOD BIOTECHNOLOGY

Recent innovations in agricultural biotechnology have led to new plant varieties, engineered by recombinant DNA technology and better responding to market demands and environmental challenges. In fact, “omics” tools in agricultural biotechnology have been used to enhance desirable phenotypic characteristics (e.g., color, taste, drought tolerance, pesticide resistance, etc.). “Omics” plays role not only in improving crop quality, consistency, and productivity, but also in the development of food crops with enhanced nutritional composition. Moreover, omics-driven systems biology helps in understanding the interactions between the “OMEs” and provide links between genes and specific characteristics.

In arable land soil is more susceptible to loss of structure, organic matter, minerals, and erosion. Thus, efforts are being made, via agricultural biotechnology, to provide a constant supply of nutrients essential to the growth of crops. An integral part of this approach is the use of biofertilizers. These are preparations containing specialized microbial inoculants that can fix, mobilize, solubilize, or decompose nutrient sources. Biofertilizers are generally applied through seed or soil and enhance nutrient uptake by plants. Widespread application of this approach, however, has been hampered by fluctuating responses of microbial inoculants across fields and crops. As a result, there is a pressing need to better understand the mechanisms underlying the interdependencies between soil microbial communities and host plant productivity. Recent genomics and exo-metabolomics studies showed that specific rhizosphere bacteria have a natural preference for certain aromatic organic acids exuded by plants. This fact suggested that plant exudation characteristics and microbial substrate uptake traits interact and form specific patterns of microbial community assembly. Furthermore, the application of genomics and transcriptomics to the study of phosphate uptake by various microalgae revealed a range of Pi transporters having specific expression patterns in relation to the availability of P. Currently, “omics” approaches are being used to study complex rhizospheric intercommunications, which is crucial to the development of new biofertilizers, thus promoting stable plant growth, better crop productivity and yield.

In the related field of food biotechnology, application of transcriptomics and metabolomics demonstrated that *Bacillus pumilus* LZP02 promote the growth of rice roots by enhancing carbohydrate metabolism and phenylpropanoid biosynthesis. Further, the application of “omics” in starch bioengineering provides better understanding on the most important enzymes for its biosynthesis. This facilitates the prediction on how starch-related phenotypes can be modified and ensures further progress in the research field of rice starch biotechnology. “Omics” also help solving issues related to food quality and traceability, to protect the origin of food, and discover biomarkers of potential food safety problems. The advance investigation of wine microbiome has a great influence for wine industry helping in better understanding of factors transforming grapes to wine, including flavor and aroma. “Omics” characterization of the complex relationships between these microorganisms, the substrate and environment, is essential to shaping wine production. Finally, combining “omics” technologies with genome editing of food microorganisms can be exploited to generate improved probiotic strains, develop innovative bio-therapeutics and alter microbial community structure in food matrices.



2019-1-BG01-KA203-062371

“OMICS” TECHNOLOGIES AND COVID-19

The coronavirus disease 2019 (COVID-19) is characterized by the Severe acute respiratory syndrome coronavirus 2 [i.e., SARS-CoV-2, which binds to the ACE2 receptor in the lung and other organs]. Due to its worldwide spread a global pandemic has been announced and much of the world's economy has been slowed. Up to April 2021 there are more than 138 million confirmed cases and 2.5 million confirmed deaths globally (<https://www.worldometers.info/coronavirus/>). Thus, there is an urgent need for an effective countermeasure to mitigate the spread of the pandemic.

Therefore, efforts are ongoing to fast-track the development and production of safe and effective vaccines against SARS-CoV-2. Preceding knowledge of SARS and Middle East respiratory syndrome (MERS) has facilitated targeting the spike protein as the viral antigen (via the ACE2 receptor). Furthermore, the genome sequencing of the SARS-CoV-2 in January 2020 made it possible to accelerate the development of next generation mRNA and vaccine platforms that encode for the antigen. Once injected into a host, the mRNA, encapsulated into lipid nano-particles, remains in the cytoplasm. When DNA is used, encased in an attenuated adenovirus vector, it enters the nucleus. The host cell translates these genetic materials into the spike protein. It settles on the surface of the cell and provokes an adaptive immune response mediated by T cells (e.g., CD4+ and CD8+) and B cells (i.e., antibodies). These vaccines were reported to be effective against SARS-CoV-2 in recent clinical trials, which underlines the significance of genomics to this new era of vaccine development.

Since the beginning of the outbreak of SARS-CoV-2 a protein interaction was created map and using a proteomics-based approach targets for drug repurposing was revealed. By applying proteomics analysis 26 of the 29 SARS-CoV-2 proteins were cloned, affinity tagged and expressed in human cells and the associated proteins were identified. A total of 66 human proteins or host factors were discovered as potential drug targets of 69 compounds. Two sets of these pharmacological agents showed antiviral activity. Moreover, computational immunoproteomics studies were performed supporting the lab-based investigations and allowing the evaluation of diagnostic products specificity, the forecast of potential vaccines adverse effects and the reduction of animal models utilization.

Furthermore, recently a method was developed based on metabolomics that is able to distinguish COVID-19 patients from healthy controls via the analysis of 10 plasma metabolites. Additionally, the data from lipidomics study suggests that monosialodihexosyl ganglioside enriched exosomes could be involved in pathological processes. Proteomics and metabolomics investigations in COVID-19 patient sera revealed that SARS-CoV-2 infection causes metabolic dysregulation of macrophage and lipid metabolism, platelet degranulation, complement system pathways, and massive metabolic suppression. The analysis of plasma metabolomic signatures showed similar profile to those described for sepsis syndrome. Furthermore, transcriptomics results showed upregulation of genes related to oxidative phosphorylation both in peripheral mononuclear leukocytes and bronchoalveolar lavage fluid. All this information suggests a critical role of mitochondrial activity during SARS-CoV-2 infection. Understanding the clinical presentation of COVID-19 as well as metabolomic, proteomic, and genetic profiles could help in discovering specific diagnostic, prognostic, and predictive biomarkers, ensuring



2019-1-BG01-KA203-062371

the development of more successful medical therapy. Moreover, differentiating metabolic biomarkers of severe vs. mild disease states in the lung during respiratory infections could lead to the discovery of novel therapeutics that modulate symptom and disease severity.

NUTRI-OMICS RESEARCH

More than a decade has passed since the idea of nutrigenomics was first introduced. Nutrigenomics, also referred to as nutritional genomics, nutritional omics, or nutri-omics, can be defined as an area of food and nutrition research making use of profound analyses of molecules or other physical phenomena.

Considering the complexity of the human body and its various interactions with food, it is conceivable that holistic analyses of food- body interactions are an important prerequisite for better understanding the effect of dietary components. The main assumption is that the combination of different omics platforms will provide a deeper insight into the influence of food components and the mechanism of their actions.

1. Genomics and nutrition

Human genes have different variants in the population. Having in mind that response to nutrition is a multigenic process, it is not surprising that non-related individuals may respond differently. The key hypothesis proposed in recent years stated that genetic variation underlies variation in nutrition-related disorders and risk for disease. In fact, nutrigenetics tries to explain how and to what extent nutrition-related traits and disorders are influenced by genetic variation. In other words, the key questions of nutrigenetics are:

- Which genes are involved in determining a defined characteristic?
- What is the functional identity of the variation by which people differ for this characteristic?
- How can this knowledge be used for the benefit of the population?

Usually, researchers define 'candidate genes' for a characteristic or disorder and search in those genes for variation. This is described as the 'hypothesis-driven' approach because the selection of the genes is based on their presumed function. For example, the genes for lipoproteins are used as candidates for obesity and the genes for insulin signaling - as candidates for diabetes. Other attractive candidates in humans are the orthologous genes underlying animal models with a Mendelian segregation of a characteristic. A well-known example is the db/db mouse. It is a model for diabetes in which the gene



2019-1-BG01-KA203-062371

for the leptin receptor was identified. The human homologue was then tested in persons with an impaired glucose tolerance. One variant (allele) of the gene was identified and was proven that it was associated with a significantly higher frequency of diabetes occurrence in comparison to the general population.

Other methods applied is related with studying the preferred segregation of an allele from parents to an affected child (transmission disequilibrium test) or to more affected siblings (sib-pair analysis). If a gene is found to be linked to the characteristic, it remains to be proven whether the associated or linked allele itself is the risk factor, or whether it could be used only as a marker for a nearby causative variation. If the genetic variation is linked with amino acid variation in the protein, a functional test could be developed and the allelic proteins can be compared for their enzymatic activity, DNA-binding affinity, etc.

Another approach to search for risk-conferring genes for a certain characteristic or disorder is the total genome scan. Hundreds of polymorphisms are detected with a random distribution across the genome but with a known chromosomal location. For every polymorphic site it is then identified whether an association exists between the allele and the characteristic. Using this approach risk loci can be identified for many human nutrition-related disorders like obesity and diabetes.

After the second world war in many Western countries, it was observed that the incidences of spina bifida gradually decreased. It was suggested that the improvement of the diet had brought forth this effect. A large epidemiologic study was started. The results obtained showed that the enrichment of the diet with vitamins was able to prevent the occurrence and recurrence of spina bifida in humans with more than 50%. Folic acid was discovered to be the active substance and now in many countries pre-conceptional folic acid supplementation is recommended as a general preventive measure. Simultaneously, it has become clear that concerning the risk for a child with spina bifida, the women in the population can be roughly divided into three groups: [i] not at risk even at a low folate intake, [b] at risk at low dietary folate but can be helped by increased folate intake, and [c] at risk despite extra folate intake. Genetic predisposition was presumed, and candidate genes taken from the folate metabolism were examined for genetic variation. This led to the detection of the 667C->T and 1298A->C alleles in the gene for methylenetetrahydrofolate reductase (MTHFR) as risk factors. However, these risk alleles are discovered to some extent in all three groups of women illustrating that genetic testing in the context of a personalized diet would be inadequate. Moreover, it became clear that not all the relative risk could be explained by these alleles. Therefore, the search for genetic variation in other genes is continued and new candidate genes may be found.

Currently, several hundreds of genes related to specific nutrition-related characteristics and disorders have already been identification via using genetic experiments in various populations and the investigation of animal and *in vitro* model systems. Finally, the nutrigenomics and nutrigenetics research should lead to the comprehensive understanding of the etiology of those characteristics and disorders with respect to the interacting genes, the dietary components, and the relative risk conveyed by both genetic variation and diet. This combined knowledge will permit the genetic profiling of every individual and thereby assess her/his risk for developing nutrition-related disorders. Based on this personal risk



2019-1-BG01-KA203-062371

profile of genetic factors, a ‘personalized diet’ could then be proposed by which the onset of a disorder could be prevented or at least delayed. Already some companies offer genetic testing for alleles that were found to be associated with certain characteristics. For example, the ApoE allele is tested as a risk factor for cardiovascular disorders and an allele of the alcohol dehydrogenase gene is used to predict potential alcohol sensitivity and (ab)use. However, it is important to note that the presence of certain allele may be said to increase the risk for a disorder by 100%, while in absolute terms the risk would usually still be very small like an increase from 0.001 to 0.002. Additionally, testing only one genetic factor might give an incomplete or even a wrong view of the situation. Nowadays, we have no idea of all the genetic factors and how they interact. Till more information is gathered, the advice based on simple testing probably has to remain simple, sounding like “drink less alcohol” or “eat less saturated fat”.

2. Transcriptomics and nutrition

Transcriptomics is the most widely employed in food research because of the many advantages of the DNA microarray technology, including the comprehensiveness of the expression data, established protocols, and high reliability and reproducibility of the data. It has been also reported the alternations of global gene expression in response to different dietary changes such as nutritional deficiency, fasting, overeating, and ingestion of specific food factors. As one example of the performed attempts to acquire reference data a transcriptome analysis of the liver of rats treated with mild caloric restriction has been performed. When using animal experiments, it was shown that a modest reduction of food intake or altered intake pattern could be sufficient for occurrence of significant metabolic changes. Therefore, it is important to distinguish between the direct effects of the food component and the secondary effects caused by the change in eating behavior. When rats were fed a diet with 5 to 30% less food than that consumed by an ad libitum-fed group for 1 week or 1 month it has been observed restriction – dependent changes in the expression level of cyp4a14. In fact, the cyp4a14 gene was induced by even a low level of caloric restriction. This data suggests that the gene can be used as a biomarker for the beneficial effects of food on energy metabolism.

3. Proteomics and nutrition

Many food and nutrition research studies have been carried out by using proteomics approaches. For instant, the effect of mild caloric restriction described above was also examined by proteomics. Nine significantly up-regulated proteins and nine down-regulated proteins have been detected after the proteome comparison of the liver of rats treated with 30% food restriction with control ones. Ten percent restriction caused the up-regulation of 9 proteins and the down-regulation of 2 proteins. An interesting discovery was the up-regulation of prohibitin, whose involvement in the regulation of longevity was recently revealed. This result suggests that prohibitin can be applied as an effective biomarker of the beneficial effects of food factors. Although the current stage of proteomics research is much less



extensive compared with that of transcriptomics, the study described here together with other nutritional proteomics research findings indicates that proteomics is a highly promising tool for the discovery of biomarkers.

4. Metabolomics and nutrition

About ten thousand different types of major metabolites exist in animal bodies, while the number of proteins is thought to exceed 100,000. This feature of metabolites is expected to result in more comprehensive features of metabolomics analysis than proteomics ones. However, the analysis of metabolites is in fact fraught with difficulties and usually requires the use of sophisticated techniques and high skill level personnel. Another difficulty originates from the width of the metabolite's abundance. Regardless of these obstacles, metabolomics is a powerful tool in food and nutrition science.

Many investigations showed that gut and serum metabolism is influenced by the change in gut microbiome composition. The gut microbiome and its alternation in relation to the diet are essential when evaluating dietary intervention trials with metabolomic end effects. The comparison between germ-free mice colonized by human baby flora and conventional mice showed the complexity of diet modifiable microbiome/metabolome covariation. In this study, the effect of the intestinal microbiome on plasma metabolites was studied. Data revealed that more than 10% of the plasma metabolome is directly dependent upon the microbiome. Some examples for microbial dependent compounds in plasma include cinammic acid, glycine conjugated compounds, hippuric acid, and other plasma metabolites. The gut microbiome also directly affects the host's ability to metabolize lipids, carbohydrates, and proteins, and can carry out several phase II detoxification reactions. There is also evidence that gut microorganisms utilize nonnutritive phytochemicals. For example, several investigations demonstrated that levels of three microbiome-dependent diet-derived metabolites, choline, trimethylamine N-oxide, and betaine, could predict risk for cardiovascular disease in mice.

5. Nutrition and other OMICs

Many other omics platforms are also a focus of nutriomics research. It was shown that epigenetic alteration could be caused by the nutrition diet during fetal development and could affect the predisposition to lifestyle-related diseases in later life. For example, children of mothers who suffered over- or undernutrition during pregnancy have increased risk of developing obesity, diabetes, hypertension, cardiovascular diseases, etc. Many studies have proven the involvement of epigenetic modifications in such acquired susceptibility. Another promising target of omics approaches in food science is associated with the RNA transcripts that do not encode proteins. MicroRNA (miRNA) is a subtype of these non-coding RNAs. Precursors of miRNAs (pri-miRNA) are long products which are cleaved to yield mature miRNAs of 22 nucleotides in length. Mature miRNAs regulate the levels of



2019-1-BG01-KA203-062371

mRNA degradation, mRNA translation, and even gene transcription. In the case of human cells, about 1,000 miRNAs are identified and are supposed to regulate the expression of more than half of the protein-coding genes. Considering the accumulating data supporting the key role of miRNA in the development of diseases and the maintenance of health, the information on the status of miRNAs is no doubt vital for the understanding of the interaction between food components and the body. Global miRNA analysis now can be easily performed by the use of commercial miRNA arrays.

With the introduction of new technologies and acquired knowledge, the number of fields in omics and their applications in various areas are rapidly increasing in the postgenomics era. Such emerging fields - including pharmacogenomics, toxicogenomics, regulomics, spliceomics, metagenomics, and environomics - present promising solutions to combat global challenges in biomedicine, agriculture, and the environment.

References

- Amer B, Baidoo EEK. 2021. Omics-Driven Biotechnology for Industrial Applications. *Frontiers Bioeng. Biotech.*, 9, Art. 613307: 1-19.
- Azizipour N, Avazpour R, Rosenzweig DH, Sawan M, Ajji A. 2020. Evolution of Biochip Technology: A Review from Lab-on-a-Chip to Organ-on-a-Chip. *Micromachines*, 11: 599.
- Canzler S, · Schor J, · Busch W, et al. · 2020. Prospects and challenges of multi-omics data integration in toxicology. *Archives of Toxicology*, 94:371–388.
- Horgan RP, Kenny LC. 2011. ‘Omic’ technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist*, 13:189–195.
- Joseph P. 2017. Transcriptomics in toxicology. *Food Chem Toxicol.*, 109(1): 650–662.
- Karahalil B. 2016. Overview of Systems Biology and Omics Technologies. *Curr. Med. Chem.*, 23, 1-10.
- Karczewski KJ, Snyder MP. 2018. Integrative omics for health and disease. *Nat Rev Genet.*, 19(5): 299–310.



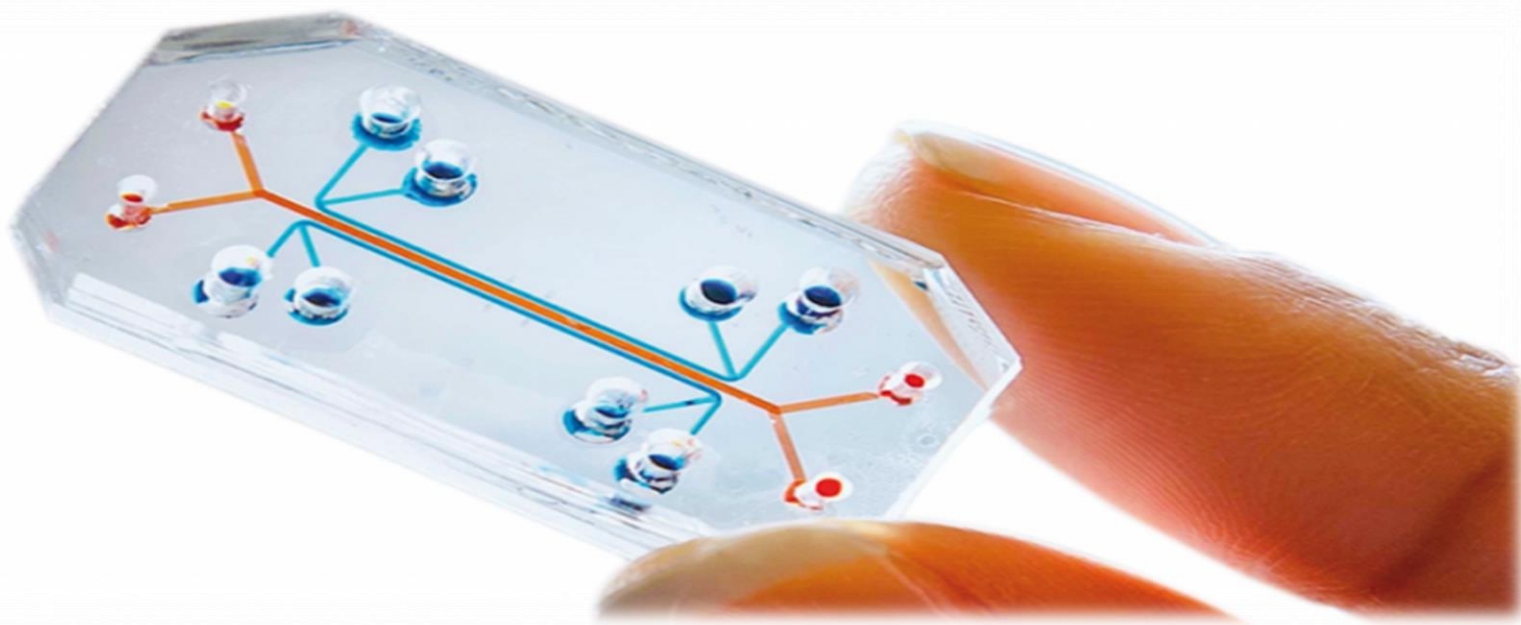
2019-1-BG01-KA203-062371

- Kato H, Takahashi S, Saito K. 2011. Omics and Integrated Omics for the Promotion of Food and Nutrition Science. *J. Trad. Complement. Medicine*, 1 (1): 25-30.
- Kim DH, Kim YS, Son NI, Kang CK, Kim AR. 2017. Recent omics technologies and their emerging applications for personalized medicine. *IET Syst. Biol.*, 11 (3): 87-98.
- Kodzius R, Gojobori T. 2016. Single-cell technologies in environmental omics. *Gene* 576: 701–707.
- Mariman ECM. 2006. Nutrigenomics and nutrigenetics: the 'omics' revolution in nutritional science. *Biotechnol Appl Biochem.*, 44(3):119-28.
- Mousumi D, Prasad GBKS, Bisen PS. 2010. Omics technology. In: *Molecular Diagnostics: Promises and Possibilities*, Dordrech Heidelberg London, Springer, pp 11-31.
- Ryan EP, Heuberger AL, Broeckling CD, Borresen EC, Tillotson C, Prenni JE. 2013. Advances in Nutritional Metabolomics. *Current Metabolomics*, 1: 109-120.
- Sauer UG, Deferme L, Gribaldo L, et al. 2017. The challenge of the application of 'omics technologies in chemicals risk assessment: Background and outlook. *Regulatory Toxicology and Pharmacology*, 91: S14 - S26
- Simões T, Novais SC, Natal-da-Luz T, et al. 2018. An integrative omics approach to unravel toxicity mechanisms of environmental chemicals: effects of a formulated herbicide. *Scientific Reports* 8:11376
- Titz B, Elamin A, Martin F, et al. 2014. Proteomics for systems toxicology. *Computational and Structural Biotechnology Journal*, 11: 73–90.
- Vlaanderen J, Moore LE, Smith MT, et al. 2010. Application of Omics Technologies in Occupational and Environmental Health Research; Current Status and Projections. *Occup Environ Med.*, 67(2): 136–143.
- Yilmaz B, Yilmaz F. 2018. Lab-on-a-Chip Technology and Its Applications. In: *Omics Technologies and Bio-engineering: Towards Improving Quality of Life*, pp. 145, Elsevier Inc.



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.

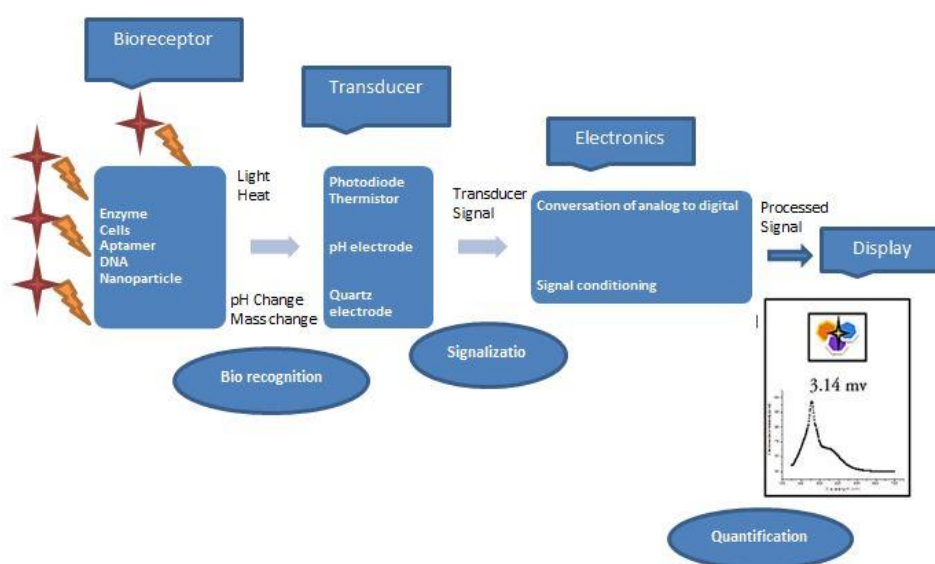


Figure 1. Schematic depiction of biosensor

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.

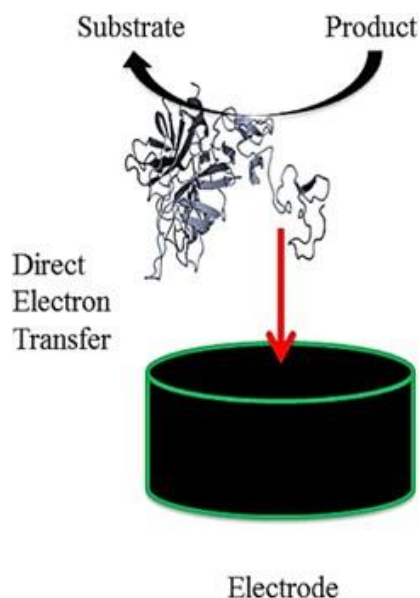


Figure 2. Third generation of glucose biosensor



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

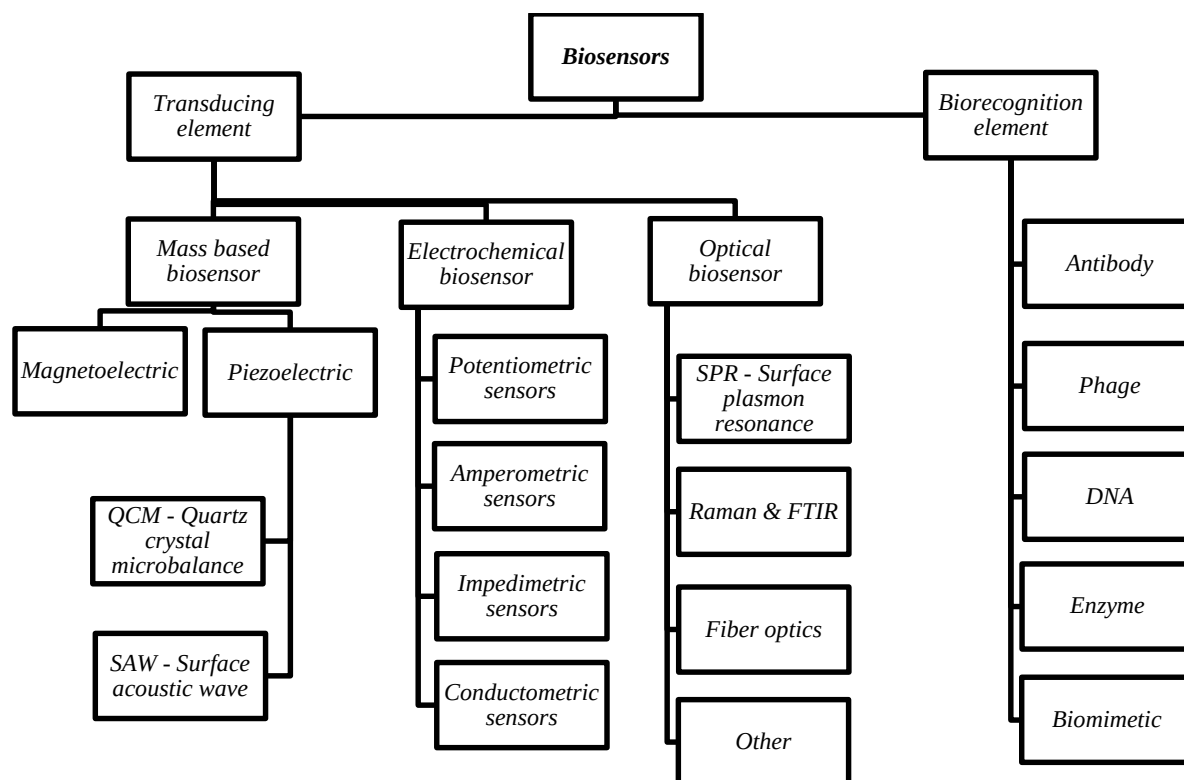


Figure 3. Classification of biosensors based on transducer and bio-recognition elements used in food analysis

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.

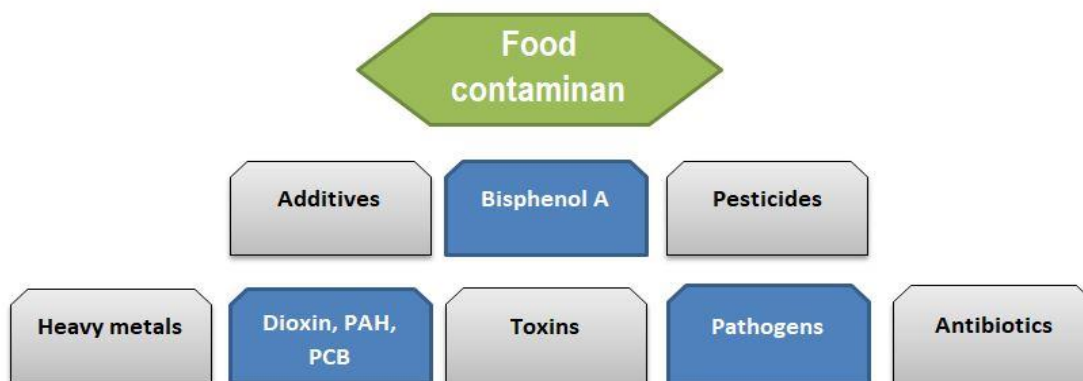


Figure 4. Predominant food contaminants and the target analytes in the food manufacturing industries

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



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+} .

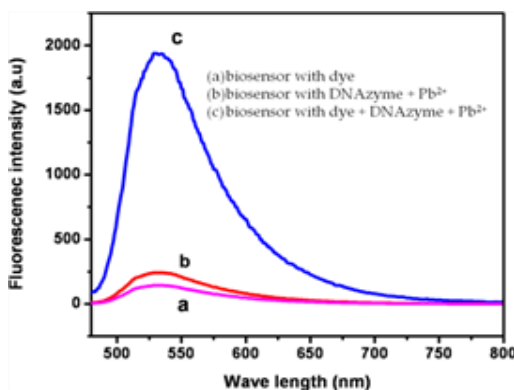


Figure 5. Fluorescence emission spectra for detection of Pb^{2+}



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 ($5 \text{ } \mu\text{g L}^{-1}$), which is lower than the recommended maximum concentration of $7 \text{ } \mu\text{g L}^{-1}$ 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.

- 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, Sezgintürk 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.



**LO5 OPEN ACCESS
SCIENTIFIC RESOURCES:
DIGITAL DATABASES**

Basic level

AUTHOR:

FABIANO CHALHOUB & ZOI GEORGIU



Contents

Open access scientific resources	4
An introduction to «Open Access» resources	4
An introduction to data.....	5
What are “data”	5
A brief history of data	5
Data prior to the invention of computers	6
Data in modern age	6
Understanding the basic vocabulary	7
Terminology	7
What are “Data” in the computer age.....	8
What is “Metadata”	9
What is a “Database”	10
What are “Tables” in a database	10
What are “Columns” in a database.....	11
What is a “Record”	11
What are “Indexes”	12
What is an “Object”	12
Structured data	12
Unstructured data	13
Big data	13
Analytics	14
Repository	14
Basic structure of a database	15
Introduction.....	15
Overview of a database architecture	16
Common database terms	17
What is the difference between major DBMS systems?	18
Databases in the scientific word.....	19



2019-1-BG01-KA203-062371

Introduction to existing databases dedicated to science.....	19
Final thoughts.....	23
References	24



Open access scientific resources

AN INTRODUCTION TO «OPEN ACCESS» RESOURCES

An Open access or OA is a set of principles and a range of practices through which research outputs are distributed online, free of cost or other access barriers. With open access strictly defined (according to the 2001 definition), or libre open access, barriers to copying or reuse are also reduced or removed by applying an open license for copyright.

The main focus of the open access movement is "peer reviewed research literature". Historically, this has centered mainly on print-based academic journals. Whereas conventional (non-open access) journals cover publishing costs through access tolls, such as subscriptions, site licenses or pay-per-view charges, open-access journals are characterized by funding models which do not require the reader to pay to read the journal's contents. Open access can be applied to all forms of published research output, including peer-reviewed and non-peer-reviewed academic journal articles, conference papers, theses, book chapters, monographs and images.

However, when it comes to define "free" access, one has to distinguish "gratis" from "libre".

In order to reflect real-world differences in the degree of open access, the distinction between *gratis* open access and *libre* open access was added in 2006 by Peter Suber and Stevan Harnad, two of the co-drafters of the original Budapest Open Access Initiative (BOAI) definition of open access publishing. *Gratis* open access refers to online access free of charge and *libre* open access refers to online access free of charge plus some additional re-use rights. *Libre* open access is equivalent to the definition of open access in the BOAI, the Bethesda Statement on Open Access Publishing and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. The re-use rights of *libre* OA are often specified by various specific Creative Commons licenses; these almost all require attribution of authorship to the original authors.

The document released in February 2002 by the BOAI contains the following very widely used definition:

- By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution and the only role for copyright in this



2019-1-BG01-KA203-062371

domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

In light of the information above, the use of open source scientific resources must follow the rules commonly adopted. Publishing open source scientific resources must also clearly mention if they are *libre* or *gratis* and must be attributed to the original author.

An introduction to data

WHAT ARE “DATA”

According to the Merriam-Webster dictionary, there are three different definitions of *data*:

1. Factual information, such as measurements or statistics, used as a basis for reasoning, discussion, or calculation
2. Information in digital form that can be transmitted or processed
3. Information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful

In this document we will cover most of the three definitions.

A BRIEF HISTORY OF DATA

Since humans began to communicate, they have experienced the need to retain information for the long term. Keeping information was necessary for our ancestors to ensure their survival. Transmitting information across generations allowed them to keep track of potential dangers, but also to have an inventory of the best places to collect food, the best spots for fishing, the most interesting animals to hunt and where to find the best shelters. All of this information was transmitted orally. With the evolution of knowledge and the invention of writing, they began to store information on indelible media.

Without going into detail about the evolution of the representation of information, some significant examples will be provided that have helped in the structuring of thought, which lead to the discovery of the computer tools we use daily.



2019-1-BG01-KA203-062371

DATA PRIOR TO THE INVENTION OF COMPUTERS

As human societies emerged, collective motivations for the development of writing were driven by pragmatic exigencies. These include organizing and governing societies through the formation of legal systems, contracts, deeds of ownership, taxation, trade agreements, treaties, census records, keeping history, maintaining culture, keeping track of scientific discoveries, codifying knowledge through curricula and lists of texts that are artistically exceptional or deemed to contain foundational knowledge, and many other needs.

For example, around the 4th millennium BC, the complexity of trade and administration in Mesopotamia outgrew human memory, and writing became a more dependable method of recording and presenting transactions in a permanent form.

Cuneiform was one of the earliest systems of writing, invented by Sumerians in ancient Mesopotamia. It is distinguished by its wedge-shaped marks on clay tablets, made by means of a blunt reed for a stylus, as demonstrated in Fig. 1.



Figure 1. Cuneiform writing

Over time, the development of knowledge, the multiplication of information, the limitation of human memory, the necessity of writing and keeping record of huge quantities of information has become essential. However, despite keeping record of almost every kind of information or data on various media, it became more and more complex to retrieve it in a simple manner. One had to read tens of reports and books to be able to synthesize on a subject.

DATA IN MODERN AGE

Today, the quantity of data produced every year and kept digitally, e.g., to-do lists, recipes, reminders, logbooks, maps, photos, e-mails, scientific data, political reports, videos, etc. is so exponential that it creates the need to structure the way we can retrieve these phenomenal quantities.

Computers gained popularity and became cost effective to use by individuals and private companies in the early 80's. However, the 60's can be considered as the new era in the field of databases. The introduction of the term "database" coincided with the availability of direct-access storage or DAS, from the mid-60s onward. This new technology represented a contrast with the past punch cards and the tape-based systems, allowing shared interactive use rather than daily batch processing. Two main data models were developed – network model "CODASYL" (Conference on Data System Language) and hierarchical model "IMS" (Information Management System).



2019-1-BG01-KA203-062371

The first generation of database systems was “navigational, in opposition to the sequential access due to the previous technologies used to store data, i.e. tapes and punch cards. Applications typically accessed data by following pointers from one record to another. Storage details depended on the type of data to be stored.

Adding an extra field to a database required rewriting the underlying access/modification scheme. Emphasis was on records to be processed, not the overall structure of the system. A user would need to know the physical structure of the database in order to query for information. One database that proved to be a commercial success was the “SABRE” system that was used by IBM to help American Airlines manage its reservations data. This system is still utilized by the major travel services for their reservation systems.

In modern Information technology, confusion has always existed among users between databases and internet web search engines accessed by browsers. A database usually contains structured data, in contrast to the World Wide Web (www), which usually contains unstructured data. Even if retrieving information from both databases and “www” are seamless and look similar, the content and the way queries are addressed are completely different. Structured and unstructured data will be explained later in this document.

Understanding the basic vocabulary

TERMINOLOGY

Like any other science, computer science has its own language. In order to fully comprehend the information that will be provided in this document, it is essential to become familiar with the vocabulary related to this topic.

Moreover, the communication with a DBA (Database Administrator) will be eased. When a Biochemist will have to express his needs in terms of structuring or managing data in a Database, he will be tempted to use his own technical language. Then the DBA will have to understand the request and transform it into a computer language, which will be understandable by biochemists.

2019-1-BG01-KA203-062371

What are “Data” in the computer age

As mentioned in section 2.1, according to the domain that is being referred to, *data* might have different meanings. In the case of computing and databases, *data* is defined as any sequence of one or more symbols. *Data* requires interpretation to become information. In information technology the “bit” is the smallest quantity of data. A bit is binary. Binary numbers are a representation of numbers using only two digits, 0 and 1 (Fig. 2). It is a base-2 numeral system, i.e.:

- 0 0 0 1 = numerical value 2^0
- 0 0 1 0 = numerical value 2^1
- 0 1 0 0 = numerical value 2^2
- 1 0 0 0 = numerical value 2^3

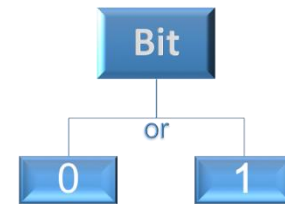


Figure 2. A bit can be 0 or 1

A sequence of “bits” constitutes a “Byte”. Bytes are made of a multiple of 4 bits (a byte of 4 bits is called a Nibble) as in the example above. Today, the byte is a unit of digital information that most commonly consists of eight bits. Historically, the byte was the number of bits used to encode a single character of text in a computer. With a byte of eight bits the maximum decimal number is 256. Historically, the byte is also the unit of computer information or data-storage capacity used to measure the quantity of data (Table 1).

Table 1

Abbreviation	Unit	Value	Size in bytes
b	bit	0 or 1	1/8 of a byte
B	bytes	8 bits	1 byte
KB	kilobytes	10^3 bytes	1,000 bytes
MB	megabyte	10^6 bytes	1,000,000 bytes
GB	gigabyte	10^9 bytes	1,000,000,000 bytes
TB	terabyte	10^{12} bytes	1,000,000,000,000 bytes
PB	petabyte	10^{15} bytes	1,000,000,000,000,000 bytes
EB	exabyte	10^{18} bytes	1,000,000,000,000,000,000 bytes
ZB	zettabyte	10^{21} bytes	1,000,000,000,000,000,000,000 bytes
YB	yottabyte	10^{24} bytes	1,000,000,000,000,000,000,000,000 bytes

Table 2. ASCII Table

BINARY BYTE OF 8 BIT	DECIMAL	HEXADECIMAL	CHARACTER	BINARY BYTE OF 8 BIT	DECIMAL	HEXADECIMAL	CHARACTER	BINARY BYTE OF 8 BIT	DECIMAL	HEXADECIMAL	CHARACTER	BINARY BYTE OF 8 BIT	DECIMAL	HEXADECIMAL	CHARACTER
0 0 0 0 0 0 0 0	0	00	NULL	0 0 1 0 0 0 0 0	32	20	SPACE	0 1 0 0 0 0 0 0	64	41	@	1 0 1 0 0 0 0 0	96	96	`
0 0 0 0 0 0 0 1	1	01	SOH	0 0 1 0 0 0 0 1	33	21	!	0 1 0 0 0 0 0 1	65	42	A	1 0 1 0 0 0 0 1	97	97	a
0 0 0 0 0 0 1 0	2	02	STX	0 0 1 0 0 0 1 0	34	22	"	0 1 0 0 0 0 1 0	66	43	B	1 0 1 0 0 0 1 0	98	98	b
0 0 0 0 0 0 1 1	3	03	ETX	0 0 1 0 0 0 1 1	35	23	#	0 1 0 0 0 0 1 1	67	44	C	1 0 1 0 0 0 1 1	99	99	c
0 0 0 0 0 1 0 0	4	04	EOT	0 0 1 0 0 1 0 0	36	24	\$	0 1 0 0 0 1 0 0	68	45	D	1 0 1 0 0 1 0 0	100	9A	d
0 0 0 0 0 1 0 1	5	05	ENQ	0 0 1 0 0 1 0 1	37	25	%	0 1 0 0 0 1 0 1	69	46	E	1 0 1 0 0 1 0 1	101	9B	e
0 0 0 0 0 1 1 0	6	06	ACK	0 0 1 0 0 1 1 0	38	26	&	0 1 0 0 0 1 1 0	70	47	F	1 0 1 0 0 1 1 0	102	9C	f
0 0 0 0 0 1 1 1	7	07	BELL	0 0 1 0 0 1 1 1	39	27	'	0 1 0 0 0 1 1 1	71	48	G	1 0 1 0 0 1 1 1	103	9D	g
0 0 0 0 1 0 0 0	8	08	BS	0 0 1 0 1 0 0 0	40	28	(	0 1 0 0 1 0 0 0	72	49	H	1 0 1 0 1 0 0 0	104	9E	h
0 0 0 0 1 0 0 1	9	09	TAB	0 0 1 0 1 0 0 1	41	29	)	0 1 0 0 1 0 0 1	73	4A	I	1 0 1 0 1 0 0 1	105	9F	i
0 0 0 0 1 0 1 0	10	0A	LF	0 0 1 0 1 0 1 0	42	2A	*	0 1 0 0 1 0 1 0	74	4B	J	1 0 1 0 1 0 1 0	106	A0	j
0 0 0 0 1 0 1 1	11	0B	VT	0 0 1 0 1 0 1 1	43	2B	+	0 1 0 0 1 0 1 1	75	4C	K	1 0 1 0 1 0 1 1	107	A1	k
0 0 0 0 1 1 0 0	12	0C	FF	0 0 1 0 1 1 0 0	44	2C	,	0 1 0 0 1 1 0 0	76	4D	L	1 0 1 0 1 1 0 0	108	A2	l
0 0 0 0 1 1 0 1	13	0D	CR	0 0 1 0 1 1 0 1	45	2D	-	0 1 0 0 1 1 0 1	77	4E	M	1 0 1 0 1 1 0 1	109	A3	m
0 0 0 0 1 1 1 0	14	0E	SO	0 0 1 0 1 1 1 0	46	2E	.	0 1 0 0 1 1 1 0	78	4F	N	1 0 1 0 1 1 1 0	110	A4	n
0 0 0 0 1 1 1 1	15	0F	SI	0 0 1 0 1 1 1 1	47	30	/	0 1 0 0 1 1 1 1	79	50	O	1 0 1 0 1 1 1 1	111	A5	o
0 0 0 1 0 0 0 0	16	10	DLE	0 0 1 1 0 0 0 0	48	31	0	0 1 0 1 0 0 0 0	80	51	P	1 0 1 1 0 0 0 0	112	A6	p
0 0 0 1 0 0 0 1	17	11	DC1	0 0 1 1 0 0 0 1	49	32	1	0 1 0 1 0 0 0 1	81	52	Q	1 0 1 1 0 0 0 1	113	A7	q
0 0 0 1 0 0 1 0	18	12	DC2	0 0 1 1 0 0 1 0	50	33	2	0 1 0 1 0 0 1 0	82	53	R	1 0 1 1 0 0 1 0	114	A8	r
0 0 0 1 0 0 1 1	19	13	DC3	0 0 1 1 0 0 1 1	51	34	3	0 1 0 1 0 0 1 1	83	54	S	1 0 1 1 0 0 1 1	115	A9	s
0 0 0 1 0 1 0 0	20	14	DC4	0 0 1 1 0 1 0 0	52	35	4	0 1 0 1 0 1 0 0	84	55	T	1 0 1 1 0 1 0 0	116	AA	t
0 0 0 1 0 1 0 1	21	15	NAK	0 0 1 1 0 1 0 1	53	36	5	0 1 0 1 0 1 0 1	85	56	U	1 0 1 1 0 1 0 1	117	AB	u
0 0 0 1 0 1 1 0	22	16	SYN	0 0 1 1 0 1 1 0	54	37	6	0 1 0 1 0 1 1 0	86	57	V	1 0 1 1 0 1 1 0	118	AC	v
0 0 0 1 0 1 1 1	23	17	ETB	0 0 1 1 0 1 1 1	55	38	7	0 1 0 1 0 1 1 1	87	58	W	1 0 1 1 0 1 1 1	119	AD	w
0 0 0 1 1 0 0 0	24	18	CAN	0 0 1 1 1 0 0 0	56	39	8	0 1 0 1 1 0 0 0	88	59	X	1 0 1 1 1 0 0 0	120	AE	x
0 0 0 1 1 0 0 1	25	19	EM	0 0 1 1 1 0 0 1	57	3A	9	0 1 0 1 1 0 0 1	89	5A	Y	1 0 1 1 1 0 0 1	121	AF	y
0 0 0 1 1 0 1 0	26	1A	SUB	0 0 1 1 1 0 1 0	58	3B	:	0 1 0 1 1 0 1 0	90	5B	Z	1 0 1 1 1 0 1 0	122	B0	z
0 0 0 1 1 0 1 1	27	1B	FSC	0 0 1 1 1 0 1 1	59	3C	;	0 1 0 1 1 0 1 1	91	5C	[	1 0 1 1 1 0 1 1	123	B1	{
0 0 0 1 1 1 0 0	28	1C	FS	0 0 1 1 1 1 0 0	60	3D	<	0 1 0 1 1 1 0 0	92	5D	\	1 0 1 1 1 1 0 0	124	B2	
0 0 0 1 1 1 0 1	29	1D	GS	0 0 1 1 1 1 0 1	61	3E	=	0 1 0 1 1 1 0 1	93	5E	]	1 0 1 1 1 1 0 1	125	B3	}
0 0 0 1 1 1 1 0	30	1E	RS	0 0 1 1 1 1 1 0	62	3F	>	0 1 0 1 1 1 1 0	94	5F	^	1 0 1 1 1 1 1 0	126	B4	~
0 0 0 1 1 1 1 1	31	1F	US	0 0 1 1 1 1 1 1	63	40	?	0 1 0 1 1 1 1 1	95	60	_	1 0 1 1 1 1 1 1	127	B5	DEL

2019-1-BG01-KA203-062371

An example of usage is the ASCII (American Standard Code for Information Interchange) table of characters commonly used for alphabetical characters (Table 2). The first 32 characters are called control characters. Initially, they were not designed to represent printable information, but to control devices that use ASCII code, such as printers, or to provide meta-information about data streams, e.g., those stored on magnetic tape.

What is “Metadata”

Metadata, or, put simply, meta-information, is used to reference the data about the data. Having data is not enough to simply put them online. Data are not usable until they can be explained in a manner that both humans and computers can process.



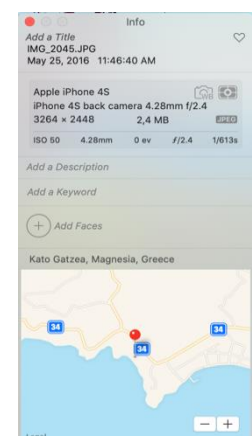
Figure 3. Picture taken in Greece

Back home, you invite your best friends for a party and want to share with them the beauties you have seen during your trip. You start showing the pictures, but you cannot recall which day, at what time and where some of them were taken. This is where the metadata of the pictures can help. In a few words, it is the description of the data. In this example, the picture is the data and the description of the picture is the metadata (Fig. 4).

In Biotechnology, one must understand that metadata are by far more important than data. It is very simple to understand the reason why metadata are a crucial component directly related to data. Imagine an experiment that will lead to a specific result. This experiment, to be valid, must be documented. This documentation should include all the conditions, under which the experiment was conducted. This might include the description of the kind of raw material used, its

Metadata may be implied, specified or given. It includes data relating to physical events, or processes, and will also have a temporal component. In almost all cases this temporal component is implied. It may be slightly tricky to understand, however, the following example will provide a clearer explanation of this term.

Imagine that you are traveling with your favorite smartphone in some paradisiac island. You start taking pictures (Fig. 3) to keep nice records of your trip. A week later, your trip reaches its end and you have to go back home.



**Figure 4.
Metadata of the
picture**

source, in which conditions it was collected, the types of machines to process the experiment, temperature, date, time, etc. For the result of this experiment to be comparable to other results of similar experiments, all the conditions must be similar. Raw data without metadata are useless.

The biggest challenge in Biotech, and any other science, is to standardize metadata. In most of the Biotech databases, this is not respected. One must absolutely be conscious of this phenomenon and thoroughly respect the standards.

What is a “Database”

In general, a database is defined as a collection of data items, such as phone books, price lists, inventory lists, customer’s addresses, etc. Nonetheless, in technical terms, a database is referred to as “a self-describing collection of integrated records”. It implies computer technology, completed with a specific computer language, such as SQL (Structured Query Language).

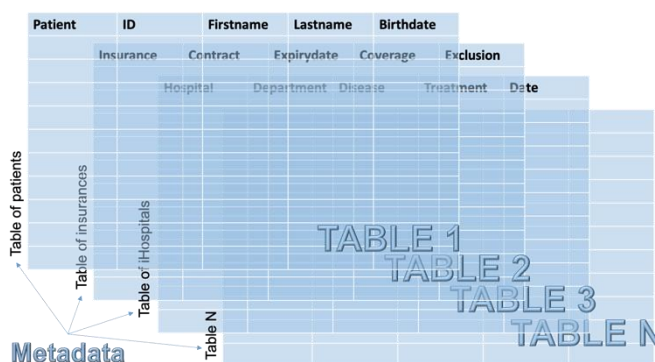


Figure 5. Database partial structure

A database consists of multiple tables (Fig. 5) and of both data and metadata. Metadata is the data that describes the structure of the data within a database. If you know how your data is arranged, then you can retrieve it. Since the database contains a description of its own structure, it is referred to as *self-describing*. The database is integrated because it includes not only data items but also the relationships among them.

The database stores metadata in an area called the data dictionary, which describes the tables, columns, indexes, constraints and other items that make up the database.

Because a flat file system i.e. “Spreadsheet” has no metadata, applications written to work with flat files must contain the equivalent of the metadata as part of the application program.

What are “Tables” in a database

A table is a collection of related data held in a table format composed of columns and rows within a database. It resembles a spreadsheet (Fig. 6).



2019-1-BG01-KA203-062371

What are “Columns” in a database

A column is a set of data values, all of a single type, in a table. Columns define the data in a table. Most databases allow columns to contain complex data like images, whole documents or even video clips. Therefore, a column allowing data values of a single type does not necessarily mean it only has simple text values. Some databases go even further and allow the data to be stored as a file on the operating system, while the column data only contains a pointer or link to the actual file. This is done for the purpose of keeping the overall database size manageable – a smaller database size means less time taken for backups and less time required to search for data within the database.

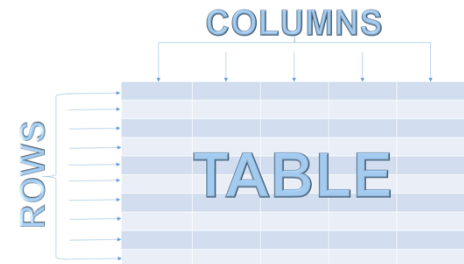


Figure 6. Table with rows and columns

In a table, each column is typically assigned a data type and other constraints, which determine the type of value that can be stored in that column. For example, one column might accept email addresses, and another might accept phone numbers with a constraint of 10 digits.

What is a “Record”

A record is a representation of a physical or conceptual object. Say, for example, that you want to keep track of the customers of a business. You assign a record for each customer. Each record has multiple attributes, such as name, address, and telephone number. Individual names, addresses and so on are the data.

What are “Indexes”

Structured data are stored in the form of records in a database. Every record has a key field, which helps it to be recognized uniquely, i.e. the ID of a patient. No other patient can have the same ID number, but another patient may have the same first name and last name.

Indexing a database is a technique to efficiently retrieve records from the database files, based on some attributes on which the indexing has been performed. To make it simple, indexing in database systems is similar to what we usually see in books. At the beginning or the end of a book, an index may be found (which is different from a table of contents), which provides all the page numbers for a specific topic. For example, an Atlas may be divided into chapters containing maps, chapters containing data on population and chapters dedicated to countries production or agricultural data. If you are looking for a specific country and you would like to have an overview of all the data regarding this specific country, the index might be very helpful as it will show you the page related to that country in each chapter (Fig. 7).

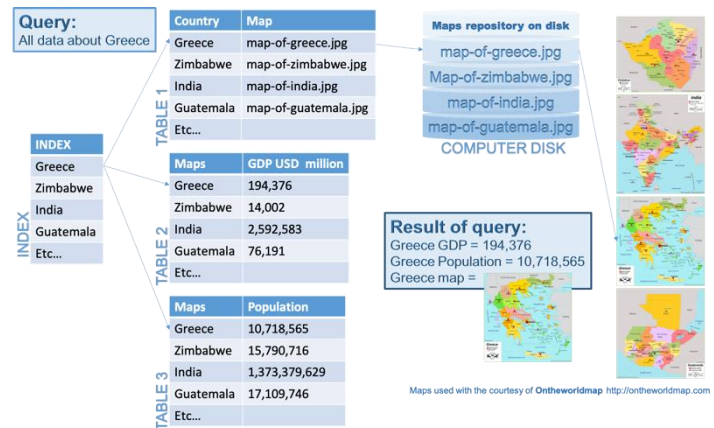


Figure 7. Example of index

What is an “Object”

In computer science, an *object* can be a variable, a data structure, a function, or a method and, as such, is a value in memory referenced by an identifier. In the relational model of database management, an object can be a table or column, or an association between data and a database entity, such as relating a person's age to a specific person.

Structured data

According to SNIA (Storage Networking Industry association), *structured data* is defined as:

“Data that is organized and formatted in a known and fixed way. The format and organization are customarily defined in a schema. The term structured

Structured data

First Name	Last Name	Citizenship	City	Etc...
Fyodor	Dostoevsky	Russian	Moscow	...
Albert	Camus	French	Mondovi	...
Dante	Alighieri	Italian	Florence	...
Nikos	Kazantzakis	Greek	Kandiye	...

Figure 8. Structured data



2019-1-BG01-KA203-062371

data is usually taken to mean data generated and maintained by databases and business applications.”

Three conditions are needed to describe data as structured:

- It must conform to a data model,
- It must have a well define structure,
- It must follow a consistent order and can be easily accessed and used by a person or a computer program.

Structured data is usually stored in well-defined schemas such as Databases. It is generally tabular with columns and rows that clearly define its attributes (Fig. 8).

SQL (Structured Query language) is often used to manage structured data stored in databases.

Unstructured data

Information that is not organized in a predefined model is called *unstructured data* or *unstructured information*. In computer science, files like text files, photos, video files, audio files and presentations are considered unstructured files. Typically, a PDF file contains unstructured data (Fig. 9).

Unstructured data

Some famous writers like M. Fyodor Dostoevsky, who is Russian origin, born in Moscow wrote the famous book *Crime and Punishment*, are worldwide known. We can also cite M. Dante Alighieri who was born in Florence, Italy, and wrote the non less renown book *The Divine Comedy*. Other countries like Greece have great authors. The citizen of Kandiye in Crete M. Nikos Kazantzakis wrote the internationally known book *Zorba The Greek* which was adapted for the cinema in 1965 and starring Anthony Queen. M. Albert Camus, born in the town of Mondovi in Algeria and French citizen whose book *The Plague* was edited in 1947 unfortunately died at the age of 46 in a car accident.

Figure 9. Extract of a PDF file

It is estimated that 80 to 90% of the worldwide total dematerialized data is unstructured. Usual query algorithms are unable to simply and efficiently extract the required information from an unstructured file, such as in the example of Fig. 9. The same information contained in Fig. 9 can easily be retrieved with a query. However, today, *unstructured data analytics tools* powered by *artificial intelligence* (AI) are available, which were specifically created to access the insights available from *unstructured data* (see 3.1.12 Analytics).

Big data

According to SNIA (Storage Networking Industry association), *big data* is defined as:

“A characterization of datasets that are too large to be efficiently processed in their entirety by the most powerful standard computational platforms available.”



2019-1-BG01-KA203-062371

In other words, *Big Data* refers to huge quantities of structured or unstructured data that cannot be processed by usual software as traditional database query language or any other kind of fetching engine.

Confusion exists concerning the current usage of the terms *Big Data* and *Analytics*. *Big Data* is the information, while *Analytics* is the way to extract the desired information from huge quantities of available information.

Analytics

In computer technology, *Analytics* is a method to extract value from big data.

In the field of healthcare, *Big Data Analytics* has led to many improvements by providing personalized medicine and predictive analytics. As the volume of data is dramatically increasing, traditional databases and search engines will not be able to handle and retrieve specific information. Patient data is generated by MRI's, X-rays, blood tests machines, monitoring sensors and many more sources of data complex to process. Extensive information in healthcare is now in electronic form; it fits under the *big data* umbrella as most of it is unstructured and difficult to use.

Big data in health research is particularly promising in terms of exploratory biomedical research, as data-driven analysis can move forward more quickly than hypothesis-driven research. Subsequently, trends seen in data analysis can be tested in traditional, hypothesis-driven follow-up biological research and eventually clinical research.

Repository

A data repository or data warehouse is a centralized place to store and maintain data. A data repository can consist of one or more *structured data* files, such as databases or *unstructured data* files, which can be distributed over a network and preserved over the long-term.



Basic structure of a database

This section is dedicated to the overview of the main building blocks constitutive of a database.

INTRODUCTION

Since the invention of computers, the amount of data stored and managed electronically has increased drastically. It is estimated that the quantity of data will reach 175 zettabytes (10^{21} Bytes) by 2025 growing from a few petabytes (10^{15} Bytes) in the year 2000. One common way of simplifying the lives of users and making the most of their resources is by storing and retrieving it more efficiently. For example, while a flat file works just fine for storing your personal data, such as an address book or some recipes, it is not as suitable for storing a city phone directory or, more precisely, the genomic data in the Biotech field. In addition, if you want to store several genomic species worth of data, it is very difficult to search and retrieve data from a flat file. Databases offer a solution to this problem by making the storage, handling and retrieval of data much easier.

The software used to manage a database is called a database management system (DBMS). This specialized software acts as a go between to help end users access the database. Usually, users do not interact directly with a database because this may result in its disorganization. Instead, they use a DBMS that reads data from or writes data to the database.

The growing complexity of big quantities of data required some companies to use data management tools based on the relational model, such as the classic RDMBS. RDBMS stands for Relational Database Management System. Nevertheless, major Internet companies, such as Google, Yahoo and Amazon, or all the popular Social Media, each faced challenge in dealing with huge quantities of data in real-time, something that conventional RDBMS solutions could not cope with. That explains the soaring popularity of NoSQL database systems that sprang up alongside.

NoSQL systems are distributed, non-relational databases designed for large-scale data storage and for massively-parallel, high-performance data processing across a large number of commodity servers. They arose out of a need for agility, performance and scale, and can support a wide set of use cases, including exploratory and predictive analytics in real-time. Built by top internet companies to keep pace with the data deluge, NoSQL databases scale horizontally and are designed to scale to hundreds of millions and even billions of users performing updates as well as reads.

Some of the common applications of NoSQL databases are social media, large scale e-mail providers and governmental healthcare systems.

Usually, a social application can scale from zero to millions of users in a few weeks and to better manage this growth, one needs a DB that can manage a massive number of users and data, but can also easily scale horizontally.

In this course, we will focus on DBMS and RDBMS only. These are the two kinds of Databases commonly used in the Biotech world up to date.

OVERVIEW OF A DATABASE ARCHITECTURE

Databases can store all kinds of information, from numbers and text, to email, web content, phone records, biological, geographical data, etc. Databases are officially classified according to how they store this data. Relational databases store data in tables. Object oriented databases store data in object classes and subclasses. We are going to focus on relational databases, as they are most commonly used. However, most of the basic topologies of databases need to have backend servers in order to host the database management system, a storage system attached to the servers to store the structure and the data of the database and, of course, computers, laptops, desktops or terminals as an interface to allow users to access the database, its management system and its content. Also required is a network to exchange between all the hardware components and a Cloud attachment to allow remote users to access the database. Fig. 10 summarizes in a simple way the minimum required to run a database.

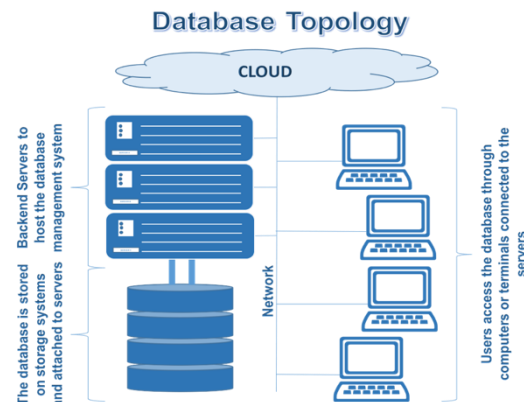


Figure 10. Hardware architecture for a database

Another basic way to describe it, is to show the three level architecture of a database. It is a virtual view of the necessary layers to make a database function properly. Fig. 11 demonstrates the three-level view architecture. It is called the ANSI-SPARC model. Nonetheless, despite the fact that this model never became a formal standard, it presents the idea of logical data independence that has been widely adopted.

Information stored inside a relational database is contained within tables. These tables are composed of rows of data and each row contains fields or columns. In a well-designed database definition, called a schema, only similar data is stored within each table and duplication of columns is kept to a minimum. Developers can connect, or join, data from two tables to link different types of information to each other.

Indexes can be created on fields in the database table to make it easier for the DBMS to retrieve data. Indexes are usually configured for frequently searched columns, like a person's name or a date value. The drawback to using indexes is that they take up storage disk space and can slow things down, if too many of them are maintained, because every time a row in the database is updated, the index also has to be updated.

Most databases support Structured Query Language (SQL), a standard language for interacting with information contained in a database. SQL allows users and applications to interact with specific subsets of data from one or more tables using several statements as SELECT, INSERT, UPDATE and DELETE.

Relational databases also provide a layered approach to storage, allowing the definition of what database objects reside in specific data files and where those data files are placed within the operating system's file structure. On top of managing the physical storage location of database objects, many database systems give some control over how the data is stored within the data files.

COMMON DATABASE TERMS

Certain database terms derive from ways that databases automate write actions. Database developers often automate writing to certain fields or other tables, such as writing a copy of the row being inserted – along with a timestamp or username – to a history or audit table. Most DBMS systems provide several ways to automatically manage database write actions.

Database triggers are the most common method of taking action on data as it is being written to the database. Triggers are usually associated with a particular table and configured to execute at a certain point during a specific write action, such as before or after an update, or after a row is inserted. Triggers

Three-Level Architecture of a Database

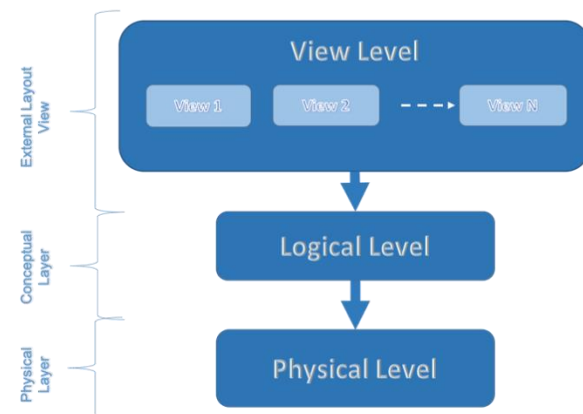


Figure 11. Three-level view architecture



2019-1-BG01-KA203-062371

can be used to format data, populate a column with data derived from existing information, or even write to another table based on the row being inserted or updated.

A stored procedure is another way of interacting with a relational database. Stored procedures are more complex than triggers and are not tied to a single specific table. Typically created by a developer, they use a combination of SQL and a programming language, such as Java or SQL (depending on the database platform). Stored procedures provide developers a lot of control over how data is validated or massaged by an application. A stored procedure could be used to manage how a user logs in to an application. The procedure might first validate the username and password, then log the success or failure of the attempt to another table, along with other information, including the computer name and a timestamp. An alert could even be sent to the user informing them that their password has expired and must be changed.

Functions are simpler than stored procedure, and can sometimes even be used from within SQL queries. Functions are usually used in a database to perform a set of actions that return one or more values, such as calculating the sum of a column for rows that match a certain condition. While these actions can be performed using SQL, building them into a function can make them easier to use in other code. Both functions and stored procedures can perform common actions in a streamlined and consistent manner, easing the workload for database administrators and developers.

WHAT IS THE DIFFERENCE BETWEEN MAJOR DBMS SYSTEMS?

The DBMS is generally driven by what the user applications need to support. That said, here is a brief comparison of the three most widely used platforms.

Microsoft SQL Server is widely used in enterprise applications and integrates easily with other Microsoft tools. Microsoft SQL Server 2019 Express is the latest version of Microsoft's free offering and is often bundled with applications that use SQL Server.

MySQL has been a favorite for open source developers for the better part of two decades. Often used as a back end for open-source blog or content management systems, MySQL has a massive installed base across the globe. In 2008, MySQL AB was acquired by Sun Microsystems, which was itself acquired by Oracle Corp. in 2009, bringing MySQL under the umbrella of one of its largest competitors. However, the MySQL Community Edition remains free and is well supported by the community. MySQL is available for numerous operating systems, including Linux, UNIX, Mac OS X and Windows.

Oracle Database is considered by many to be the standard in enterprise-level database platforms and supports numerous enterprise applications. Oracle Database Express Edition is available free of

charge and is also free to distribute (though it is not technically free software), making it another popular option for developers or hobbyists on Windows or Linux.

Now that you have learned the fundamental database terms and concepts, you are that much closer to speaking the same language as your organization's database developers.

Databases in the scientific world

This part deals with the basics of databases used in the scientific world

INTRODUCTION TO EXISTING DATABASES DEDICATED TO SCIENCE

This section is dedicated to the overview of the most common open access databases used in science.

Continuous developments in the fields of biotechnology and information technology have led to the exponential growth of data. Studies conducted by researchers at the European Bioinformatics Institute (EMBL-EBI) have demonstrated that this growth of information is doubling approximately every year. These extensive amounts of data are stored, organized and constantly updated in scientific databases, where they are readily available for scientists, including biologists and bio-informaticians, to use for research purposes. The information available in biological databases is obtained from a range of scientific fields, including metabolomics, microarray gene expression and proteomics. Apart from storing, organizing and sharing huge volumes of data, the main goal of biological databases is to offer web application programming interfaces (APIs) for computers to exchange and integrate data from many different database resources via an automated method.

Biological databases can be defined as data collections, which are structured in such a way making their contents easy to explore, handle and update. Examples of such databases are presented in Fig. 12. In 1972, the first protein structure database, known as the Protein Data Bank (PDB), was created. This database originally contained only 10 entries, which has now expanded to contain more than 10,000 entries, signifying the rapid growth of biological data. A biological database may contain several types of data, including protein sequences, textual descriptions, attributes, and tabular data. Generally, they can be divided into primary, secondary, and



Figure 12. Example of names of Biotech databases



2019-1-BG01-KA203-062371

composite databases. Primary databases include data about the sequence or structure alone, whereas secondary databases include data originating from the primary database. Data, such as the conserved sequence and active site residues of protein families, can be found in secondary structure databases. Furthermore, entries of the PDB, which is a primary database, can be found in secondary structure databases, stored in an organized way.

Broadly speaking, biological databases can be categorized into sequence, structure, and pathway databases:

- **Sequence databases:** The most commonly used biological databases. These include protein and nucleotide sequence databases, which contain wet lab results and are the main source for experimental results. GenBank and EMBL are examples of sequence databases.
- **Structure databases:** These databases contain information regarding protein structure and molecular interactions. PDB is an example of a structure database.
- **Pathway databases:** These databases are based on data derived from the comparative study of metabolic pathways. The Kyoto Encyclopedia of Genes and Genomes (KEGG) and Biocyc are two indicative pathway databases.

A typical search in a nucleotide sequence database may, for example, generate data concerning the scientific name of the source organism from which it was isolated, contact name, the input sequence with details of the molecule type and, frequently, literature citations related to the sequence.

Certain tools have been developed to facilitate scientists in data processing and retrieval from biological databases. These tools, which are termed bioinformatics tools, are software programs created for the extraction of meaningful data from the vast number of biological databases and for conducting sequence or structural analysis. Bioinformatics tools are used to obtain data from genomic sequence databases and for the visualization, analysis and retrieval of data from proteomic databases. These tools are largely divided into:

- **Homology and similarity tools:** These tools are used for the detection of similarities between the sequences of unknown structural and functional sequences, whose function and structure are already known.
- **Protein function analysis tools:** Programs applied for the comparison of one protein sequence to a secondary (or derived) protein, which permit the estimation of the biochemical function of a query protein.
- **Structural analysis tools:** These tools allow the comparison of structures with the known structure databases and the establishment of the 2D/3D structure of a protein.
- **Sequence analysis tools:** Programs used for the additional, more comprehensive assessment of a query sequence, involving evolutionary analysis and identification of mutations.



Biological databases may also be categorized, based on the scope of data coverage, into:

- Comprehensive databases: These databases comprise various types of data from a number of species. Examples of comprehensive databases are GenBank and EMBL.
- Specialized databases: These databases include particular types of data or data from particular organisms. An example of specialized databases is WormBase, which contains information on nematode biology and genomics.

In relation to the level of biocuration, which is defined as the activity of organizing, demonstrating and making biological information readily available to both humans and computers, biological databases are classified as primary and secondary or derivative databases. Primary databases consist of raw data as archival repository, while secondary or derivative databases consist of curated information as added value. In regard to the method employed for curating the data, biological databases may be further classified as expert-curated databases or community-curated databases, which are curated in a co-operative way by numerous researchers.

Additional categorization of biological databases can also be made based on data type. The data types that accordingly classify databases include DNA, RNA, protein, expression, pathway, disease, nomenclature, literature and standard and ontology. Some of the most important and widely used biological databases are the following: *GenBank*, the *UCSC Genome Browser* and *Ensembl*, which are sequence databases/portals; *WormBase* and The Arabidopsis Information Resource (TAIR), which are model organism databases; and the *PDB*, Online Mendelian Inheritance in Man (OMIM), *MetaCyc* and *KEGG*, which are characterized as non-sequence-centric databases.

Data manipulation is an essential part of the experimental process of all studies, regardless of their scale. The online availability of biological data combined with the decreasing costs of automated genome sequencers have made it possible for small biology laboratories to become big-data generators. Even if a laboratory is not equipped with such instruments, it can still become a big-data user by gaining access to public repositories containing biological data, such as the US National Center for Biotechnology Information in Bethesda. A large part of the construction in big-data biology is virtual, based on cloud computing, in which data and software are located in massive, off-site centers that can be accessed on demand. Therefore, it is not necessary for users to purchase their own hardware. The cloud computing system allows potential users to create virtual spaces for data, software and results that are freely accessible by everyone, or to keep spaces locked up behind a firewall permitting access to a chosen group of collaborators.

The use of biological databases can be advantageous in several research areas. For example, databases may aid experimental design by allowing the automatic analysis and easy processing of experimental data and making the examination of experimental results simple and quick. Drug discovery is another area that may be simplified by using databases. In this specific area, databases can be scanned in order to find new candidates for drugs by training a classifier on a dataset where functioning and non-



2019-1-BG01-KA203-062371

functioning drugs have been identified. Moreover, machine learning techniques may be applied to design virtual assays that are able to identify promising new drugs, which can subsequently be analyzed in a laboratory setting. (REF. 4) And most importantly, new scientific experiments can be carried out and new results generated by analyzing existing data sets.

Without the existence of databases, sharing and integration of large quantities of data would be virtually impossible. Although many life scientists have advanced computational skills, a large percentage are not familiar with developing or adapting the relevant software. Nonetheless, the involvement of life scientists in this process is crucial, since they can provide feedback to computer science specialists focusing on different needs and approaches to science. The ability to have access to the actual data sets originally used in a specific study provides researchers with the opportunity to reproduce and expand on such study. This is why it is important for data to become freely available to scientists at any time without restrictions, a notion supported by Open Science and numerous related initiatives. One of these initiatives is known as ELIXIR, a project designed to help scientists across Europe safeguard and share their data and to reinforce current resources, including databases and computing facilities, in individual countries.

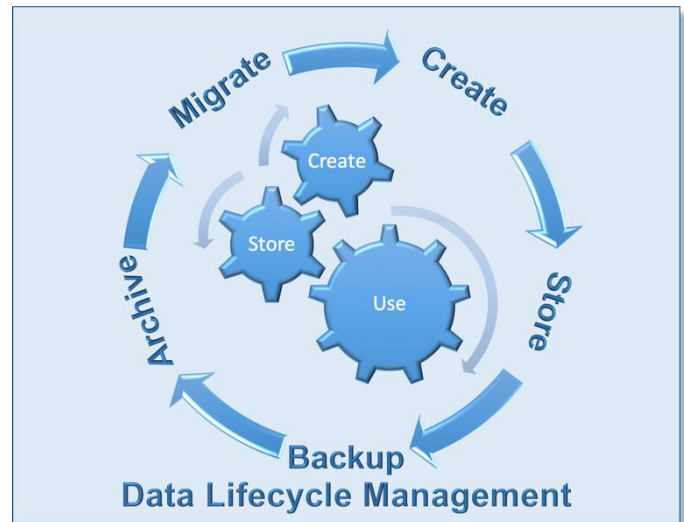
Although the creation of biological databases has brought about many benefits, such as the promotion of scientific quality production enabled by networking, they still require improvement in terms of knowledge optimization. It is crucial to manage transdisciplinary knowledge in such a way that will lead to an increase in its quality and quantity. Data heterogeneity is another common issue faced in biological data integration. In the field of biology, several different methods exist for the representation of similar data. This complicates data integration and processing, which, in turn, makes it harder to acquire unified views of such data. An example of this problem is the use of various alternate names when referring to genes, regardless of the existence of full guidelines issued in 1979 proposing the adoption of gene nomenclature standard, leading to difficulties in data sharing. The implementation of standards enables the re-use of data, however, their absence causes significant loss of productivity and contributes to a decrease in data accessible by researchers. Therefore, it is imperative to find a solution to this matter in order to eliminate the challenges faced by scientists when using biological databases to conduct their research.



2019-1-BG01-KA203-062371

FINAL THOUGHTS

Dealing with data implies a drastic discipline to keep access on a long term to the stored information. Technology evolves, which means that the hardware and software used today is not the standard of tomorrow. This means that to be able to read any data written today we will have to execute two different kinds of migrations. A logical migration and a technological migration. Logical migration is related to the kind of format in which the data is stored. Technological migration is related to the kind of hardware used. As an example, if you try to open a Word file written in 1993 with Word version 6 with the latest version Word 2019, it will not work. This example shows a lack of logical compatibility. To avoid this issue and keep an ascended compatibility, the file should have been migrated by the time to the latest version in order to keep it up to date and readable with the latest versions of software.



The same thing applies to hardware, i.e. servers, storage, networks, etc... Another example could be the kind of server and operating system used to run a database. In case you decide to change your hardware and to migrate from, let's say, Windows to UNIX, a different kind of hardware will be needed to run UNIX and a different version of database to run on UNIX. Windows run on Intel based platforms (and Intel like) and Unix runs on SPARC based platforms, which means that you will have to migrate to a UNIX - SPARC compatible version of the database.

Keeping in mind this constant evolution of hardware, operating systems, software and formats, performing the appropriate logical and technological migrations on time could save you a lot of time and troubles.

Last but not least, it is important to keep backing up your data. Once every three to six months, perform a restore test to see if you are capable of retrieving your backups. This is crucial for two reasons:

- It will keep you up to date on how to restore your data
- It is the best testing method to see if your data was properly backed up



References

- Baxevanis AD, Bateman A. 2015. The importance of biological databases in biological discovery. *Curr Protoc Bioinformatics.*, 50(1):1.1.1-1.1.8.
- Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2014. GenBank. *Nucleic Acids Res.*, 42:D32–D37.
- Brooksbank C, Bergman MT, Apweiler R, Birney E, Thornton J. 2014. The European Bioinformatics Institute's data resources 2014. *Nucleic Acids Res.*, 42:D18–D25.
- Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, et al. 2016. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.*, 44(D1):D471-80.
- Figueiredo MSN, Pereira AM. 2017. Managing knowledge – the importance of databases in the scientific production. *Procedia Manuf.*, 12:166–73.
- Harris TW, Baran J, Bieri T, Cabunoc A, Chan J, Chen WJ. 2014. WormBase 2014: new views of curated biology. *Nucleic Acids Res.*, 42:D789–D793.
- Howe D, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W, et al. 2008. Big data: The future of biocuration: Big data. *Nature.*, 455(7209):47–50.
- Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. 2021. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.*, 49(D1): D545–51.
- Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, et al. 2019. The BioCyc collection of microbial genomes and metabolic pathways. *Brief Bioinform.*, 20(4):1085–93.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res.*, 12(6):996-1006.
- Lapatas V, Stefanidakis M, Jimenez RC, Via A, Schneider MV. Data integration in biological research: an overview. *J Biol Res (Thessalon).* 2015;22(1):9.
- Marx V. 2013. Biology: The big challenges of big data: Biology. *Nature.*, 498(7453):255–60.
- Nature Structural Biology 10, 980. 2003; doi: 10.1038/nsb1203-980



2019-1-BG01-KA203-062371

Oliveira AL. 2019. Biotechnology, big data and artificial intelligence. *Biotechnol J.*, 14(8):e1800613.

Razvi SRH, Rampogu S. 2016. Bioinformatics in the present day. *MOJ proteom bioinform* [Internet]., 3(1):11–2. Available from: <http://dx.doi.org/10.15406/mojpb.2016.03.00073>

Toomula N, Kumar A, Kumar D S, Bheemidi VS. 2012. Biological databases- integration of life science data. *J Comput Sci Syst Biol.*, 04(05):087-092. Available from: <http://dx.doi.org/10.4172/jcsb.1000081>

Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. 2020. Ensembl 2020. *Nucleic Acids Res.*, 48(D1): D682–8.

Zou D, Ma L, Yu J, Zhang Z. 2015. Biological databases for human research. *Genomics Proteomics Bioinformatics.*, 13(1):55–63.

Web sources:

https://en.wikipedia.org/wiki/Airline_reservations_system

<https://en.wikipedia.org/wiki/CODASYL>

https://en.wikipedia.org/wiki/Database_administrator

https://en.wikipedia.org/wiki/IBM_Information_Management_System

<http://www.redbooks.ibm.com/abstracts/sg245352.html>

https://en.wikipedia.org/wiki/Navigational_database

<https://en.wikipedia.org/wiki/SQL>

<https://omim.org/>

<https://www.ascii-code.com>

<https://www.budapestopenaccessinitiative.org/>

<https://www.merriam-webster.com/dictionary/data>



2019-1-BG01-KA203-062371

<https://www.snia.org/education/online-dictionary/term/big-data>

<https://www.snia.org/education/online-dictionary/term/structured-data>

www.arabidopsis.org/aboutarabidopsis.html



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.



**LO6 SOCIAL AND
POLITICAL ACCEPTABILITY
OF MODERN
BIOTECHNOLOGY TOOLS**

Basic level

AUTHOR:

TRAYANA NEDEVA & ANNA KUJUMDZIEVA



Contents

Modern Biotechnology at a Glance.....	3
Practical Applications of Modern Biotechnology	4
Social Attitude to Modern Biotechnology	5
Societal Impact.....	6
Establishment and spreading of knowledge	7
Science and Technology (S&T) impact on society	7
Perception of modern biotechnology applications	8
Benefits and Risks from Biotechnology Applications	12
Benefits vs. risks societal dispute.....	12
Risk acceptance.....	13
Risk Management.....	15
Social benefits of modern biotechnology.....	15
Acceptance and diffusion of modern biotechnology.....	16
Possible negative impacts of modern biotechnology	19
Socioeconomical Impacts.....	19
Impacts on biodiversity and sustainable agriculture	19
Impacts on public welfare	20
Biosafety issues of modern biotechnology.....	20
Conclusions	21
References	22



Modern Biotechnology at a Glance

Biotechnology is a modern advanced technology grounded on science disciplines like molecular biology, biochemistry, cell biology. It accepts a lot of new techniques such as cell and tissue cultivation, cell fusion, gene recombination, microbial fermentation, and also duplication and recombination of organisms at a cellular, chromosomal, and gene level. In this way, it can make organisms answering to human needs, produce new products and breed new plants having novel features of high output and stress resistance. Modern biotechnology is composed of four technological systems. They are genetic, cell, enzyme, and fermentation engineering, of which genetic engineering and cell engineering are the most advanced ones.

Besides the great social and economic benefits of modern biotechnology, there is a possibility it to harm human health and the environment. This can result in many socio-economic problems, such as destroying original social and economic patterns, posing threats to biodiversity and traditional crop varieties, injuring countries or communities' socioeconomic welfare, violating traditional ethical, moral, and religious values.

The modern biotechnological practice in the production of fiber, pharmaceuticals, and food has a powerful development of the late 20th and beginning of the 21st century. This emerging technology is often accepted as the next technological revolution possessing the potential to change fundamentally the way of societal arrangement regarding production and distribution of goods. Considerable investments have been already made in biotechnological research for new products' development. Currently, science and technology are expected to bring consumers a big variety of genetically modified (GM) products. Besides, many GM products have already entered the food distribution chains. Despite its promise to bring significant benefits to society, public acceptance of modern biotechnology is quite diverse at the world dimension. To enable biotechnology to contribute major benefits for human beings, the international society has paid serious attention to the biosafety of the overall biotechnological production to promote its better acceptance.



Practical Applications of Modern Biotechnology

Nowadays, biotechnology has been classified into few categories based on its applications. They are depicted in Fig. 1.

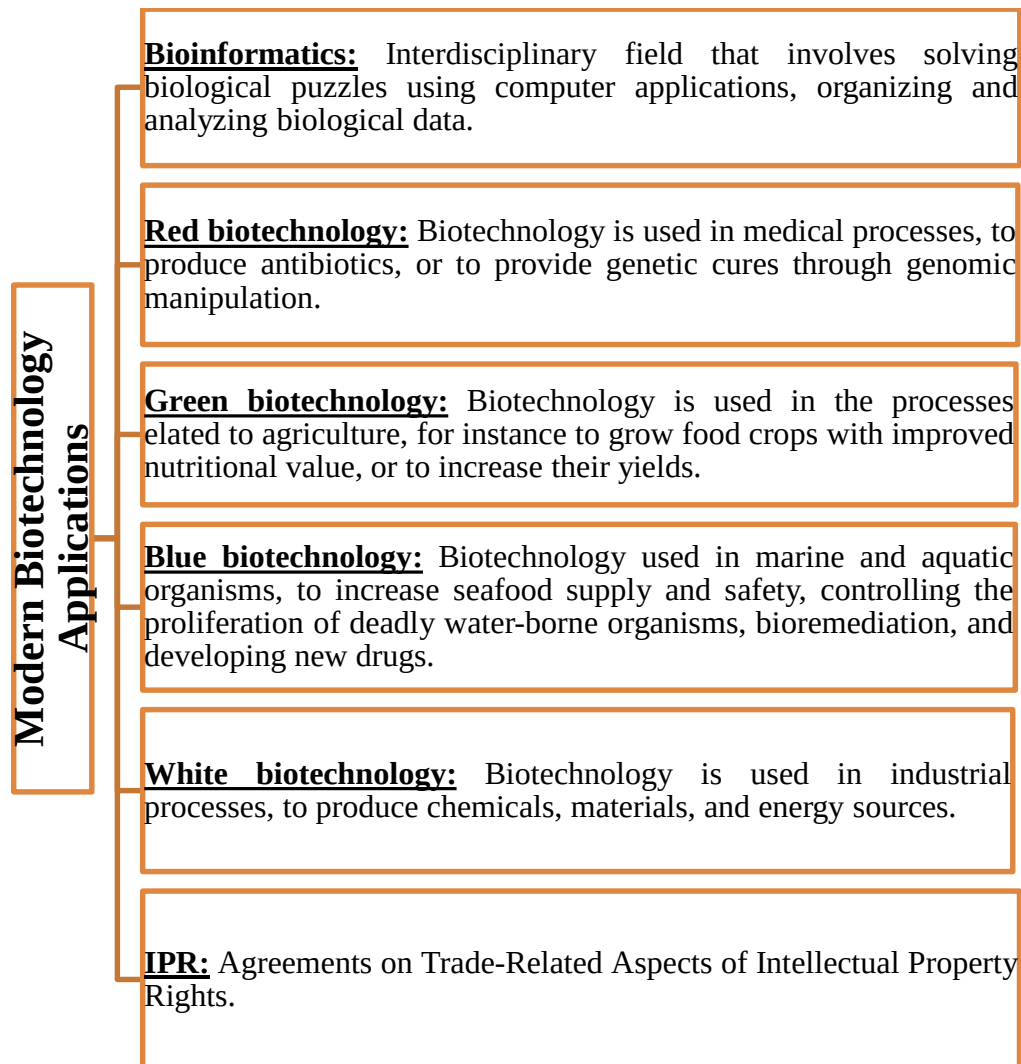


Figure 1. Practical applications of modern biotechnology



Social Attitude to Modern Biotechnology

In general, biotechnology intends to emphasize the potential benefits to society through reduction of hunger and malnutrition, suspension and healing of diseases, and increase of health and general well-being. On the other hand, there is an opinion that GM products are used as a needless interference with nature and can cause unknown and potentially disastrous consequences. At the same time in the U.S., GM crops entered the grain supply chains without raising major public concern. Besides, agricultural biotechnology is facing significant opposition in Europe and many other developed countries. Regarding the obvious public troubles about the perceived risks for humans and the environment, European Union (EU) accepted rather limited regulations on all transgenic crops in any part of the EU food system. At a global level, the situation is even more diverse: in the U.K., objection movements throw down GM crops on different occasions. Until recently, Brazil and India refused to approve any GM crop. Due to similar consumer concerns, some fast-food supply chains decided not to use GM potato in their products.

The reason against the use of gene technologies in agricultural production is grounded on the fact that some people and institutions see risks to humans and the environment, while others oppose it minding moral, ethical, and social concerns. Biotechnology frequently is running down because of its tools' application in plants and animals, especially gene transfer across species, based on understanding "realms of God" and against "Law of Nature". Also, there is an opinion that since genes are naturally occurring parts, which are an object of discovery (not of invention), patent ownership of genetic findings and processes is morally and ethically insolvent.

Public discussions on agricultural biotechnology also pose some social and political debates. There is a talk concerning the application of modern genetic technologies in production of commodities in developed countries that are currently imported to developing countries. It is claimed that such developments will have significant negative effects on the poverty situation in the Third world and will lead to global instability. At the same time, other researchers have the opposite opinion. Another source of care is the possibility that farmers will become steadily dependent on multinational corporations for their "means of production" bringing harmful economic, social and political effects.

The importance of this subject defines the need for a full understanding of public interests and concerns regarding an agreement between private and public decisions concerning food biotechnology. Consumer acceptance of biotechnology is strongly related not only to expected risks and benefits associated with GM products but also to their moral and ethical dimension. Further, public views about multinational corporations, trust in government, science, and technology development also reveal their attitudes towards biotechnology. It is found that consumers' cognitive factors (e.g., levels of risk aversion, opinions about GM foods) impacted the acceptance of GM food products, while the socio-economic factors did not have significant effects. Thus, the public perceptions of biotechnology possess a lot of dimensions and are possibly affected by multiple forces, preferences, and events. For instance, positive benefits (e.g., nutritional benefits from improved/new products, environmental benefits via reduced use of pesticides, etc.) can encourage consumer acceptance of food biotechnology. On the other hand, sensibility about risks to humans and the environment is expected to have adverse effects on public

acceptance of GM products. Thus, the activities to analyze and determine the factors influencing consumer attitudes towards biotechnology have to follow the procedure, denoted in Fig. 2.

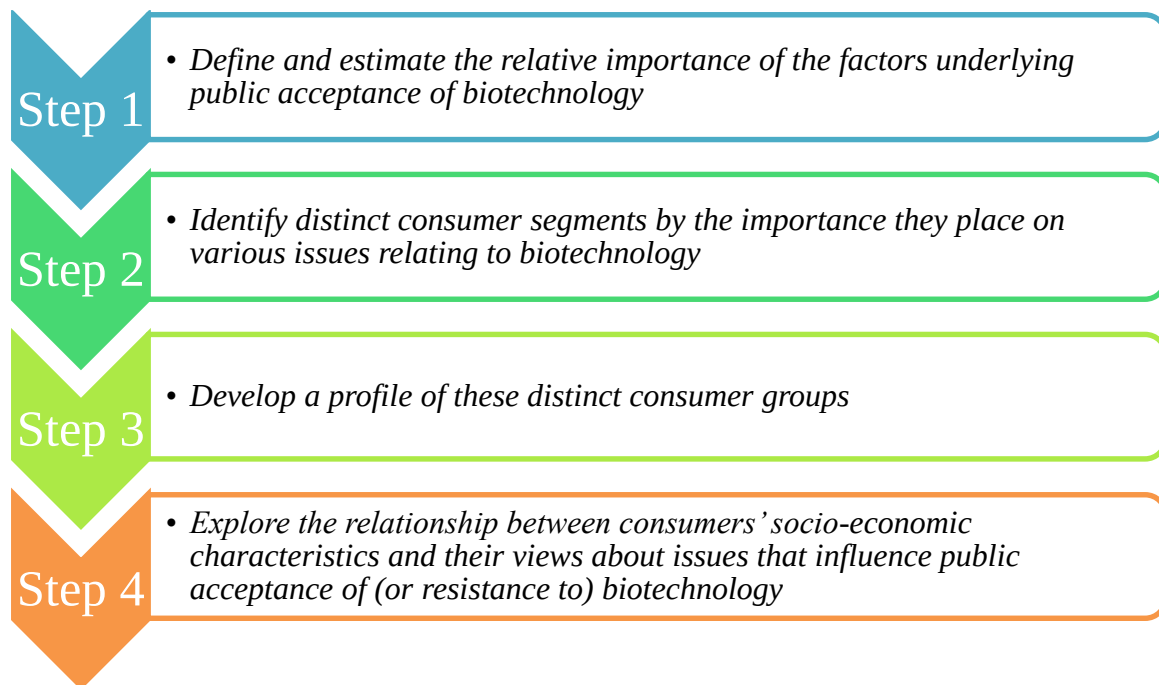


Figure 2. The procedure for analysis and determination of the factors influencing consumer attitudes towards biotechnology

SOCIETAL IMPACT

To evaluate the impact of modern biotechnology on society, it is not enough only to measure it by specific indicators. When the focus is on such dimensions like increase the number of patents on genetically modified organisms (GMOs), this will not indicate the dynamic forces shaping the impact of biotechnology on industrial and societal structures. The impact analysis is combining complex problems and has a variety of diverse inputs, which possess both short- and long-term effects and influence the results of the problem or use of technology, in an unrespecting way. To realize the sense about the nature of links between societies, individuals, and the balance of forces between social groups facing new technologies is of basic importance.



ESTABLISHMENT AND SPREADING OF KNOWLEDGE

A starting point is the result of the technological impact, which depends upon the establishment and spreading of knowledge. It is regarded as preferred values, social practices, and norms of behavior of involved actors in the development and use of biotechnology. The knowledge extends beyond the collection of information on biotechnology, such as techniques, risks, etc. It also embraces the way they are shown and commented on by individuals. Risk and safety for the public cannot be assayed in the same manner as scientific risk is done. To account for these differences, it is necessary to explain how trust is created between consumers and manufacturers. In brief, such a definition of knowledge gives a piece of analysis on the impact of biotechnology and more broadly, environmental policies and practices.

From a conceptual point of view, this approach could be treated as a knowledge-based discourse. It is useful considering the expression of the biotechnological impact in the sense of the educational, social, and political background. The environmental discourses of biotechnology knowledge, and more specifically, how it is constructed, accepted, and perhaps most importantly, spread and sustained within society are taken into account as well.

SCIENCE AND TECHNOLOGY (S&T) IMPACT ON SOCIETY

In the EU there is a discussion on biotechnology and especially concerning genetic engineering and impacts arising from this technology. This issue has to be addressed strongly linked to the result of social interactions. Therefore, the way of realization of these dynamic forces should account for the interrelations and the form of resulting complexity between actors.

Another point to estimate, are two major topics when examining potential technology impacts. These are i) the immediate impact of the use and development of the technology and ii) the technology transcending risks. That means dependence on the long-term application of the technology and how to achieve and sustain a dominant position in its impact upon society from education, political and social perspectives. This could be sustained generally through more concern about the long-term consequences of S&T activities noted below:

- *Publication of educational efforts in society:* Performance of intensive efforts to include scientists, key public opinion leaders, and media in an existing debate on biotechnology and especially about genetic engineering. The core objective is to avoid an informational vacuum, which can facilitate anti-biotechnology activity and generate negativism in knowledge structures.

- *Role of regulatory policy in society:* Focusing on effective risk communication and understanding the fact that the public perception of risk can be very different from that of the scientific community.

PERCEPTION OF MODERN BIOTECHNOLOGY APPLICATIONS

Three major topics are identified and based around: i) the science, ethics, and gender of biotechnology ii) the International political economy, trade, and the environment; iii) the state and system security and warfare through biotechnology. These general themes are coming from consideration for the international relations of biotechnology. They imply the concept of power in the International Political Economy (IPE) to frame the differing discourses and complex issues of technology impacts. Based on these topics the several indicators are regarded (Fig. 3):

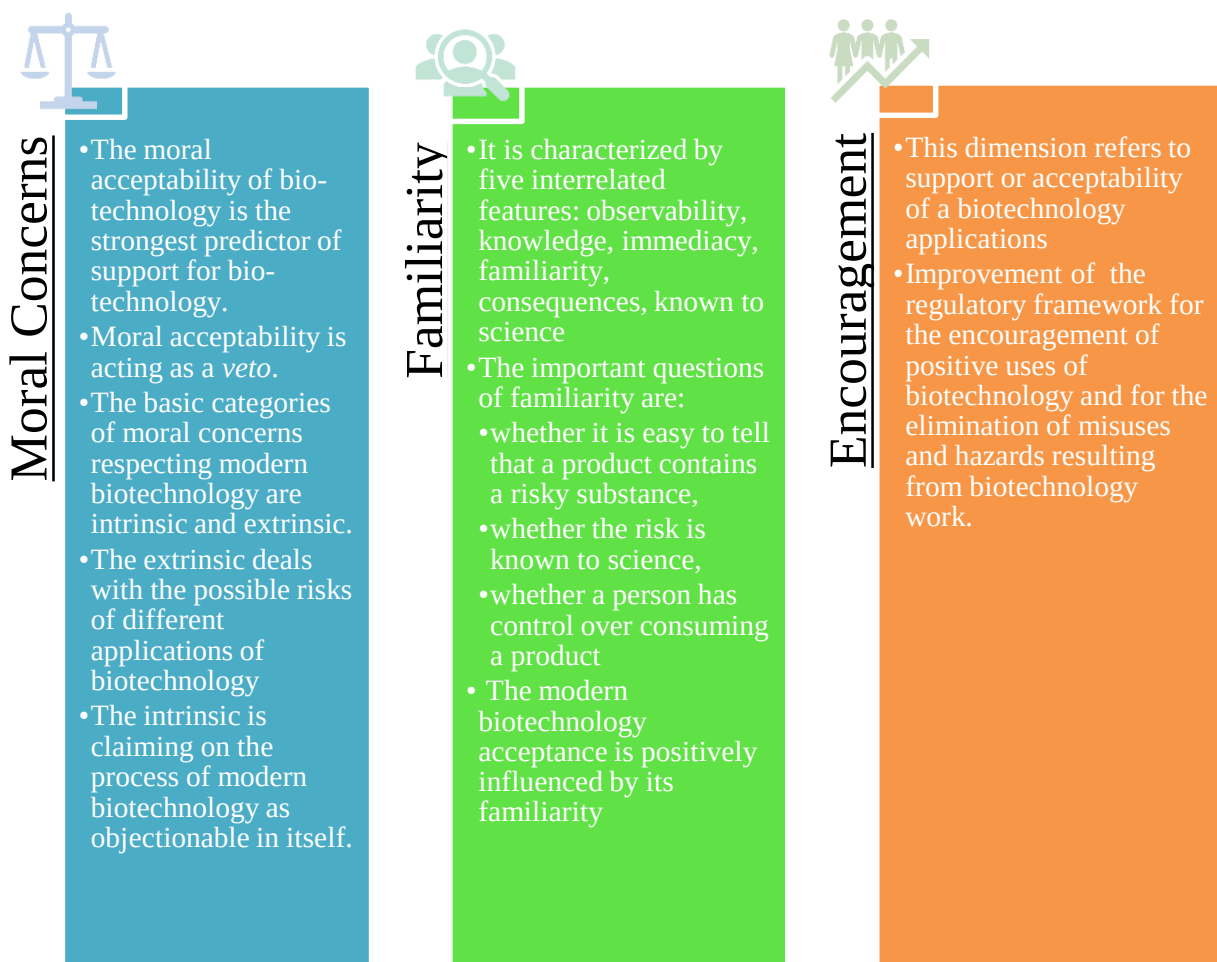


Figure 3. Social attitude indicators of modern biotechnology

Modern biotechnology is regarded as a complex emerging field that exhibits high science combined with limited knowledge of part of the society. It is related to attitudes towards the natural environment, technological progress, religious and moral beliefs, and several other sets. A matrix of variables including interest of the public domains, such as science and politics, optimism about technologies, social and cultural values, engagement with the issue of biotechnology and confidence in the industry, regulation, and other civil society groups, all these factors contribute to the public's representation of and opinions about biotechnologies (Fig. 4). Background characteristics such as gender, education, age, and religion have been found to affect people's attitudes as well.

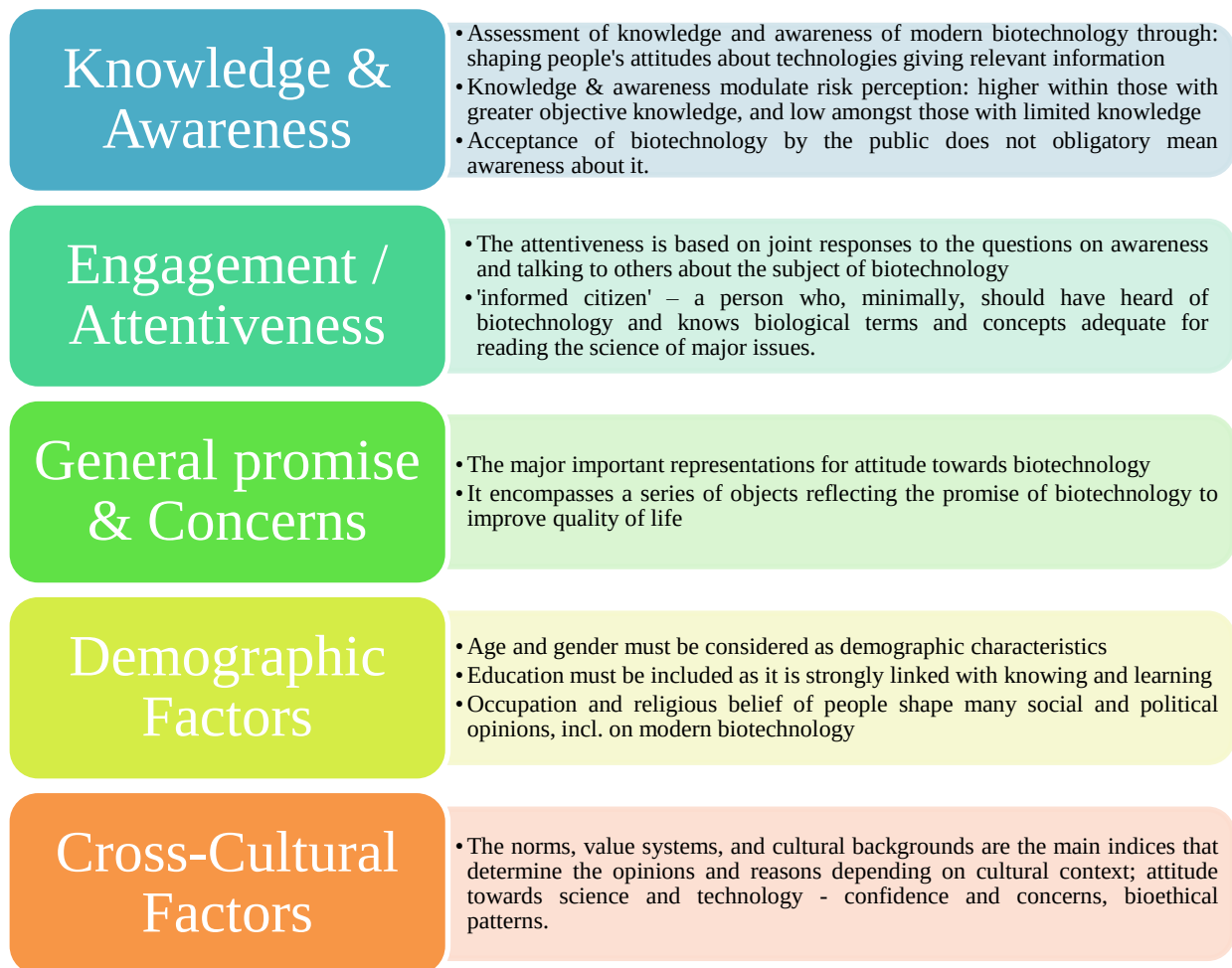


Figure 4. Factors of general attitude



2019-1-BG01-KA203-062371

The serious debate on the acceptance of genetic engineering is accompanied by many pieces of evidence that object to the focus on specific applications of the technology than genetic engineering *per se*. Studies are showing that consumer attitudes regarding gene transfers are influenced by the type of transfers. The range of acceptance encompasses the following: i) plant to plant gene transfers; ii) animal to animal transfer; iii) animal-plant or human-animal, the last being least acceptable.

A survey on perception on modern biotechnology applications, performed among people from different countries/continents indicated that medical applications (leading to the development of medicines and vaccines, and genetic testing) are very much acceptable in comparison to food or crop biotechnology applications. However, medical applications linked to xenotransplantation or animal cloning for milk production meet serious problems. The Asian people, for instance, are less troubled for medical genetically engineered products than for genetically modified food. The investigation concerning cloning of human cells and bioremediation among the Europeans is put at the intermediate level, less assumed than genetic testing but better confessed acceptable than genetically modified crop and food.

Another survey reported that genetic testing still scored the highest support within the European population, followed by human cell cloning, the production of GM enzymes for environmentally friendly soaps, and xenotransplantation. At the bottom of this range are placed GM crops and food. Also, higher support is given to technics for the manipulation of bacteria to clean oil spills followed by disease-resistant crops; afterward are applications for fat meat and better taste tomatoes and the less support receive technologies for enhancing milk production in cows.

The Malaysian consumers are more open for use of modern biotechnology applications, which do not involve inter-species gene transfers, like food production and GM crop that only involve the transfer of plant genes. For them, the gene modifications are more acceptable in the following order: i) transfer oil palm genes to reduce its fat (saturated) content, ii) transfer human genes into bacteria to produce insulin, iii) transfer bacterial genes into soybean to make it resistant to herbicides.

An individual's attitude towards a new technology depends on several related factors such as his/her perception of its risks and benefits, and socially communicated values and trusts in institutions representing these technologies. Concerning public perception of biotechnology, it is speculated that the attitude to genetic engineering is determined by the worth of potential benefits offered, knowledge on genetic engineering, and scientific world-view, from which the perceived risk (rational worries) and anxieties or fears (irrational worries) must be excluded. Additionally, various minor factors such as background factors must be added. It is considered that the main effect on acceptance seemingly is based on the knowledge level, awareness of benefits, confidence, and trust.

The studies of public attitude towards biotechnology indicate a lot of similarities with risk perception studies. Some authors have used the psychometric approach based on cognitive psychology and it is accepted as the most mature and dominant paradigm in risk perception studies. These psychometric methods are settled on the ground of the following views:



2019-1-BG01-KA203-062371

- To perform “risk” as a subjective idea, and not an objective one;
- To involve technical/physical and social/psychological outlooks in risk criteria;
- To take opinions of the “public” (i.e., laypeople, not experts) as the subject of interest.

Thus, the public acceptance of modern biotechnology is based on analyzing the cognitive structure of risk judgments, commonly using multivariate statistical procedures like factor analysis, multi-dimensional scaling, and multiple regression. This psychometric approach assumes that the public does not perceive technological risk due to a single dimension linked to expected injuries or fatalities close to a risk assessor's viewpoint but interprets risk as a multidimensional concept, concerned with broader qualitative dimensions.

The key varieties of a risk perception study are the perceived magnitude of risk or dread, risk acceptance, familiarity with the hazard, and at least - the factor benefit. The importance of another dimension, ‘interference with nature’ in risk perception studies on genetic engineering is also very important.

Biotechnology is at the confluence of science and ethics. The development of technology is linked with an ethical vision, which in turn is shaped by specific issues. Loads of biotechnology can be evaluated for their benefit to human society. But it has to be considered that biotechnology has also a dangerous part. It can provoke unanticipated consequences, which can cause harmful effects or dehumanize people. So, the ethical issues of proposed effects must be carefully investigated.

The ethical evaluation of new technologies, including biotechnology, requires a different approach to ethics. There is a necessity for changes because a new technology can have a more profound impact on the world due to i) restrictions to a rights-based ethics approach; ii) the importance and difficulty of predicting consequences; iii) opportunity now to manipulate humans themselves.

The ethical questions concerning biotechnology are very different. Due to the potential for deep change in the human future, such questions must be carefully considered. In the first place, it is necessary to articulate and predict the responsibilities towards nature and others, including future generations, and then to focus on rights and freedom. The real power and potential of biotechnology require strong caution to ensure ethical progress, as biotechnology is accepted as a significant force to improve the quality of people's lives in the 21st century.

Biotechnology is intrinsically linked to science and scientific knowledge. But sometimes there is doubt that biotechnology is closely tied to ethics. Lastly, different biotechnological trends are promoting a certain vision of life, some of which are good for life and deserved to be encouraged or pursued but others are bad and should be eliminated. This vision impacts people's choices and influences their sense of ethically appropriate biotechnology.

At times, the link between biotechnology and ethics is described as a conflicting point. Once, there is an impression that ethics is needed only in case one wants to tell others that they are doing wrong



things. To a certain degree, this is lucid, since dispute, debate, and argument are common parts of ethics consultation. But ethics is just as important in the case of consensus that the chosen pathway is good and right. Of course, there isn't any necessity for an ethical debate when the search is a cure for cancer. Thus, the solution to perform such a search is predicted by a common understanding that to cure cancer there is an ethical reason. The efforts, resources, and creativity devoted to working out better treatments are ethically eligible and the majority of these achievements due to biotechnology.

Benefits and Risks from Biotechnology Applications

The societal acceptability of benefits and risks imposed by modern biotech applications concern:

- Benefits vs. risks societal dispute
- Risk acceptance as a key element in risk perception
- Tools for measuring public attitudes towards risks and benefits: risk management

BENEFITS VS. RISKS SOCIETAL DISPUTE

To determine the acceptability of biotechnological applications it is very important to evaluate the perceived benefits and risks. The diagram on Fig. 5 summarizes them.



Figure 5. Perceived benefits and perceived risks

RISK ACCEPTANCE

Another key variable is risk acceptance which is important to measure risk perception. But it is rarely used in attitude towards biotechnology studies. Highlighted the harsh situation with modern technologies, they are always linked with some kind of risks that pose serious problems for societies. Policy-makers have considered cost-benefit analysis as the basis of decision-making methodology for societal risk acceptance. The core question to be solved in the risk-benefit analysis is: whether this

2019-1-BG01-KA203-062371

product (activity, technology) is eligible safe or how safe is safe enough? There are two main approaches for performance of risk-benefit analysis (Fig. 6).

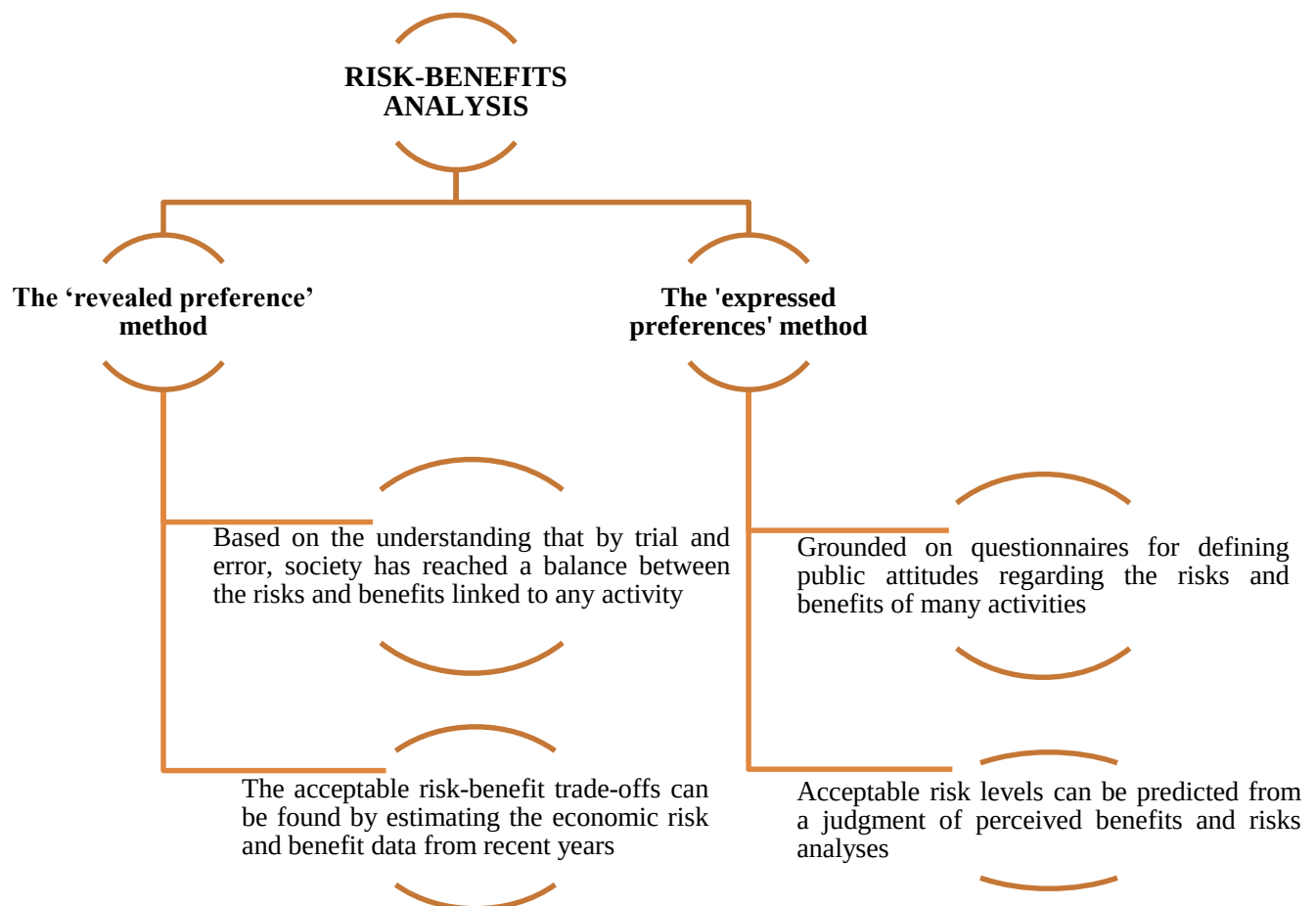


Figure 6. Approaches for risk-benefit analysis



RISK MANAGEMENT

The management of risk is the process that aims to diminish the effect and to balance the way people respond to it. If one wants to push away all risks, this is not a viable goal. People readily take risks performing a wide range of activities. Thus, the question about the acceptable risk is not entirely technical; its answer depends on values. In a simplified way, it is possible to set a simple threshold for acceptable risk, for instance - one in a million-lifetime risk. All the activities riskier than this threshold are “unacceptable” as well as all those less risky are “acceptable”.

However, there are two reasons why this approach is not working. The first is linked to the above-mentioned risk perception. The people are not reacting to risk simply; the acceptability of risk in the public's mind is not simply a matter of the presumable fatal things but involves other elements of the activity. For instance, if the activities are taken voluntarily if the process leading to the risk is new/old or familiar, if the process provokes an intuitive, emotional dread reaction or not. That is why, to define an acceptable risk, the processes of risk communication and engagement described above must be included. In the second issue, some activities are essential and must be encountered even if they cause relatively high risks.

Besides, there could be other cases where inexpensive alternatives exist to perform less risky or modified activities. In these cases, the risk might be well predicted as unacceptable, due to ready alternatives. This is an economic issue that deals with several resources dedicated to risk reduction. Besides the economic problems, it is possible to consider risks, not in terms of absolute values but to their internal differences between alternatives.

However, very often it is not possible fully to escape risk; rather one must choose among different risks. Therefore, no definite acceptable threshold of allowable risk exists, as well as there is no single convenient monetized value for risk diminution. Societal desire to invest in risk reduction depends on the wide range of factors that manage risk perception, like the fear associated with the risk, its catastrophic potential, etc.

SOCIAL BENEFITS OF MODERN BIOTECHNOLOGY

Biotechnologies are being used widely in such fields as medicine and hygiene, agriculture, forestry, breeding, fisheries, energy and chemical industries, metallurgical and mining industries, food, and light industries, environmental protection. Modern biotechnology has gradually demonstrated its huge potential contribution to productivity, though it has a short history of only tens of years so far. Modern biotechnology will become powerful means to reduce the world's constraints in the fields of food, health, energy, resources, and the environment. The development and application of modern biotechnology will, over a long time, affect human beings and society deeply.



Biotechnological industries are likely to become the leading industries in the 21st century. For example, crop varieties, produced by biotechnology, with novel traits of high productivity and stress resistance (aridity, coldness, high salt, etc.) are expected not only to increase grain production but also decrease the use of agricultural chemicals (such as pesticides, herbicides, and fertilizers), which will be beneficial to the environment. In the respect of medicine and hygiene, many drugs such as DNA vaccines, protein engineering medicine, monoclonal antibodies, anti-sense RNA drugs have been developed by biotechnology. Transgenic animals and plants producing some kinds of medicines (such as vaccines and hormones) may help to reduce illness in the process of everyday diet. Human Genome Project (HGP) and the research into some disease genes will discover the genetic reasons for some illnesses. Environmental pollution is getting worse at present, but biological transformation reactors created by biotechnology promise to absorb pollutants or wastes and decompose them into materials of low or no toxicity.

Therefore, in the present society characterized by knowledge-intensive industries, countries all over the world, especially the industrialized ones, take biotechnology seriously. Large quantities of manpower and financial resources are allocated to R & D in biotechnology. Relevant development strategies and policies are drawn up and support mechanisms created to stimulate the development of biotechnology.

Acceptance and diffusion of modern biotechnology

Although there have been many studies on public attitude or perception towards biotechnology and some researchers tried to identify factors predicting attitude using either regression or correlation, there are limited studies that try to construct a structural model predicting attitude towards biotechnology.

The first documented model was developed by Kelley (1995) who proposed a structured way for the approval of genetic engineering by the Australians (Fig. 7). Later, Pardo et al. (2002) proposed another model for explanation of European attitude towards biotechnology applications. Kelley found out that approval of genetic engineering was mainly predicted by agricultural and health goals (beneficial aspects). On the other hand, scientific and genetic engineering knowledge did not predict approval. Demographic variables also did not directly affect the approval of genetic engineering (their effects were not statistically significant). Gender was found to affect knowledge and scientific world-view while age was found to affect knowledge, goal, and scientific world-view, thereby exerting a very weak (indirect) effect on approval. Education had a large effect on knowledge but not concerning attitudes. Occupation and religion did not have any effect on any of the intermediate variables. Green supporters were less

2019-1-BG01-KA203-062371

favorable to genetic engineering (negative direct effect on the approval of genetic engineering) as compared to Labor Coalition supporters (politic).

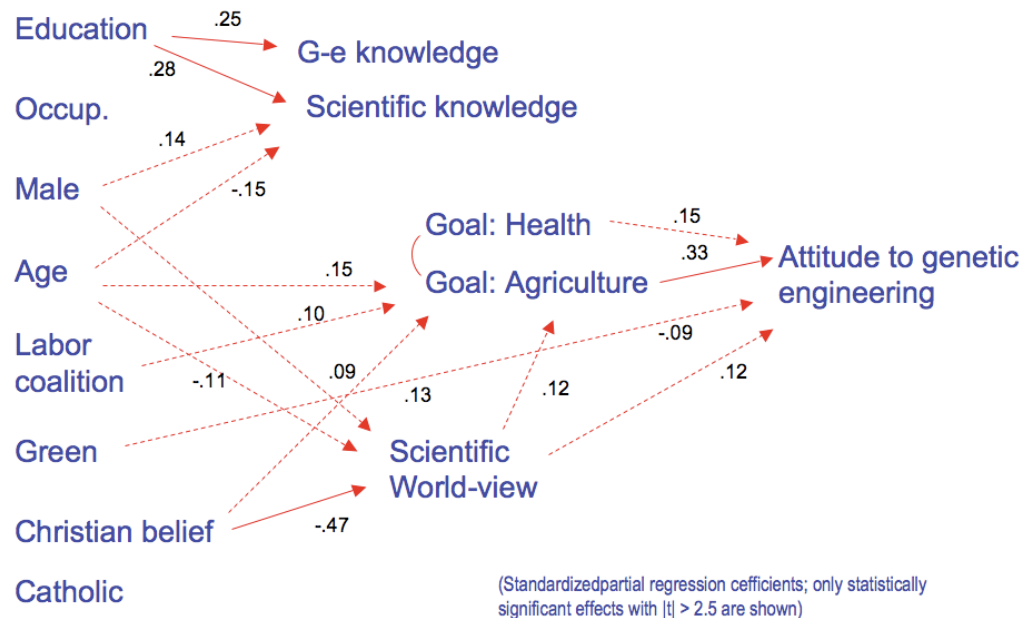


Figure 7. An ‘Attitude towards the genetic engineering’ model of Kelly (1995)

It is often speculated that biotechnology is the technology to deeply transform the economy and society. Following such opinions, the development and application of biotechnology have considerable potential for far-reaching economic, social, and environmental impacts. That is why biotechnology is of strategic importance to knowledge-based economies and their governments. Nations that failed to develop biotechnology capabilities fail in realizing economic impacts.

Biotechnology-related applications are diffused over several very different industrial sectors, e.g., food production, textile finishing, pulp and paper, agriculture, power generation, chemicals and petrochemicals, and pharmaceuticals. The products manufactured by these sectors do not usually distinguish between alternative production processes. Thus, for many products, a distinction between biotech and conventional (e.g., chemical) production processes and also the introduction of some new ones would be necessary. Moreover, biotechnology is a dynamic field and certain biotechnological production methods will probably soon be re-developed, and completely new, products will appear. Current official statistics and existing companies’ surveys can provide specialized information about the diffusion of biotechnological applications, and their economic impact. The data have been constantly renovated and different techno-economic studies are performed to provide future production foresight



2019-1-BG01-KA203-062371

and scenarios. Applying a blend of methods and sources, they frequently focus on expert judgment, case-studies, and various statistics.

An alternative to measuring the diffusion of biotechnology in use for major manufacturing sectors is the biotechnology-related-sales (BRS) concept. Existing studies have used various approaches for measurement (definition of the biotechnology aggregation level, examined sectors, scenario assumptions, etc.). Moreover, to assess the impact of biotechnology diffusion, an input-output model and relevant indicators have been introduced to study its effect on employment. The model studies in a comparative manner how biotechnology is affecting major application sectors. It compares the market penetration in 2004 and 2020 through the diffusion rates, measured with upper and lower limits. This approach helps not to overestimate the impact of biotechnology while capturing certain innovative biotechnological applications. The data of this survey indicate a nearly similar diffusion pattern in pharmaceuticals, food-processing, and agriculture. Diffusion is expected to increase in all sectors greatly by 2025, as indicated by the penetration of certain product groups. However, it was shown that there is a serious level of suspicion about the diffusion of biotechnology; the expectations for 2025 are highly uncertain. The analysis of this tendency, indicate that there are diffusion barriers specified in Fig. 8.

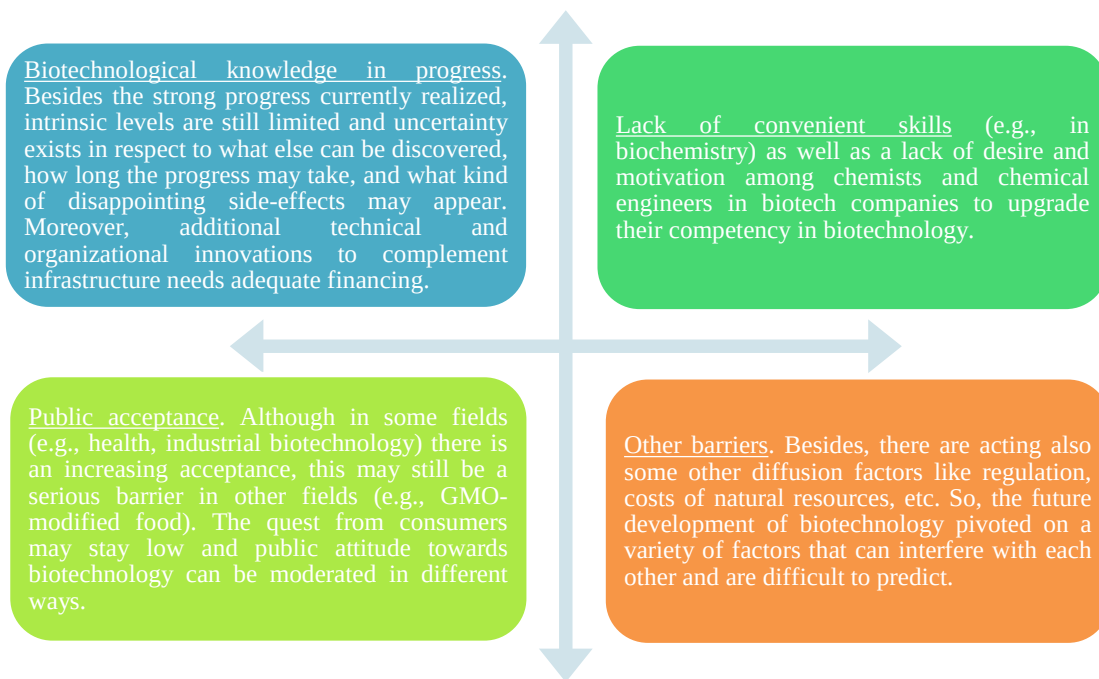


Figure 8. Main barriers interfering Biotechnology diffusion



Possible negative impacts of modern biotechnology

SOCIOECONOMICAL IMPACTS

Motivated to perceive great economic profit, a lot of famous international business corporations made considerable investments of capital in the R & D of biotechnology, and in this way are changing the previous world's social and economic trends. Such an example is the company Monsanto (formerly US-based, and now a part of the global consortium BASF). Traditionally working in the chemical industry, after 1985 the company made serious business rearrangements, invested in 3 biotechnological sectors, and became a major consortium dealing with R & D in bio-science. After 1998, about 20 million hectares of fields all over the world were planted with Monsanto's GM seeds. Soon after that, Monsanto took an advance on the crop seeds international market through a completely integrated system for trading GM seeds and agricultural chemicals. In this way, Monsanto received control on a big part of the human population food chain gaining tremendous profit feeding them with GM crop. This is an example of how in the current globalized world, a possibility for a few multinational corporations to decide on human food consumption emerges.

In the case of use in dairy, bovine growth hormone (BGH) produced by genetic engineering, the milk yield of a dairy cow can increase by 30 percent., which means 10 percent less fodder usage. So, the result is very impressive for the dairy industry.

A lot of medicines, like vaccines for treating malaria and cholera are provided to patients via new antigens synthesized through modern biotech methods. They make the rates of illness and death in many countries (especially in the developing ones) falling rapidly, concurring the population growth rates. In this context, a question arises about the impacts on the already insufficient economics and infrastructures of the said countries.

IMPACTS ON BIODIVERSITY AND SUSTAINABLE AGRICULTURE

Many GM crops contain foreign genes from other plants, animals, and microbes. All these foreign genes can be transferred to other plants in nature through the pollens of genetically transformed plants and will cause pollution to the natural pool of genetic resources. Hence, as the majority of transgenic plants possessed economically useful features, like high productivity, disease-pest resistance, stress resistance, and farmers are attracted to economic interests to grow these plants rather than traditional crop varieties, the predicted danger for the natural resources is serious. Also, the technologies for breeding of a monoculture of crop varieties in some fields may add to the loss of biodiversity. So, in a long term, this degradation of the genetic base may reduce the natural resistance of plants to diseases and insect pests and this may cause diminish in the crop yields. In this way, the global sustainable development could be compromised.



2019-1-BG01-KA203-062371

Bacillus thuringiensis (Bt) is a soil bacterium, which is used to synthesize a toxin against insects within plants. It has been used as a microbial pesticide for many years and the genes coding these insecticidal toxins could be introduced in the cell of many crops such as cotton, soybean, and rape, to create pest-resistance crops. In this case, the experts in environmental protection and the Green Peace Corporation are troubled by the fact that the toxin production is realized in the late life cycle of the transgenic plants and there is an opportunity to increase commonly the resistance of pests to the Bt toxins. So, this situation would lead to bigger usage of Bt-based pesticides, due to loss of their efficacy and causing heavy economic losses.

IMPACTS ON PUBLIC WELFARE

The exploitation of genetic engineering technology to humans could facilitate the diagnose and cure of genetically related diseases such as some cancers, hemophilia, etc. But such gene diagnosis may also have an opposing effect on the employment and marriage prospects of human beings. Through genetic engineering animals, plants and microbes can be transformed into bio-factories for the production of different kinds of chemicals. Thus, they can be treated as constant biological producers and at the same time, the rising potential for products outcome can cause socioeconomic problems. Assume that specific microbial strains are engineered to produce a valuable substance on large scale. In such a case, it is important to forecast the impact of this production on the world's major producers of the same substance. If this substance counts for a significant percentage of the income of a given country, what would be the influence and consequences of this biotech production on the country's agriculture and financial infrastructure.

Biosafety issues of modern biotechnology

Inconvenient application of biotechnology can cause many risks, like other advanced technologies. For instance, the newly introduced genes in crops made by genetic engineering can lead to allergic reactions. The economically important features of pest-resistance, herbicide-resistance, or stress-resistance, introduced in plants may get these transgenic hosts away from agricultural cultivation systems. The genes for resistance in transgenic plants can be transferred to their wild weedy relatives, and the latter can be transformed into "super" weeds, which control will be very difficult. Thus, the large-scale releases of transgenic pest-resistance crops into the environment will cause severe selection pressures to enforce the resistance of target pests.

The transgenic virus-resistance plants that are already transformed with foreign virus genes possess virus-coded proteins that may recombine with the genetic material of other viruses to yield new types of viruses with higher toxicity. If the selective killing of the target pests and pathogens is not



2019-1-BG01-KA203-062371

possible, the transgenic pest-and-disease resistance crops can at once poison other organisms and also may cause harmful action via food chains on beneficial microorganisms, insects, birds, and mammals.

Moreover, many biotechnological products are alive organisms and can move and reproduce by themselves. Once they are released into the environment and are proven to be toxic, then it is nearly impossible to discard them, and the harm they create can rise with time becoming more and more serious. In this way, environmental releases of transgenic plants on a large-scale may damage the natural ecological balance in a longtime perspective.

Biosafety concerns have hindered the R & D applications of biotechnology. The public in lots of countries has signified its negative opinion, repulsion, indeed dread regarding biotechnological products. This is voiced even through demonstrations, crushing fields crops, banding the import, or buying the biotechnological products. This makes the biosafety issues at the very top of the agenda of many countries and international societies. Laws and regulations on biosafety are formulated in many countries, and biosafety issues are regarded in many international documents and treaties, like Agenda 21 and The Convention on Biological Diversity.

Conclusions

Biotechnology is on the way to become seriously relevant to the economy and can gain considerable importance in several application sectors in the next 10 to 15 years. However, the overall effects of modern biotechnology on the economy are quite diverse. For instance, employment estimations show that increasing the part of biotechnology in application sectors and its link to suppliers indicate the trend for reduction of employment due to productivity effects. The economic importance of biotechnology rises because of its diffusion in mature manufacturing sectors. However, diffusion in these main application sectors can influence only limited parts of overall employment. Only if biotechnology increases in other sectors might a major economy-wide impact emerge.

The deepness of biotechnological transformation of the economy depends strongly on its potential to propose the creation of new products and to improve the efficiency in the production of existing ones. Hence, it is expected that biotechnology would enable the creation of many products but the potential to raise labor productivity is estimated to be restricted. One important view is that biotechnology may transform the economy and society differently, cooperating with other technologies like ICT. Major boosts for such economic growth may emerge in combination with other technologies or efficiency gains, which raise productivity and purchasing power to realize the fruits of biotechnology.



References

Agreements on Trade-Related Aspects of Intellectual Property Rights (TRIPS) available from https://www.wto.org/english/docs_e/legal_e/27-trips.pdf

Amin L, Hashim H, Sidik NM, Zainol ZA, Anua N. 2011. Public attitude towards modern biotechnology. *African Journal of Biotechnology* Vol. 10(58), pp. 12409-12417, Available online at <http://www.academicjournals.org/AJB> DOI: 10.5897/AJB11.1061

Amin L. 2013. *Modern Biotechnology in Malaysia* (UM Press)

Biology Online. Blue Biotechnology. http://www.biologyonline.org/kb/biology_articles/biotechnology/blue_biotechnology.html

Gaskell G, Alum N, Baouer M, Durant J, Allansdottir A, Bonfadelli H, Boy D, Cheveigne DS, Fjaestad B, Gutteling JM, Hampel J, Jelsoe E, Jesuino JG, Kohring M, Kronberger N, Midden Holden C. 1990. Activists urge ban on herbicide R & D. *Science*, 247:1539.

Horst W, Doelle J, Rokem S, Berovic M. 2019. *Biotechnology - Volume XIII: Fundamentals in Biotechnology*, EOLSS Publications, p 364.

ISAAA & The University of Illinois at Urbana Champaign (UIUC). 2003. *The social and cultural dimensions of Agricultural Biotechnology in Southeast Asia Report*.

Kelley J. 1995. *Public perceptions of genetic engineering: Australia, Final report to the Department of Industry, Science and Technology, May 1995*.

Lorenz P, Zinke H. 2005. White Biotechnology: Differences in US and EU approaches? *Trends. Biotechnol.* 23(12): 570-574.

Luscombe N, Greenbaum D, Gerstein M. 2001. What is Bioinformatics? A Proposed Definition and Overview of the Field. *Method. Inform. Med.* 4: 346-358.

Miller H, Huttner S. 1998. A baroque solution to nonproblem. *Nature biotechnology*, 16: 698-699.

Nielsen TH, Przestalski A, Rusanen T, Sakellaris G, Torgersen H, Twardowski T, Wagner W. 2000. Biotechnology and the European public. *Nat. Biotechnol.* (18): 935-938.

O'Mathúna DP. 2007. Bioethics and biotechnology. *Cytotechnology*. 53(1-3): 113–119. doi: 10.1007/s10616-007-9053-8



2019-1-BG01-KA203-062371

Pardo R, Midden C, Miller JD. 2002. Attitudes toward biotechnology in the European Union. *Journal of Biotechnology* 98, 9-24.

Wydra S, Nusser M. 2011. Diffusion and economic impacts of biotechnology – a case study for Germany. *Int. J. Biotechnology*, Vol. 12, Nos. 1/2, 2011 87 DOI: 10.1504/IJBT.2011.042683.



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**

Basic level

AUTHOR:
DIANA DI GIOIA



Contents

Renewable energy: biotechnology for biogas and bioethanol production	3
Biogas.....	3
Bioethanol	8
Biotechnology for the remediation of contaminated sites.....	14
Wastewater.....	14
Contaminated soils.....	21
STRENGTHS and WEAKNESSES of bioremediation technologies	27
References	28



Renewable energy: biotechnology for biogas and bioethanol production

Recent updates on the legislation on renewable energy → Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources. The 2030 climate and energy framework includes EU-wide targets and policy objectives for the period from 2021 to 2030. The key targets are:

- **Reduce CO₂ emissions (from 1990 levels) by 40%**
- **Increase renewable energy sources by 32%**
- **Improve energy efficiency by 32,5%**

BIOGAS

Biogas is the mixture of gases produced by the breakdown of organic matter in the absence of oxygen (anaerobically), primarily consisting of methane and carbon dioxide. Biogas can be produced from raw materials such as agricultural waste, manure, municipal waste, plant material, sewage, green waste or food waste. Biogas is a renewable energy source (Fig. 1). Biogas is considered to be a renewable resource because its production-and-use cycle is continuous, and it generates no net carbon dioxide. As the organic material grows, it is converted and used. It then regrows in a continually repeating cycle. From a carbon perspective, as much carbon dioxide is absorbed from the atmosphere in the growth of the primary bio-resource as is released, when the material is ultimately converted to energy.

Biogas is produced by anaerobic digestion with methanogenic or anaerobic organisms, which digest organic materials inside a closed system, or fermentation of biodegradable materials. This closed system is called an anaerobic digester or a biodigester.

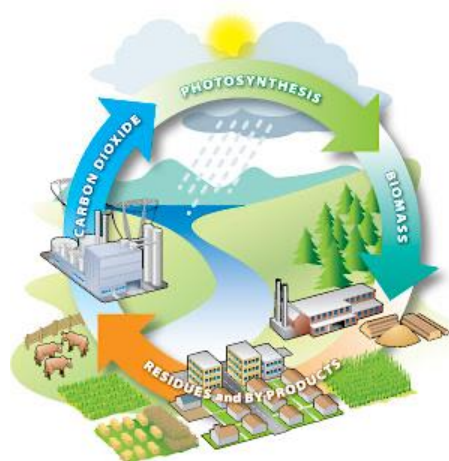


Fig. 1. Renewable energy cycle

Components	% (dry gas)
Methane (CH ₄)	50-80
Carbondioxide (CO ₂)	35-45
Sulphide (H ₂ S)	0.02-0.2
Water Vapour	saturation
Hydrogen, Ammonia	traces
Oxygen, Nitrogen	traces

Fig. 2. Main biogas components

Biogas is primarily methane (CH₄) and carbon dioxide (CO₂) and may have small amounts of hydrogen sulfide (H₂S), and moisture (Fig. 2). After purification from H₂S and moisture, the biogas can be used to produce electricity and thermal energy. In this way, biogas can be used in a co-generation system (a kind of gas engine) to convert the energy in the gas into electricity and heat.

Methane from biogas (after elimination of CO₂) can be used as a fuel and for any heating purpose. Biogas can therefore be cleaned and upgraded to natural gas standards, when it becomes bio-methane.

Biogas can be produced from a vast variety of raw materials (feedstocks). The biggest role in the biogas production process is played by microbes feeding on the biomass (details in the related power point). Materials suitable for biogas production include:

- biodegradable waste from enterprises and industrial facilities, such as surplus lactose from the production of lactose-free dairy products
- spoiled food from shops
- biowaste generated by consumers
- sludge from wastewater treatment plants
- manure and field biomass from agriculture

2019-1-BG01-KA203-062371

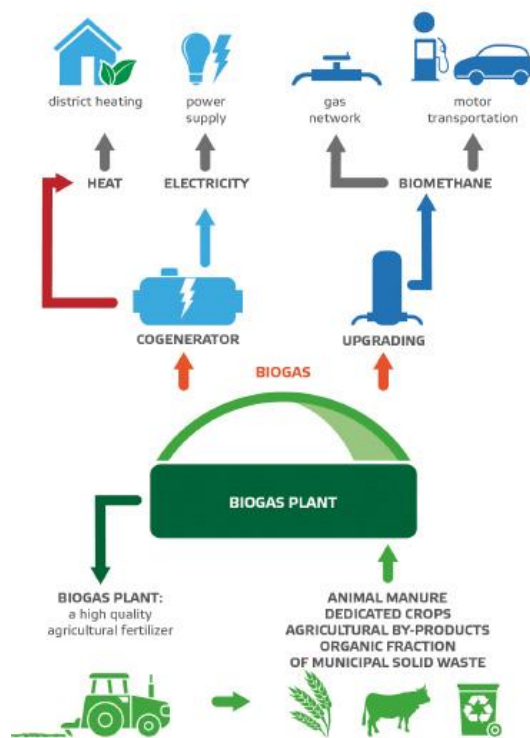


Fig. 3. Scheme of a biogas plant

Biomethane potential of different by-products

By-products	Biomethane potential (55% biogas) MNm ³ /year
Cattle waste	809,6
Pig waste	190
Slaughter waste	65,7
Crop residues and waste	1541
Agro-food industry by-products	598
FORSU (organic fraction of municipal solid waste)	561,7
Total	3766

Fig. 4. Biomethane potential of different by-products

The material is typically delivered to the biogas plant's reception pit by lorry or waste management vehicle. A delivery of solid matter such as biowaste will next undergo crushing to make its consistency as even as possible. At this point, water containing nutrients obtained from a further stage in the production process is also mixed with the feedstock to take the rate of solid matter down to only around one-tenth of the total volume.

This is also when any unwanted non-biodegradable waste, such as packaging plastic or out-of-date food waste from shops, is separated from the mixture. This waste is taken to a waste treatment facility where it is used to generate heat and electricity. Biomass that has passed through slurrification is combined with biomass delivered in the form of slurry to the biogas plant and pumped into the pre-digester tank where enzymes secreted by bacteria break down the biomass into an even finer consistency. A scheme of a biogas plant is presented in Fig. 3.

The residual solids and liquids created in biogas production are referred to as digestate. This digestate goes into a post-digester reactor and from there further into storage tanks. Digestates are well suited for uses such as fertilization of fields. Digestates can also be centrifuged to separate the solid and liquid parts.

Solid digestates is used as fertilizers in agriculture or in landscaping and can also be turned into gardening soil through a process of maturation involving composting.

Digestates are centrifuged to obtain process water for the slurrification of biowaste at the beginning of the process. This helps reduce the use of clean water. The centrifuged liquid is rich in nutrients, particularly nitrogen, that can be separated further using methods such as stripping technology and used as fertilizers or nutrient sources in industrial processes. Methane that can possibly be obtained from different by-products/waste is shown in Fig. 4.

2019-1-BG01-KA203-062371

Stages in biogas production. Anaerobic digestion is a multiple-stage process in which hydrolysis is one of the main steps. During hydrolysis the complex insoluble substrate macromolecules are hydrolysed into simpler and more soluble intermediates by bacteria. A large number of microbial species, acting in concert, are capable of utilising organic substrates such as carbohydrates, proteins and lipids to produce volatile fatty acids (VFAs), which can be converted to methane and carbon dioxide by methanogenic microorganisms. Bacteria excrete enzymes that hydrolyse the particulate substrate to small transportable molecules, which can pass through the cell membrane. Once inside the cell, these simple molecules are used to provide energy and to synthesize cellular components. Polysaccharides are converted to simple sugars; hydrolysis of cellulose by cellulase enzymes yields glucose; hemicellulose degradation results in monosaccharides such as xylose, glucose, galactose, pentoses, arabinose and mannose, while starch is converted to glucose by amylase enzymes. A scheme of the process and examples of microorganisms involved is depicted below (Fig. 5).

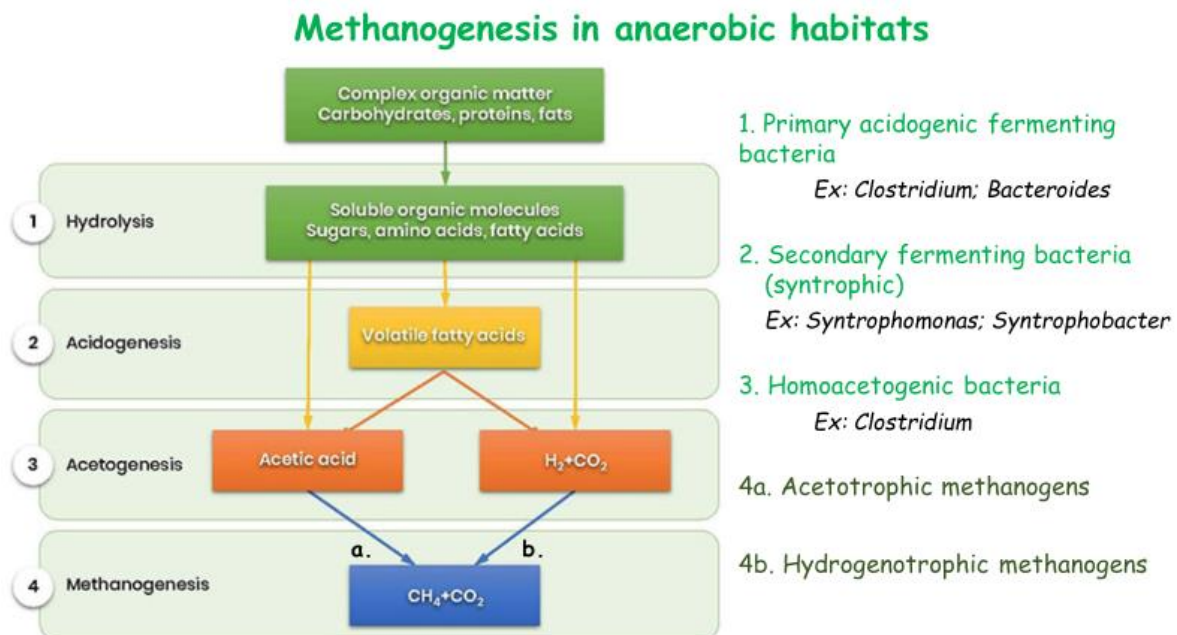


Fig. 5. Methanogenesis in anaerobic habitats

Methanogens are Archaea. They can live in different habitats and are a heterogeneous group of microorganisms. They lack cell nuclei and are therefore prokaryotes. Archaea were initially classified as bacteria, receiving the name archaeobacteria (in the Archaeobacteria kingdom), but this term has fallen out of use. Archaeal cells have unique properties separating them from the other two domains, Bacteria and Eukaryota. Archaea are further divided into multiple recognized phyla. Classification is difficult because most have not been isolated in a laboratory and have been detected only by their gene sequences

in environmental samples. Archaea and bacteria are generally similar in size and shape, although a few archaea have very different shapes. Despite this morphological similarity to bacteria, archaea possess genes and several metabolic pathways that are more closely related to those of eukaryotes, notably for the enzymes involved in transcription and translation. Other aspects of archaeal biochemistry are unique, such as their reliance on ether lipids in their cell membranes. Archaea use more energy sources than eukaryotes: these range from organic compounds, such as sugars, to ammonia, metal ions or even hydrogen gas. Salt-tolerant archaea (the Haloarchaea) use sunlight as an energy source, and other species of archaea fix carbon, but unlike plants and cyanobacteria, no known species of archaea does both. Archaea reproduce asexually by binary fission, fragmentation, or budding, unlike bacteria, no known species of Archaea form endospores. The first observed archaea were extremophiles, living in extreme environments, such as hot springs and salt lakes with no other organisms. Improved molecular detection tools led to the discovery of archaea in almost every habitat, including soil, oceans, and marshlands. Archaea are particularly numerous in the oceans, and the archaea in plankton may be one of the most abundant groups of organisms on the planet. Archaea are a major part of Earth's life. They are part of the microbiota of all organisms. In the human microbiome, they are important in the gut, mouth, and on the skin. Their morphological, metabolic, and geographical diversity permits them to play multiple ecological roles: carbon fixation; nitrogen cycling; organic compound turnover; and maintaining microbial symbiotic and syntrophic communities, methane formation.

Substrates for methanogenesis are a large array and they are listed below.

Substrates transformed into methane	ΔG° (kJ mol/ CH_4)	Microorganisms	
$\text{CO}_2 + 4\text{H}_2$	- 135	The majority of methanogens	Hydrogenotrophs
Formate	- 281	The majority of methanogens that uses $\text{CO}_2 + 4\text{H}_2$	
CO	- 196	Methanobacterium and Methanosarcina	
Methanol	- 105	Methanosarcina and other methylotrophic archaea	Methylothrophs
Methanol + H_2	- 113	Methanosphaera stadtmaniae and cuniculi	
Trimethylamine/methylamine/dimethylamine	- 76	Methanosarcina and other methanogens	
Dimethylsulphide	- 49	Some methylotrophic methanogens	Acetothrophs
Acetate	- 31	Methanosarcina and Methanosaeta	

Fig. 6. Substrates for methanogenesis

BIOETHANOL

The use of bioethanol in Europe will be described and the main microorganisms used for bioethanol production (yeasts and bacteria) will be presented and their performances compared.

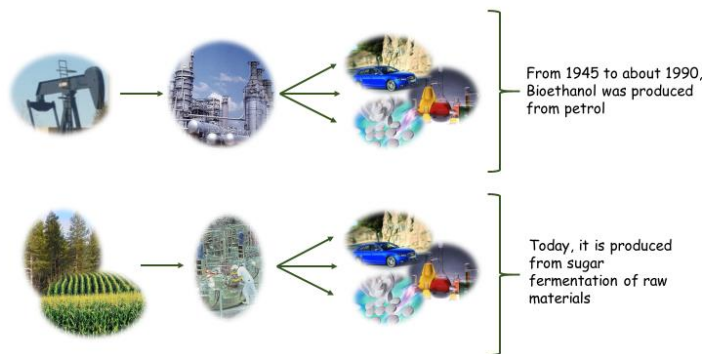


Fig. 7. Ethanol vs bioethanol production

Ethanol is widely used as a solvent, reagent, in the food industry and automotive fuel (Fig. 7 and 8). The production of ethanol by fermentation is based on the use of raw materials, microorganisms and technologies different from those used for the production of alcoholic beverages (wine and beer) in order to have the maximum yield in ethanol in the shortest time and with the lowest costs. 95% of the world actual ethanol production is BIOETHANOL (sugar fermentation process), whereas only 5% is manufactured by the chemical process of reacting ethylene with steam.

Renewable ethanol is manufactured in a biorefinery by fermenting sugars into alcohol. In the EU, these sugars typically come from a variety of agricultural sources such as wheat, corn, barley, rye, triticale, and sugar beet. The majority of renewable ethanol biorefineries are built to specifically process either grains or sugar beets. Currently, the most commonly used feedstocks in Europe are corn, sugar beet and wheat.

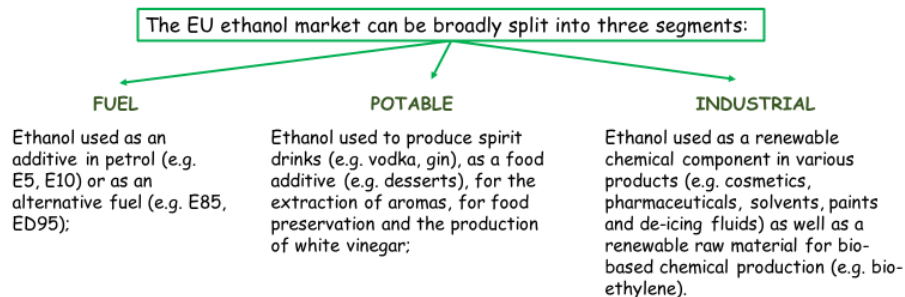


Fig. 8. Bioethanol use on the EU market

2019-1-BG01-KA203-062371

Bioethanol is the principle fuel used as a petrol substitute for road transport vehicles. Ethanol or ethyl alcohol (C_2H_5OH) is a clear colourless liquid, it is biodegradable, low in toxicity and causes little environmental pollution if spilt. Ethanol burns to produce carbon dioxide and water. Ethanol is a high octane fuel and has replaced lead as an octane enhancer in petrol. By blending ethanol with gasoline we can also oxygenate the fuel mixture so it burns more completely and reduces polluting emissions. Ethanol fuel blends are widely sold in the United States. The most common blend is 10% ethanol and 90% petrol (E10). Vehicle engines require no modifications to run on E10 and vehicle warranties are unaffected also. Only flexible fuel vehicles can run on up to 85% ethanol and 15% petrol blends (E85).

The main **sources of sugar** required to produce ethanol come from fuel or energy crops. These crops are grown specifically for energy use and include corn, maize and wheat crops (1st generation bioethanol), waste straw, willow and poplar trees, forest and agricultural waste, residues from pulp production, solid urban waste (2nd generation bioethanol), third-generation bioethanol has been derived from algal biomass including microalgae and macroalgae (Fig. 9).

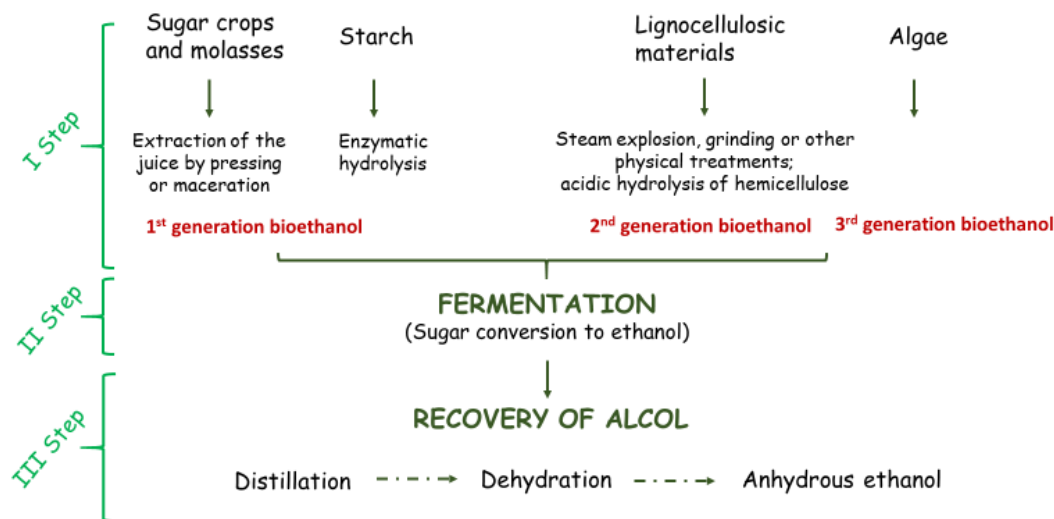


Fig. 9. First, second and third generation bioethanol

Benefits of Bioethanol. Bioethanol has a number of advantages over conventional fuels. It comes from a renewable resource i.e. crops and not from a finite resource and the crops it derives from can grow well in Europe (like cereals, sugar beet and maize). Another benefit over fossil fuels is the greenhouse gas emissions. The road transport network accounts for 22% of all greenhouse gas emissions and through the use of bioethanol, some of these emissions will be reduced as the fuel crops absorb the



2019-1-BG01-KA203-062371

CO₂ they emit through growing. Also, blending bioethanol with petrol will help extend the life of the diminishing oil supplies and ensure greater fuel security, avoiding heavy reliance on oil producing nations. By encouraging bioethanol's use, the rural economy would also receive a boost from growing the necessary crops. Bioethanol is also biodegradable and far less toxic than fossil fuels. In addition, by using bioethanol in older engines can help reduce the amount of carbon monoxide produced by the vehicle thus improving air quality. Another advantage of bioethanol is the ease with which it can be easily integrated into the existing road transport fuel system. In quantities up to 5%, bioethanol can be blended with conventional fuel without the need of engine modifications.

Bioethanol can be produced from biomass by the **hydrolysis and sugar fermentation processes**. Biomass wastes contain a complex mixture of carbohydrate polymers from the plant cell walls as cellulose, hemicellulose and lignin. In order to produce sugars from the biomass, the biomass is pre-treated with acids or enzymes in order to reduce the size of the feedstock and to disrupt the plant structure. The cellulose and the hemicellulose portions are broken down (hydrolysed) by enzymes or dilute acids into sucrose sugar that is then fermented into ethanol. The lignin which is also present in the biomass is normally used as a fuel for the ethanol production plants boilers.

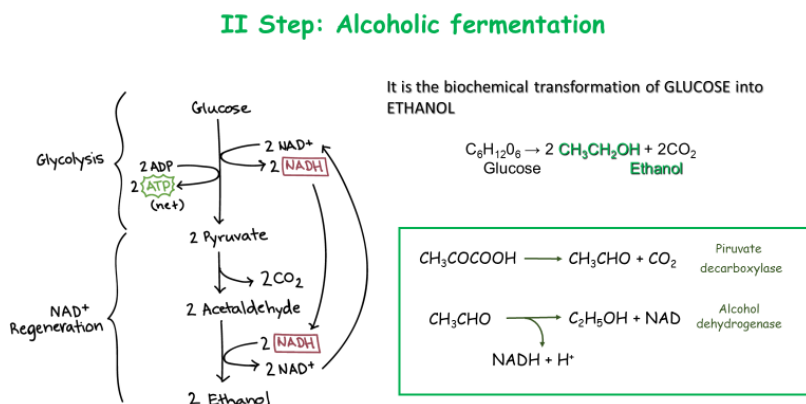
There are three principle methods of extracting sugars from biomass: concentrated acid hydrolysis, dilute acid hydrolysis and enzymatic hydrolysis. The first one works by adding 70-77% sulphuric acid to the biomass that has been dried to a 10% moisture content. The acid is added in the ratio of 1.25 acid to 1 biomass and the temperature is controlled to 50°C. Water is then added to dilute the acid to 20-30% and the mixture is again heated to 100°C for 1 hour. The gel produced from this mixture is then pressed to release an acid sugar mixture and a chromatographic column is used to separate the acid and sugar mixture. The dilute acid hydrolysis process is one of the oldest, simplest and most efficient methods of producing ethanol from biomass. Dilute acid is used to hydrolyse the biomass to sucrose. The first stage uses 0.7% sulphuric acid at 190°C to hydrolyse the hemicellulose present in the biomass. The second stage is optimised to yield the more resistant cellulose fraction. This is achieved by using 0.4% sulphuric acid at 215°C. The liquid hydrolytes are then neutralised and recovered from the process. Instead of using acid to hydrolyse the biomass into sucrose, we can use **enzymes** to break down the biomass in a similar way.

Corn, one of the most agricultural sources to obtain ethanol, can be processed into ethanol by either the dry milling or the wet milling process. In the wet milling process, the corn kernel is steeped in warm water, this helps to break down the proteins and release the starch present in the corn and helps to soften the kernel for the milling process. The corn is then milled to produce germ, fibre and starch products. The germ is extracted to produce corn oil and the starch fraction undergoes centrifugation and saccharification to produce gluten wet cake. The ethanol is then extracted by the distillation process. The wet milling process is normally used in factories producing several hundred million gallons of ethanol every year. The dry milling process involves cleaning and breaking down the corn kernel into fine particles using a hammer mill process. This creates a powder with a coarse flour type consistency. The powder contains the corn germ, starch and fibre. In order to produce a sugar solution the mixture is then

2019-1-BG01-KA203-062371

hydrolysed or broken down into sucrose sugars using enzymes or a dilute acid. The mixture is then cooled and yeast is added in order to ferment the mixture into ethanol.

Sugar Fermentation Process



The hydrolysis process breaks down the cellulosic part of the biomass or corn into sugar solutions that can then be fermented into ethanol. Yeast is added to the solution, which is then heated. The yeast contains an enzyme called invertase, which acts as a catalyst and helps to convert the sucrose sugars into glucose and fructose.

Fig. 10. Yeast alcoholic fermentation

The main features of microorganism to be used in the industrial production of ethanol are:

- **high yields of molar conversion of sugar** (moles of ethanol produced/moles of sugar consumed)
- **high production rate** (moles of ethanol produced/time × volume of culture) mol/Lh or g/Lh
- **high ethanol yields in weight** (grams of produced ethanol/volumes of culture, ≥ 120 g/L)
- **high tolerance to ethanol**
- as **low as possible production of side products** derived from side fermentations (e.g. glycerol)

Yeasts for bioethanol fermentation: Microorganisms such as yeasts play an essential role in bioethanol production by fermenting a wide range of sugars to ethanol. They are used in industrial plants due to valuable properties in ethanol yield (>90.0% theoretical yield), ethanol tolerance (>40.0 g/L), ethanol productivity (>1.0 g/L/h), growth in simple, inexpensive media and undiluted fermentation broth with resistance to inhibitors and retard contaminants from growth condition. As the main component in fermentation, yeasts affect the amount of ethanol yield. *Saccharomyces cerevisiae* is the most widely used yeast. Since thousands of years ago, *S. cerevisiae* have been used in alcohol production especially in the brewery and wine industries. It keeps the distillation cost low as it gives a high ethanol yield, a high productivity and can withstand high ethanol concentration. Nowadays, yeasts are used to generate



2019-1-BG01-KA203-062371

fuel ethanol from renewable energy sources. Some yeast strains belonging to the species *Pichia stipitis*, *S. cerevisiae* and *Kluyveromyces fragilis* were reported as good ethanol producers from different types of sugars. *S. cerevisiae* tolerates a wide range of pH thus making the process less susceptible to infection. Baker's yeast was traditionally used as a starter culture in ethanol production due to its low cost and easy availability. However, baker's yeast and other *S. cerevisiae* strains were unable to compete with wild-type yeast which caused contamination during industrial processes. Stressful conditions like an increase in ethanol concentration, temperature, osmotic stress and bacterial contamination are the reasons why the yeast cannot survive during the fermentation. Flocculent yeasts were also used during biological fermentation for ethanol production as it facilitates downstream processing, allows operation at high cell density and gives higher overall productivity. It reduces the cost of cells recovery as it separates easily from the fermentation medium without centrifugation. There are common challenges to yeasts during sugar fermentation which are rise in temperature (35–45°C) and ethanol concentration (over 20%). Yeasts growth rate and metabolism increase as the temperature increases until it reaches the optimum value. Increase in ethanol concentration during fermentation can cause inhibition to microorganism growth and viability. Inability of *S. cerevisiae* to grow in media containing high level of alcohol leads to the inhibition of ethanol production. The other problems in bioethanol fermentation by yeast are the ability to ferment pentose sugars. *S. cerevisiae* is the most commonly used in bioethanol production. However, it can only ferment hexoses but not pentoses. Only some yeasts from genera *Pichia*, *Candida*, *Schizosaccharomyces* and *Pachysolen* are capable of fermenting pentoses to ethanol. The efficiency of ethanol production on an industrial scale will be increased by using yeasts that are tolerant to inhibitors. The common challenges of yeasts can be overcome by using ethanol-tolerant and thermotolerant yeast. Ethanol-tolerant and thermotolerant strains which can resist stresses can be isolated from natural resources such as soil, water, plants and animals. Ethanol fermentation at high temperature is a beneficial process as it selects thermo-tolerant microorganisms and does not require cooling costs and cellulase. For example, *K. marxianus* is thermotolerant yeast which is capable of co-fermenting both hexose and pentose sugars and can survive the temperature of 42-45°C.

***Zymomonas mobilis* for bioethanol fermentation:** *Zymomonas mobilis* is a Gram negative, facultative anaerobic, non-sporulating, polarly-flagellated, rod-shaped bacterium. It is the only species of the genus *Zymomonas*. It has notable bioethanol-producing capabilities, which surpass yeast in some aspects. It was originally isolated from alcoholic beverages like the African palm wine, the Mexican pulque, and also as a contaminant of cider and beer (cider sickness and beer spoilage) in European countries. *Zymomonas mobilis* degrades sugars to pyruvate using the Entner-Doudoroff pathway. The pyruvate is then fermented to produce ethanol and carbon dioxide as the only products (analogous to yeast). The advantages of *Z. mobilis* over *S. cerevisiae* with respect to producing bioethanol are:

- ✓ higher sugar uptake and ethanol yield (up to 2.5 times higher),
- ✓ lower biomass production,
- ✓ higher ethanol tolerance up to 16% (v/v),
- ✓ does not require controlled addition of oxygen during the fermentation.

2019-1-BG01-KA203-062371

However, in spite of these attractive advantages, several factors prevent the commercial usage of *Z. mobilis* in cellulosic ethanol production. The foremost hurdle is that its substrate range is limited to glucose, fructose and sucrose. Wild-type *Z. mobilis* cannot ferment C5 sugars like xylose and arabinose which are important components of lignocellulosic hydrolysates. Unlike *E. coli* and yeast, *Z. mobilis* cannot tolerate toxic inhibitors present in lignocellulosic hydrolysates such as acetic acid and various phenolic compounds. Concentration of acetic acid in lignocellulosic hydrolysates can be as high as 1.5% (w/v), which is well above the tolerance threshold of *Z. mobilis*. Engineered *Z. mobilis* could overcome its inherent deficiencies and therefore expanded its substrate range to include C5 sugars like xylose and arabinose. Acetic acid resistant strains of *Z. mobilis* have been developed by rational metabolic engineering efforts, mutagenesis techniques or adaptive mutation. Moreover, an extensive adaptation process was used to improve xylose fermentation in *Z. mobilis*. By adapting a strain in a high concentration of xylose, significant alterations of metabolism occurred. One noticeable change was reduced levels of xylitol, a byproduct of xylose fermentation which can inhibit the strain's xylose metabolism. An interesting characteristic of *Z. mobilis* is that its plasma membrane contains hopanoids, pentacyclic compounds similar to Eukaryotic sterols. This allows it to have an extraordinary tolerance to ethanol in its environment, around 13%.

A comparison between production of bioethanol by *S. cerevisiae* and *Z. mobilis* is shown in Fig. 11.

Ethanol production: comparison between *S. cerevisiae* and *Z. mobilis*

<i>Saccharomyces cerevisiae</i>	<i>Zymomonas mobilis</i>
• metabolic versatility • produces EtOH with high yields and rate until EtOH reaches 11-12% (v/v)	• uses a limited number of substrates with respect to yeasts (glucose, fructose and sucrose) • glucose uptake and fermentation rate to EtOH 8 times higher than <i>Saccharomyces</i> • no oxygen request • grows at higher temperature than yeast (40 °C), more favorable for the recovery of EtOH
Ethanol productivity rate: 0.67 g of ethanol/g of glucose × h	Ethanol productivity rate: 5.67 g of ethanol/g of glucose × h

Fig. 11. Comparison between different microbial agents operating the bioethanol fermentation

Biotechnology for the remediation of contaminated sites

WASTEWATER

Water is a resource to be protected. Please find below the main location where freshwater is located on earth and some of the reasons why water is slowly running out. Almost 70 % of the water today is consumed for agriculture, about one-quarter for commercial uses, and roughly 10% is utilized for domestic purposes. Therefore, the main sector that uses water is agriculture/farming. Agricultural water is mainly used for irrigation as well as pesticide and

fertilizer applications and for animal farming. There are three sources for agriculture water: i) Groundwater from underground wells; ii) surface water that is derived from open canals, streams, irrigation ditches, and diverted from reservoirs; iii) rainwater which is usually collected in barrels, tubs, and large cisterns. **Water is often poisoned.** The main causes of poisoning are: domestic use of water (organic matter, surfactants...), agriculture and industrial use of water (fertilizers and pesticides, water deriving from industrial processes), atmosphere (contamination of rainwater by toxic substances present in the atmosphere, deriving from industries, airplanes, motor vehicle engines).

Directive 2000/60/EC of the European Parliament established a framework for Community actions in the field of water policy. This legislation establishes a framework for the protection of inland surface waters, coastal waters and groundwater. The objectives of the directive are:

- the safeguard against further deterioration;
- the improvement of the state of the ecosystems;
- the promotion of sustainable water use;
- the reduction of groundwater pollution;
- the reduction of discharges;
- mitigation of the effects of floods and droughts.



Fig. 12. Water location and use

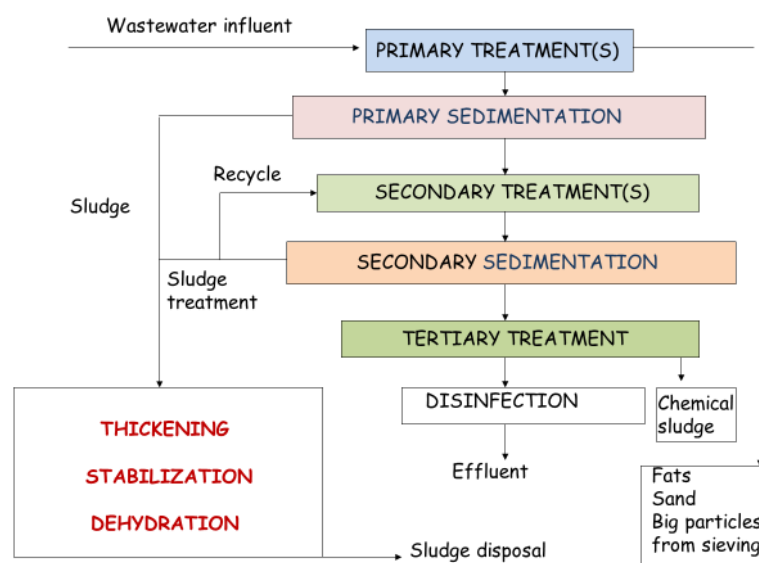
2019-1-BG01-KA203-062371

The depuration of wastewater is obtained through different treatments that are often applied in sequences: primary, secondary (biological), tertiary treatments as indicated in the picture) (Fig. 13 and 14).

Biological depuration of wastewater

Primary treatment (mechanical treatments)	Sieving Trituration Equalization of flow rates Omogenization of pollutants Desanding (sedimentation or centrifugation) Defatting Primary sedimentation
Secondary treatment (biological treatment)	Activated sludge Percolating filters Biocircles Anaerobic digestion Phytodepuration
Tertiary treatment (chemical treatment)	Chlorination Ozonization

Fig. 13 (above) and 14 (below). Primary, secondary and tertiary treatment for the biological depuration of wastewater



The **activated sludge process** is one of the most commonly used for **secondary wastewater treatment** of civil and industrial origin. It is a suspended-growth biological treatment process, using a dense microbial culture in suspension to biodegrade organic material under aerobic conditions and forming spontaneously a biological floc (referred to as activated sludge). Diffused or mechanical aeration maintains the aerobic environment in the reactor. Typical retention times are 5-14 hours in conventional units rising to 24-72 in low rate systems. The activated sludge process depends on aerobic biological action (Fig. 15). The microorganisms break down the complex organic substances into simple molecules including water and CO₂. This process results in the removal of soluble and suspended organic matter from wastewater. The growth of microorganisms in the presence of dissolved oxygen removes the majority of pollutant matter; in turn, protozoa grow and feed on these organisms. The resulting balance is of a living culture in suspended form in the activated sludge flocs. The main elements of the system include an aeration tank (secondary treatment) in which the wastewater is thoroughly mixed with continuously activated sludge and oxygen. From this part of the process, it passes into a clarifier tank (secondary sedimentation), where the settled sludge is removed from the purified water to be recycled by the return activated sludge pumps. For this system to work, two requirements must be met: the aeration device must be capable of both transferring oxygen from the atmosphere to the liquid, and distributing this oxygen throughout the wastewater to the suspended living microorganism. This type of



Fig. 15. Aerobic oxidation tank

system is suited for low-strength waste, typically on the order of 50-200 mg L⁻¹ BOD. Pre- or post-treatment of wastewater may also be applied. After the aeration basin, the mixture of microorganisms and wastewater (mixed liquor) flows into a settling basin or clarifier where the sludge is allowed to settle. Some of the sludge volume is continuously recirculated from the clarifier, as Returned Activated Sludge, back to the aeration basin to ensure adequate amounts of microorganisms are maintained in the aeration tank. The microorganisms are again mixed with incoming wastewater where they are reactivated to consume organic nutrients. Then the process starts again.

The activated sludge process, under proper conditions, is very efficient. It removes 85 to 95 percent of the solids and reduces the biochemical oxygen demand (BOD) about the same amount. The efficiency of this system depends on many factors, including wastewater climate and characteristics. Toxic wastes that enter the treatment system can disrupt the biological activity. Wastes heavy in soaps or detergents can cause excessive frothing and thereby create aesthetic or nuisance problems. In areas where industrial and sanitary wastes are combined, industrial wastewater must often be pretreated to remove the toxic chemical components before it is discharged into the activated sludge treatment process. Nevertheless, microbiological treatment of wastewater is by far the most natural and effective process for removing wastes from water.

Microbial populations in activated sludges: The microbial community participating to the biological depuration process forms floc agglomerates which are called **activated sludges**. The activated sludge of a of a secondary treatment plant is a microbial culture that develops around organic and inorganic particles and that metabolizes the organic matter present in wastewater. Activated sludge flocs tend to settle down in the “secondary” sedimentation phase because of gravity. Several groups of microorganisms are responsible for the depuration process:

- **Bacteria**
 - **Fungi**
 - **(Algae)**
 - **Protozoa**
 - **Rotifers and Nematodes**
- } *Decomposers: microorganisms involved in decomposition of organic matter (heterotrophic, saprotrophs = obtain nutrients directly from decomposition of dead organic matter)*
- } *Predators*

Bacteria are primarily responsible for removing organic nutrients from the wastewater. They also develop a sticky layer of slime around the cell wall that enables them to clump together to form bio-solids or sludge that is then separated from the liquid phase. The successful removal of wastes from the water depends on how efficiently the bacteria consume the organic material and on the ability of the bacteria to stick together, form floc, and settle out of the bulk fluid. The flocculation (clumping) characteristics of the microorganisms inactivated sludge enable them to amass to form solid masses large enough to settle to the bottom of the settling basin.

FUNGI are also heterotrophic organisms helping in the degradation of organic matter.

Protozoa play a critical role in the treatment process by removing and digesting free swimming dispersed bacteria and other suspended particles. This improves the clarity of the wastewater effluent. Like bacteria, some protozoa need oxygen, some require very little oxygen, and a few can survive without oxygen. The types of protozoa present are classified as follows:

Amoebae → Little effect on treatment & die off as amount of food decreases

Flagellate → Feed primarily on soluble organic nutrients

Ciliates → Clarify water by removing suspended bacteria.

Rotifers and Nematods are multi-cellular organisms which are larger than most protozoa and do not basically remove organic material from the wastewater. Although they can eat bacteria, they also feed on algae and protozoa. A dominance of metazoa is usually found in longer age systems; namely, lagoon treatment systems. Although their contribution in the activated sludge treatment system is small, their presence does indicate treatment system conditions.

In addition to activated sludge plants, there are other types of secondary wastewater treatment processes. Some of them can be **IMMOBILIZED CELLS PROCESSES**, as indicated in Fig. 16.

2019-1-BG01-KA203-062371

- ✓ Cells involved in the depuration process develop on a solid support rather than in the form of suspended flocs;
- ✓ a biological film is formed, some mm thick, through which soluble substrates and oxygen spread by diffusion;
- ✓ the biological depuration process takes place in the presence of three phases: solid, liquid and gas, consisting of the support, the wastewater and air;
- ✓ aerobic and anaerobic technologies have been developed.

Percolating filters



Biocircles

**Fig. 16. Immobilized cell technologies for the depuration of wastewater**

Percolating filters are a biological purification technology through microorganisms that develop, in an aerobic environment, on appropriate support materials, through which sewage percolates. The percolating filter tank is filled with inert, natural or synthetic materials (e.g. stones) through which wastewater is fed from above. The **DEPURATION PHASES** include: 1) Primary treatments to prevent obstruction of the bed; 2) Formation of the biofilm (3-4 mm thick); 3) Cell detachment from biofilm and setting-up of the biofilm again; 4) Final sedimentation. The main advantages of this technology are the low cost of set-up and maintenance and the fact that they can tolerate variation of the organic load of the influent. The main drawbacks include the large areas for the set-up and the problems of bad odors. Biocircles are a modified version of the percolating filters, in which the surfaces bearing the biofilm rotates around an axis, half immersed in the liquid to be treated; the rotation allows the oxygenation of the biomass adhered to the disc. Anaerobic systems work well when the entering flow rate is low and the organic load entering the system is sufficiently high. The required depuration efficiency is not high (anaerobic microorganisms are characterized by a lower growth rate and a slower metabolism with respect to aerobic ones, so the organic matter is not completely degraded). These conditions are **typical of some industrial wastewater**, anaerobic reactors do not work well for the treatment of large scale civil effluents.

Another important system for the depuration of wastewater is **phytodepuration** (Fig. 17), which is a: purification technology characterised by biological treatments, in which plants growing in water-saturated soil develop a key role helped by the direct action of the bacteria that colonize the root system

2019-1-BG01-KA203-062371

and rootstock. These treatments are seen both as alternatives as well as support to traditional systems based on biological processes and chemical and physical reactions. The term “Wetlands” indicates “Phytodepuration” systems of wastewaters designed to artificially create the same ecological conditions that are naturally established in watery areas. “Phytodepuration” systems engineered, designed and built to reproduce natural self-depurative processes in a controllable environment. In comparison to natural wetlands, phytodepuration systems allow for the choice of the site, the flexibility in the dimension, control of hydraulic flows and retention times. Phytodepurifying functions can be preferred and additionally exploited with opportune strategies, like the choice of plant species and substratum and control of the flow of water. With Phytodepuration systems, pollutants are removed through a combination of chemical, physical and biological processes. The most effective processes are sedimentation, precipitation, adsorption, assimilation from plants and microbial activity. Phytodepuration technology adds the medium’s adsorbing ability to the traditional biological oxidation depurative treatment (filtering action by plant roots that also provide a large surface suitable for developing microbial masses involved in the treatment) and removal of nutrients due to their growth. Different strategies are described in Fig. 18:

Phytodepuration

Phytodepuration is a natural wastewater treatment system consisting of a waterproof basin filled with gravel and aquatic plants. Depuration takes place through the combined action of gravel, plants and microorganisms.

Phytodepuration technology has developed throughout Europe, particularly in the Nordic countries (United Kingdom, Denmark, the Netherlands, Austria, Sweden) since the mid-1980s. In Italy, as well as in the Mediterranean countries, phytodepuration started to be tested starting from the early 90s.

➤ Advantages

- Simple management;
- No (or very low) energy consumption and/or use of chemical reagents;
- Ability to adapt to variations of hydraulic load and pollutants in entrance;
- High depuration capacity of: organic load and bacterial load.

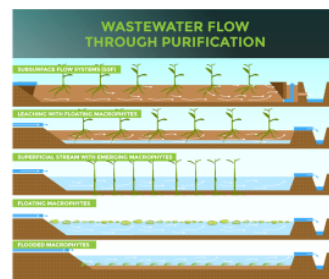


Fig. 17. General features of phytodepuration

2019-1-BG01-KA203-062371

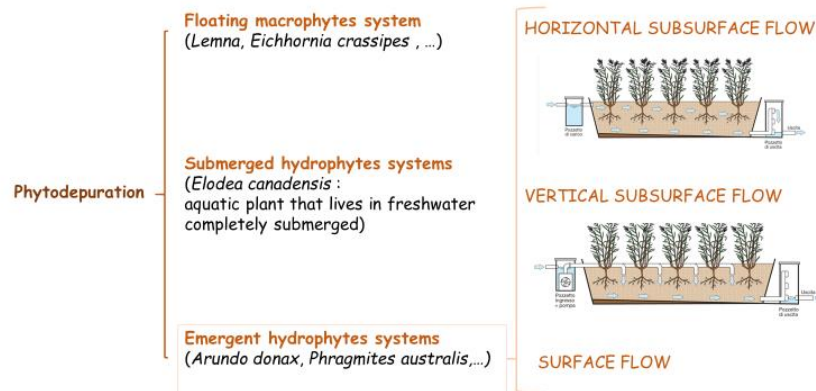


Fig. 18. Different phytodepuration strategies

Floating macrophytes, including water hyacinth (*Eichhornia crassipes*), are dominant invasive organisms in tropical aquatic systems and they may play an important role in modifying the gas exchange between water and the atmosphere. Hydrophytes are plants which live in water and adjust with their surroundings. They either remain fully submerged in the water or most of their body parts remain under the water. In the picture you can distinguish different examples of emergent hydrophyte system.

CONTAMINATED SOILS

Chemical compounds released in the environment

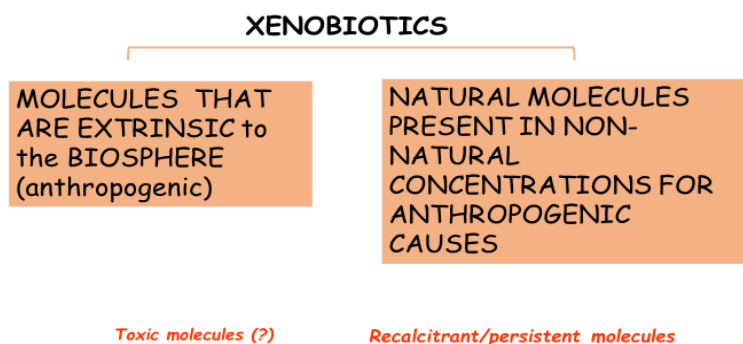


Fig. 19. Definition of xenobiotics

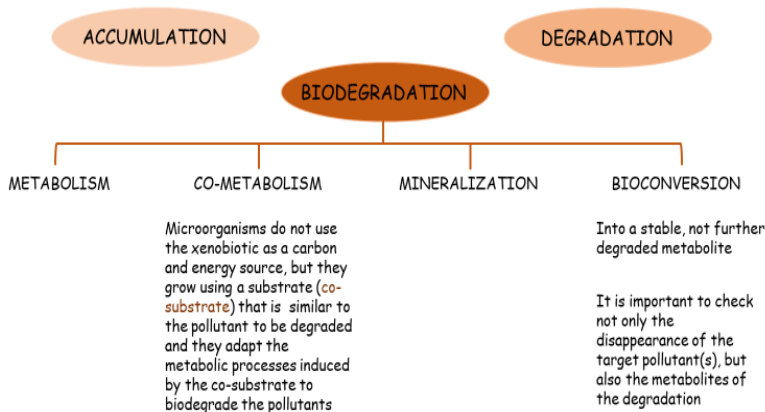


Fig. 20. Fate of xenobiotics in the environment

What are xenobiotic compounds (Fig. 19)? Which fate can they have in the environment (Fig. 20)? Xenobiotics are not necessarily molecules extrinsic to the biosphere, but they can also be natural molecules present in non-natural concentrations in the environment. They are not necessarily toxic molecules, but, generally, they are recalcitrant to the biodegradation.

Anthropogenic molecules can derive from different sources:

1) Petrochemical industry: - *fuels* (aliphatic and aromatic hydrocarbon mixtures), - *pure chemicals* for the chemical and pharmaceutical industry (alcohols, ethers, esters, aldehydes, frequently substituted with Cl⁻, amino or nitro groups)

2) Pulp and paper industries: use wood as raw material and produce pulp, paper and other cellulose-based products. - *the bleaching of paper* with chlorine-based products produces halogenated molecules including chloro-lignin

3) Synthetic plastic industry: styrene, vinyl chloride, solvents, cross-linking agents to produce polymers.

4) Pesticide industry: benzene and heterocyclic derivatives, urea, organophosphorus compounds

5) Pharmaceutical and cosmetic industry

6) Textile industry: reagents to produce synthetic fibers, detergents to soften fibers and pesticides for insect/moth control

7) Paint industry

8) domestic use of chemicals (personal hygiene products, cleaning products, ...)

A typical example of natural molecules present in non-natural concentrations is **olive mill wastewater**: they derive from olive milling to produce olive oil, but they have a high COD (see below) and need to be treated before being discarded in the environment.

Several questions have to answer before deciding whether a contaminated soil can be subjected to a bioremediation treatment. The first aspect to be considered is if we are referring to a chronic contaminated soil or a recent sudden contamination has occurred. In the first case, we can take some time to decide what to do, study in depth the situation, do some preliminary laboratory test and decide the best strategy. In the second case, we have to act promptly following the experiences that have been collected for the same contaminants and the same soils.

Please read the questions that need to be answered before DECIDING THE STRATEGY TO APPLY (Fig. 21).

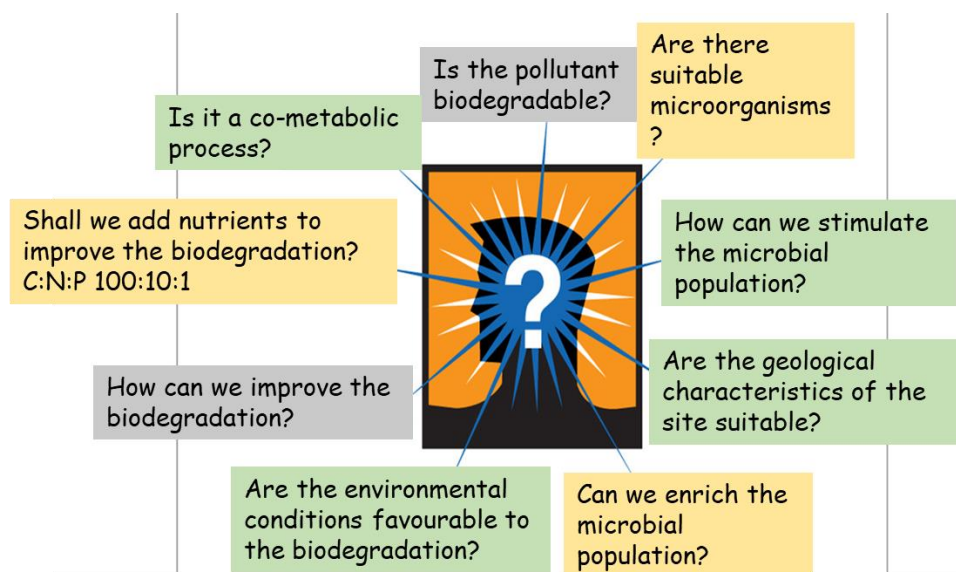


Fig. 21. How to decide which bioremediation strategy is the best one for your contaminated soil

A large array of bioremediation technologies can be used for the decontamination of contaminated sites (Fig. 22).

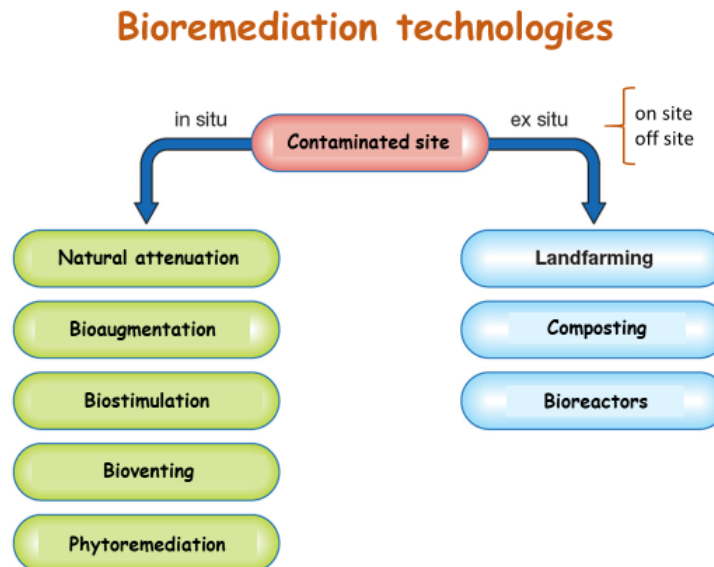


Fig. 22. *In situ* and *ex situ* bioremediation technologies

IN SITU means that the soil is not removed from its original location and is treated in the same location where the contamination occurred.

NATURAL ATTENUATION: within this term "a variety of physical, chemical, or biological processes" are included. These processes, under favorable conditions, act without human intervention to reduce the concentration, the toxicity and the mobility of contaminants in soil or groundwater. These *in situ* processes include biodegradation. A precise monitoring of the process is done, in order to make sure that the soil is remediated. The use of natural attenuation is often proposed as a remedial solution for benzene, toluene, ethylbenzene, and xylene (BTEX). More recently, natural attenuation has been proposed for chlorinated solvents, nitroaromatics, heavy metals, and other contaminants for which the scientific understanding and field experience are robust enough. Natural attenuation is applied when there is solid evidence that natural attenuation processes are transforming the contaminants to harmless products.

BIOAUGMENTATION: it is the addition of microbial cultures in order to speed up the rate of degradation of a contaminant. Indigenous microorganisms present in the contaminated areas may already be able to break down the contaminants, but their action may be inefficient and slow. Bioaugmentation requires studying the indigenous varieties present in the location to determine if biostimulation is possible. The same indigenous bacterial cultures can be isolated, cultivated and implemented into the location to boost the degradation of the contaminants. If the indigenous variety do not have the metabolic



2019-1-BG01-KA203-062371

capability to perform the remediation process, exogenous varieties with different degradation pathways can be introduced.

BIOSTIMULATION: is the addition of nutritional supplements for the indigenous microbiota to promote its metabolism. Usually, it refers to the addition of rate limiting nutrients like phosphorus, nitrogen, oxygen, electron donors to severely polluted sites to stimulate the existing bacteria to degrade the hazardous and toxic contaminants. The addition of these nutrients improves the degradation potential of the indigenous microorganisms. Among all the bioremediation techniques, biostimulation is considered to be the most efficient method for remediation of hydrocarbons, especially petroleum products and its derivatives. This is mostly due to the easy availability of carbon source which is one of the rate-limiting nutrients required by the indigenous microorganisms for their metabolic activities from the petroleum contaminants. In addition to the mentioned rate limiting nutrients, implementation of other nutrients rich organic matter can also trigger the remediation process extensively. In this context, it has been shown that the addition of bio solids (nutrient rich organic matter) obtained from the treatment of domestic sewage and inorganic fertilisers, rich in nitrogen and phosphorus, can improve and speed up the degradation rate of petroleum hydrocarbons.

BIOVENTING: also referred to as “Bioenhanced Soil Venting”, it is an *in situ* technology based on the natural stimulation of the indigenous biological activity with the introduction of oxygen through an air fluid; it is successfully applied to any aerobically biodegradable organic substance, in particular for the remediation of sites polluted by petroleum derivatives. The intake of air is at low flow rate as it is only designed to provide the oxygen needed to support microbial activity. Inside the polluted area, toxic compounds are removed from the air flow while organic compounds are biodegraded aerobically. The air is injected directly through one or more wells connected with vacuum pumps which provides the forced circulation of the air in the unsaturated contaminated soil. Passive bioventilation systems are also applied which exploit the natural exchange of air to transport oxygen, a one-way valve is installed on top of the external vent that allows air to enter only when the pressure inside the contaminated soil is higher than atmospheric.

PHYTOREMEDIATION: It is the treatment of contaminated soil with the use of plants to clean up soil contaminated with hazardous contaminants (Fig. 23). It is more properly defined as "the use of green plants and the associated microorganisms, along with proper soil amendments and agronomic techniques to either contain, remove or render toxic environmental contaminants harmless", because plants act in synergy with rhizospheric microorganisms that, very often, strongly collaborate with plants in the realization of the process. Phytoremediation is proposed as a cost-effective plant-based approach of environmental remediation that takes advantage of the ability of plants to concentrate elements and compounds from the environment and to detoxify various compounds. The concentrating effect results from the ability of certain plants called hyperaccumulators to bioaccumulate chemicals. The remediation effect is quite different. Toxic heavy metals cannot be degraded, but organic pollutants can be and are generally the major targets for phytoremediation. Several field trials confirmed the feasibility of using plants for environmental cleanup.

2019-1-BG01-KA203-062371

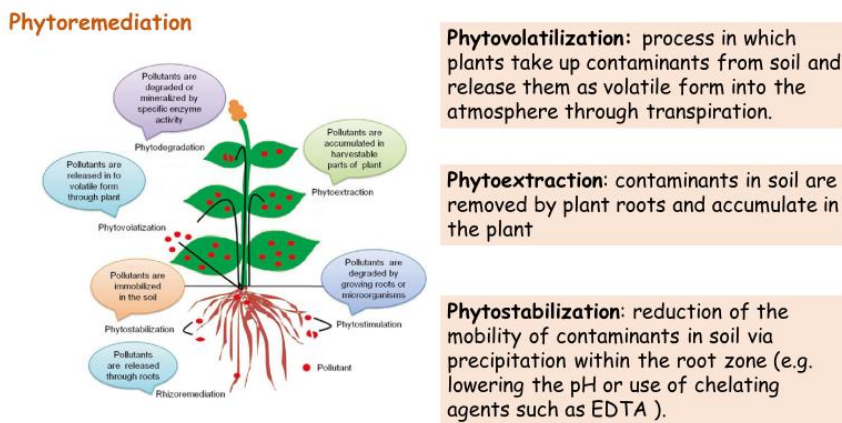


Fig. 23. Phytoremediation technology

EX SITU means that the soil is removed from its original location and treated very close to the site where the contamination has occurred (**on site**) or far away (**off site**).

LANDFARMING: it is a full-scale bioremediation technology, which requires excavation and placement of the contaminated soils, sediments, or sludges in a site where they can be treated; this is typically an *on site* remediation technology used to enhance the microbial degradation of hazardous compounds. Usually, liners and plastic covers are employed to control leaching of contaminants underground in order to avoid contamination of aquifers. Soil conditions are often controlled to optimize the rate of contaminant degradation, in particular:

- Moisture content (usually by irrigation or spraying).
- Aeration (by tilling the soil with a predetermined frequency, the soil is mixed and aerated).
- pH (buffered near neutral pH by adding crushed limestone or agricultural lime).
- Other amendments (e.g., soil bulking agents, nutrients, etc.).

Contaminated soil is usually treated in lifts that are up to 1 meter thick. When the desired level of treatment is achieved, the lift is removed and a new one is placed. Very often only the top of the remediated lift is removed, then the new lift is constructed by adding more contaminated soil to the remaining material and mixed. This serves to inoculate the freshly added material with an actively degrading microbial culture, and can reduce treatment times. This technique has been successfully used for years in the management and disposal of oily sludge and other petroleum refinery wastes. Generally, the higher the molecular weight of the target molecule (i.e., the more rings within a polycyclic aromatic hydrocarbon), the slower the degradation rate. Also, the more chlorinated or nitrated the compound, the more difficult it is to be degraded. Factors that may limit the applicability and effectiveness of the process

2019-1-BG01-KA203-062371

include: (a) large space requirements; (b) the conditions advantageous for biological degradation of contaminants cannot be reached, which increases the length of time to complete remediation, particularly for recalcitrant compounds; (c) inorganic contaminants are not biodegraded; (d) the potential of large amounts of particulate matter released by operations; and (e) the presence of metal ions may be toxic to microbes and may leach from the contaminated soil into the ground. Land farming, combined with other biological treatments, is widely used and has been successfully applied to many waste types, especially for disposal of oily sludge and other petroleum refinery wastes.

COMPOSTING: Composting is a process that works to speed up the natural decay of organic material by providing the ideal conditions for microorganisms to thrive (Fig. 24). The end-product of this concentrated decomposition process is nutrient-rich product (compost) that can help crops, garden plants and trees to grow.

Composting is the *aerobic* decomposition and transformation of the organic matter of urban solid waste by *aerobic microorganisms*. The process gives rise to a stable product that is called *COMPOST*.

The aim of composting

- production of a soil amendment used in agriculture
- production of a substrate for the production of mushrooms
- disposal of activated sludge derived from water treatment

Raw material for compost production

- straw and lignocellulosic materials
- agricultural by-products
- animal sludge
- organic fraction of urban solid waste
- activated sludge
- soil remediation (quality of the compost should be checked!)

Who is involved?

- mesophilic bacteria
- thermophilic bacteria
- yeasts
- Actinomycetes (Actinobacteria)

Fig. 24. Composting

proven effective in degrading or altering many types of contaminants, such as chlorinated and non-chlorinated hydrocarbons, wood-preserving chemicals, solvents, heavy metals, pesticides, petroleum products, and explosives. Compost used in bioremediation is referred to as “tailored” or “designed” compost in that it is specially made to treat specific contaminants at specific sites. The ultimate goal in any remediation project is to return the site to its pre-contamination condition, which often includes revegetation to stabilize the treated soil. In addition to reducing contaminant levels, compost advances this goal by facilitating plant growth. In this role, compost provides soil conditioning and also provides nutrients to a wide variety of vegetation.

Compost bioremediation refers to the use of a biological system of micro-organisms in a mature, cured compost to sequester or break down contaminants in water or soil. Micro-organisms consume contaminants in soils, ground and surface waters, and air. The contaminants are digested, metabolized, and transformed into humus and inert by-products, such as carbon dioxide, water, and salts. Compost bioremediation has

BIOREACTORS: the treatment of a contaminated soil in a bioreactor is the finest remediation technology although it is the most expensive (Fig. 25). It is an *ex situ off site* technology: the soil, after its removal from the original site, can be treated far away from its original location. The soil is treated in slurry phase inside a bioreactor made of different materials (glass, steel, concrete or other materials) and all the parameters of the remediation process are monitored and controlled to make the process as effective as possible (pH, redox potential, temperature,

concentration of the pollutant(s), presence of degradation metabolites). Microorganisms can therefore work in their optimal conditions. The treatment of a contaminated soil in a bioreactor is usually applied for soils contaminated by particularly recalcitrant molecules that are difficultly removed by other remediation technologies (e.g. highly chlorinated molecule).

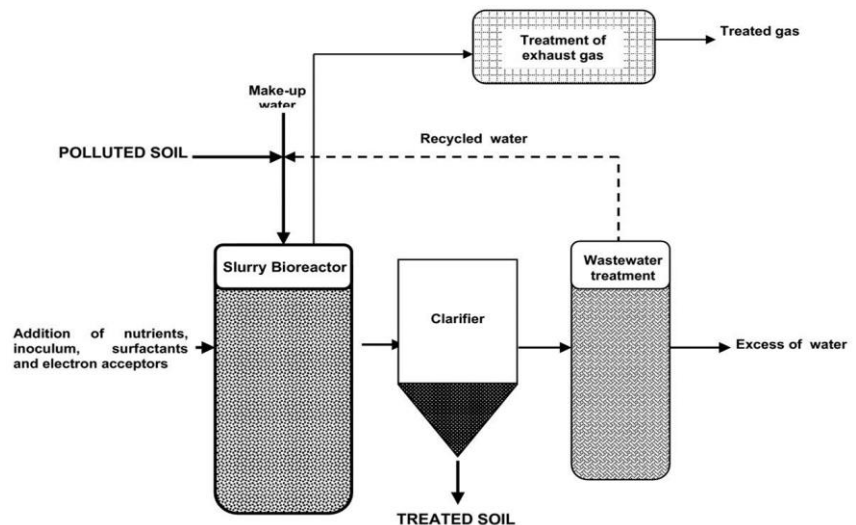


Fig. 25. Example of soil treatment in a bioreactor

STRENGTHS and WEAKNESSES of bioremediation technologies

STRENGTHS

- Reduced costs with respect to chemical and physical strategies (lower energy costs);
- Reduced environmental impact: the soil can be re-used in situ;
- The problem (i.e. the contamination) is solved (pollutants disappear, they are not simply moved from one site to another);
- Acceptability by the public opinion.

WEAKNESSES

- Problems of the bioavailability of pollutants;
- Problems in case pollutants are more than one;
- Problems of adequate environmental conditions (pH, temperature, oxygen availability).



References

- Meegoda JN, Li B, Patel K, Wang LB. 2018. A review of the processes, parameters, and optimization of anaerobic digestion. *International journal of environmental research and public health*, 15(10): 2224
- Wang P, Wang H, Qiu Y, Ren L, Jiang B. 2018. Microbial characteristics in anaerobic digestion process of food waste for methane production—A review. *Bioresource technology*, 248: 29-36
- Azhar S HM, Abdulla R, Jambo SA, Marbawi H, Gansau JA, Faik AAM, Rodrigues KF. 2017. Yeasts in sustainable bioethanol production: A review. *Biochemistry and Biophysics Reports*, 10: 52-61
- Prasad RK, Chatterjee S, Mazumder PB, Gupta SK, Sharma S, Vairale MG, .Gupta DK. 2019. Bioethanol production from waste lignocelluloses: A review on microbial degradation potential. *Chemosphere*, 231: 588-606
- Sivagami K, Sakthivel KP, Nambi IM. 2018. Advanced oxidation processes for the treatment of tannery wastewater. *Journal of environmental chemical engineering*, 6(3): 3656-3663
- Krzeminski P, Tomei MC, Karaolia P, Langenhoff A, Almeida CMR, Felis E, Fatta-Kassinos D. 2019. Performance of secondary wastewater treatment methods for the removal of contaminants of emerging concern implicated in crop uptake and antibiotic resistance spread: A review. *Science of the Total Environment*, 648: 1052-1081
- Abatenh E, Gizaw B, Tsegaye Z, Wassie M. 2017. The role of microorganisms in bioremediation- A review. *Open Journal of Environmental Biology*, 2(1): 038-046
- Wan X, Lei M, Chen T. 2020. Review on remediation technologies for arsenic-contaminated soil. *Frontiers of Environmental Science & Engineering*, 14(2): 1-14.



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);

2019-1-BG01-KA203-062371

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).

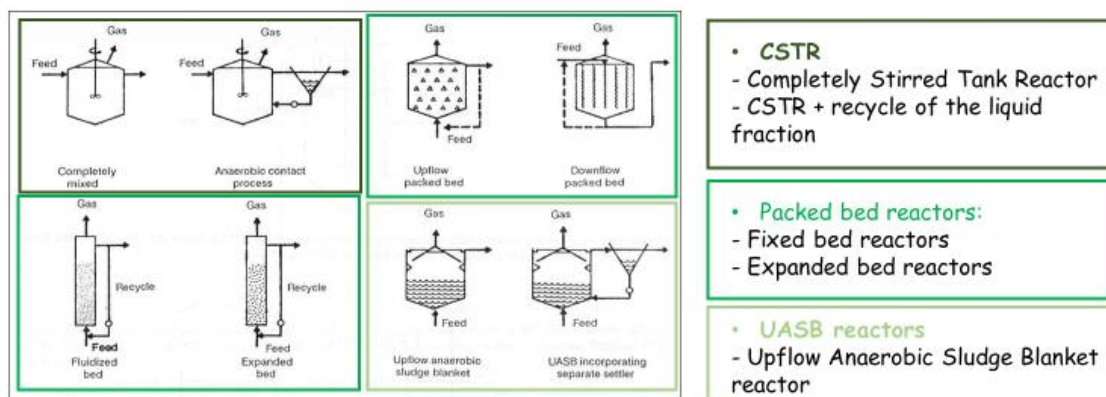


Fig. 1. Reactors for biogas production



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).

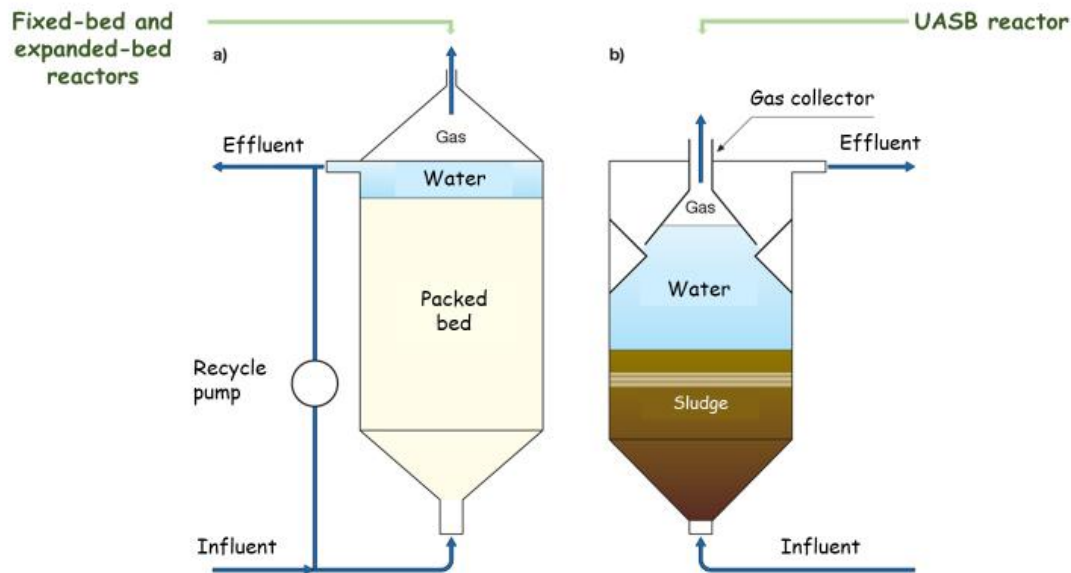


Fig. 2. Fixed bed/Expanded bed reactors (left) and UASB reactor

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).

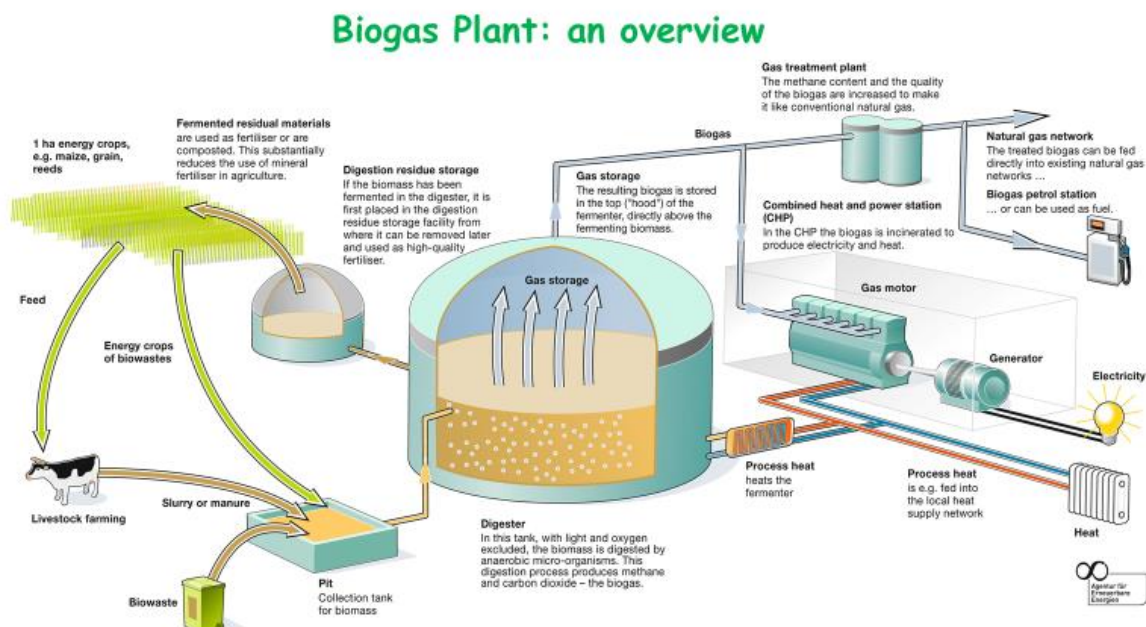


Fig. 3. A biogas producing plant

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)

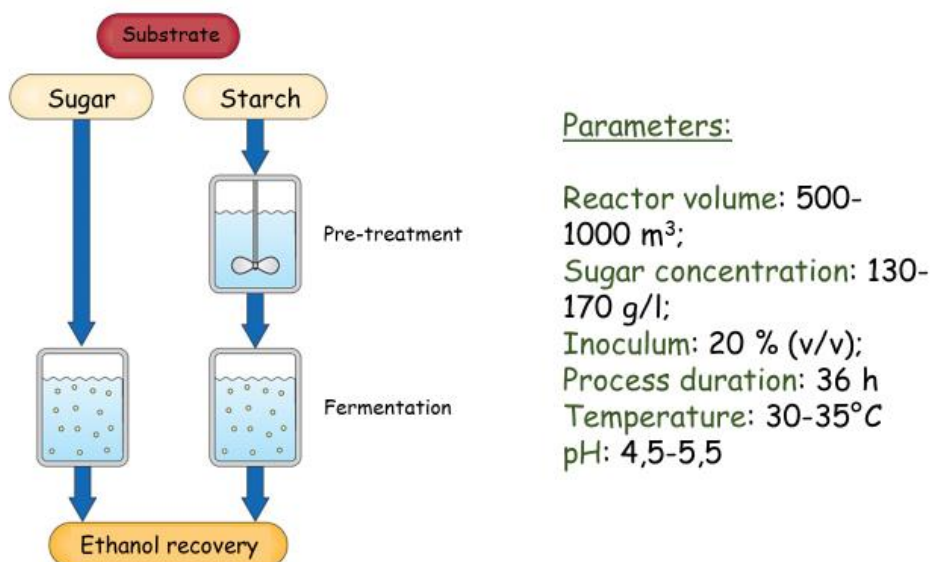


Fig. 4. Schematic production of bioethanol from sugar crops

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

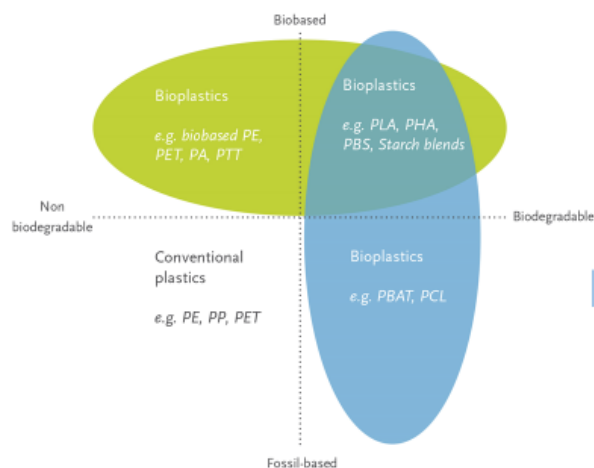
2019-1-BG01-KA203-062371

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)

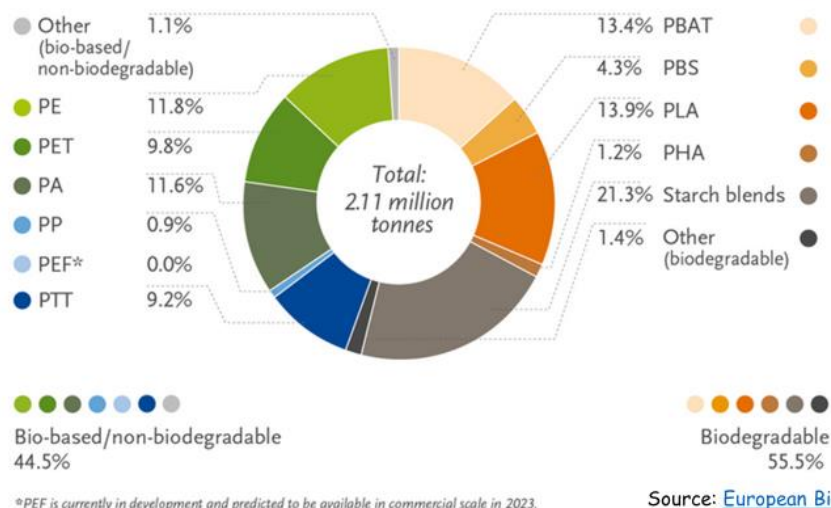


Fig. 7. Main bioplastics produced worldwide

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.

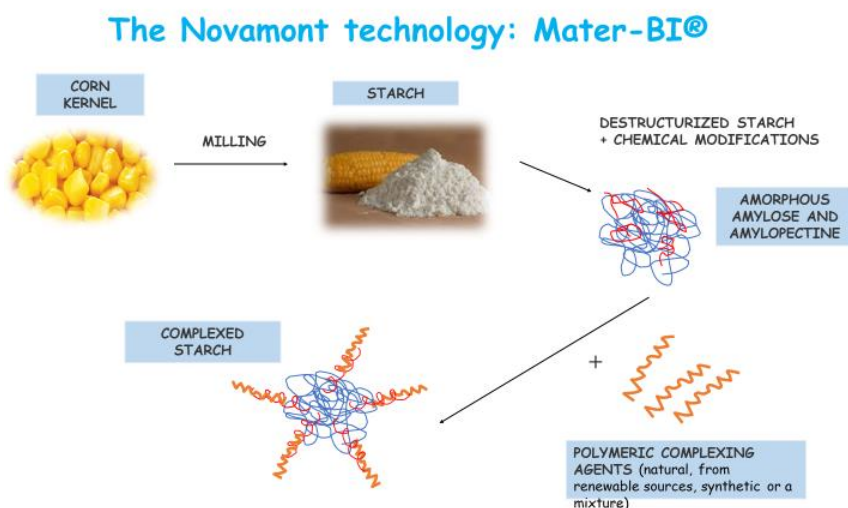


Fig. 8. Novamont's starch-based technology

2019-1-BG01-KA203-062371

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 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

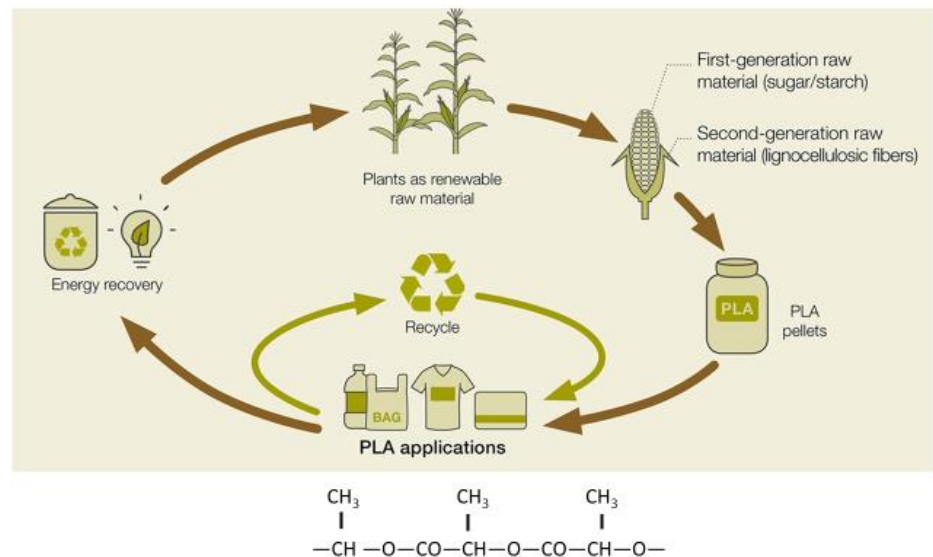


Fig. 9. PLA life cycle

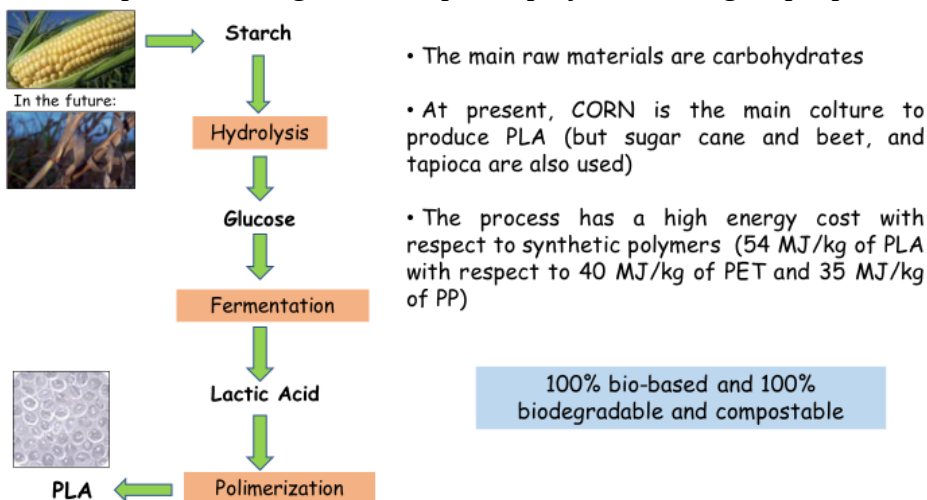


Fig. 10. PLA production from starch

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.

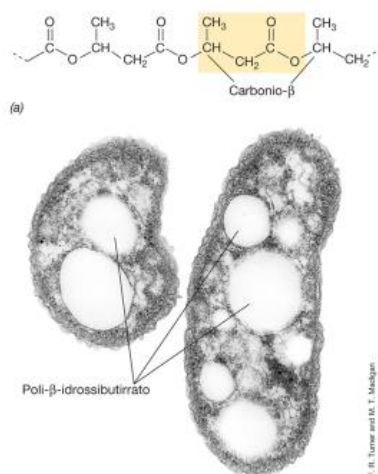
3. Polyhydroxyalkanoates

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).

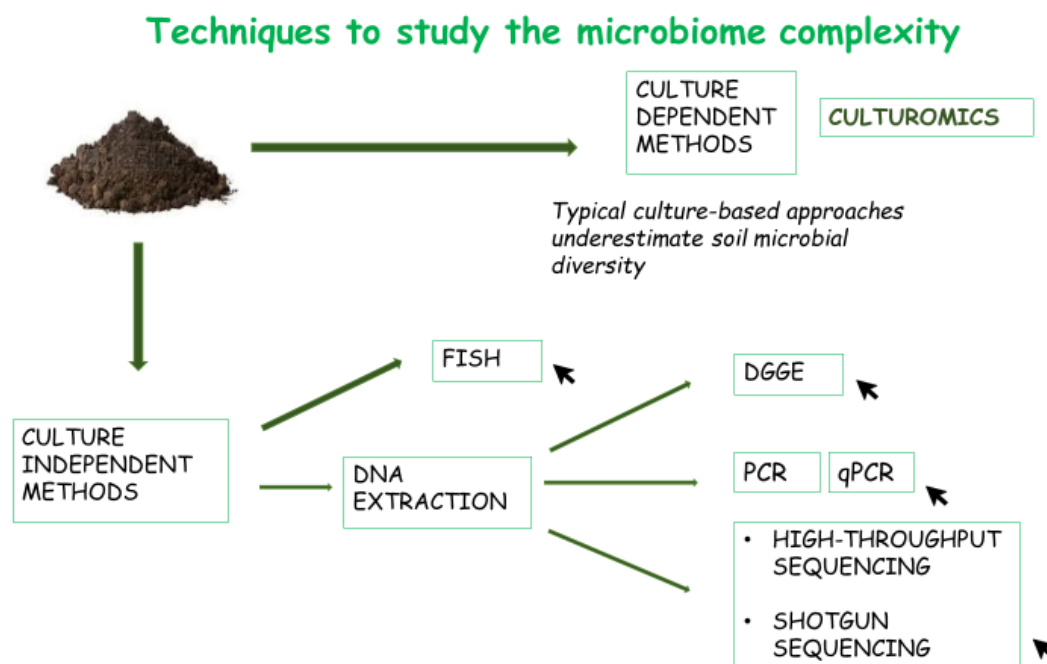


Fig. 12. Techniques to study the microbiome complexity

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...), 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

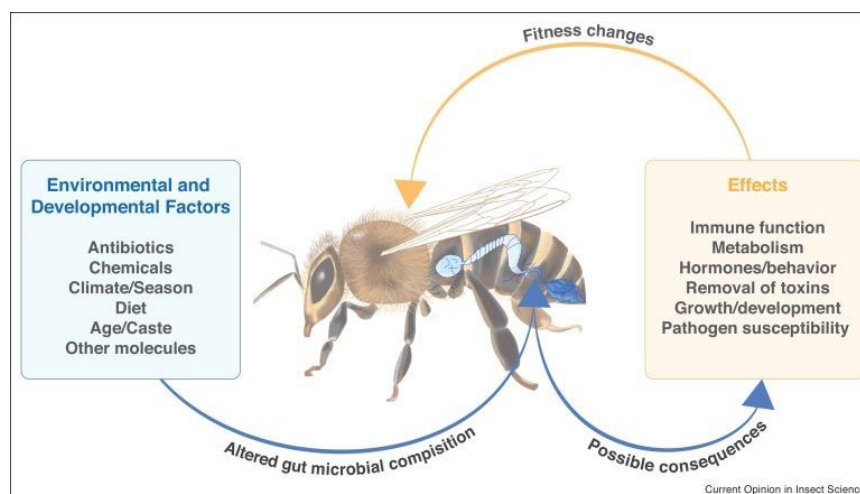


Fig. 14. Consequences of altered gut microbial compositions in bees



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

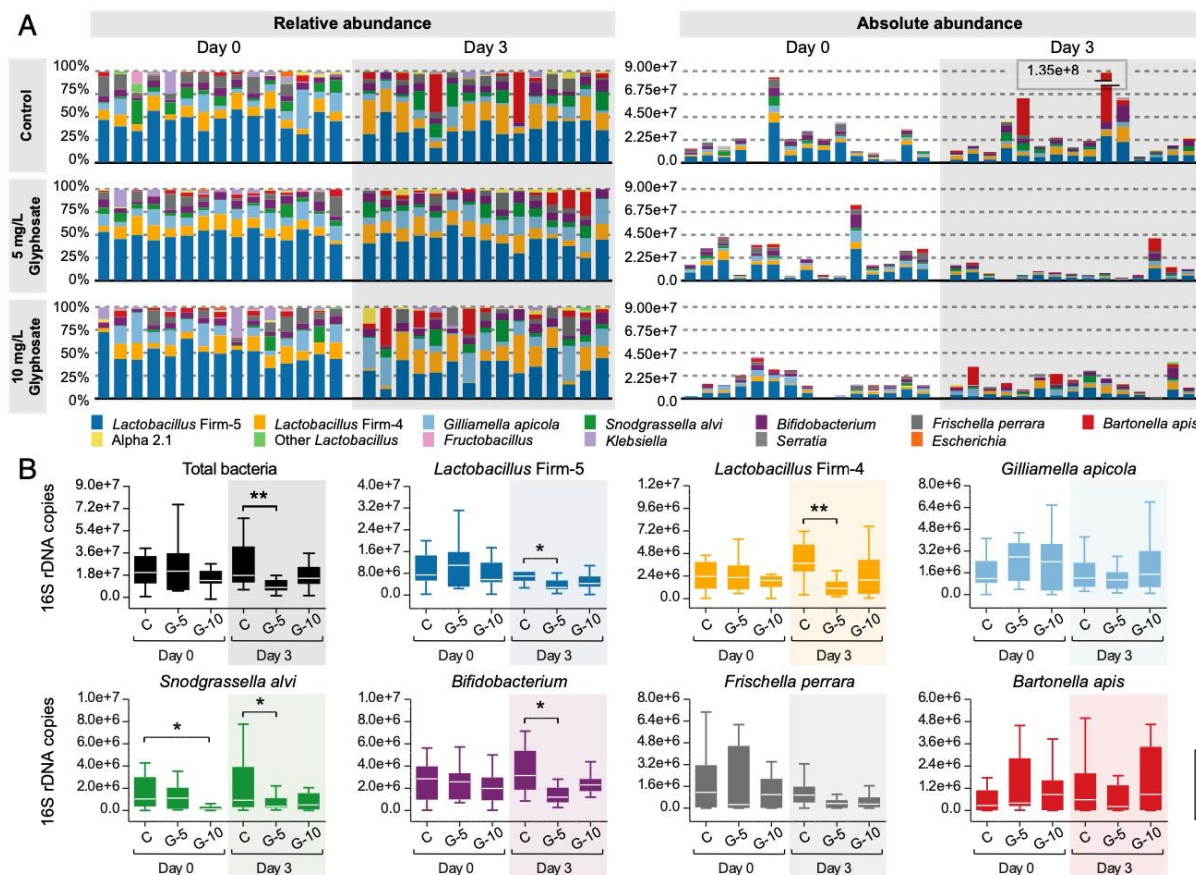


Fig. 15. Analyses on the bee gut microbiota

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?

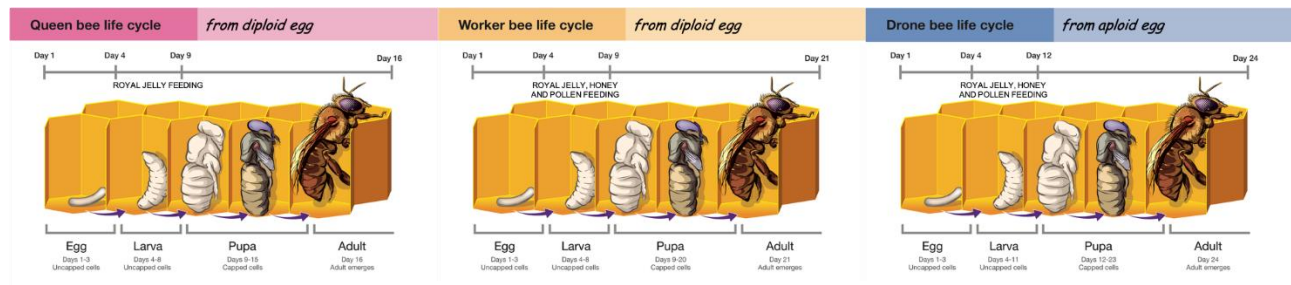


Fig. 16. The growth cycle of a bee

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.

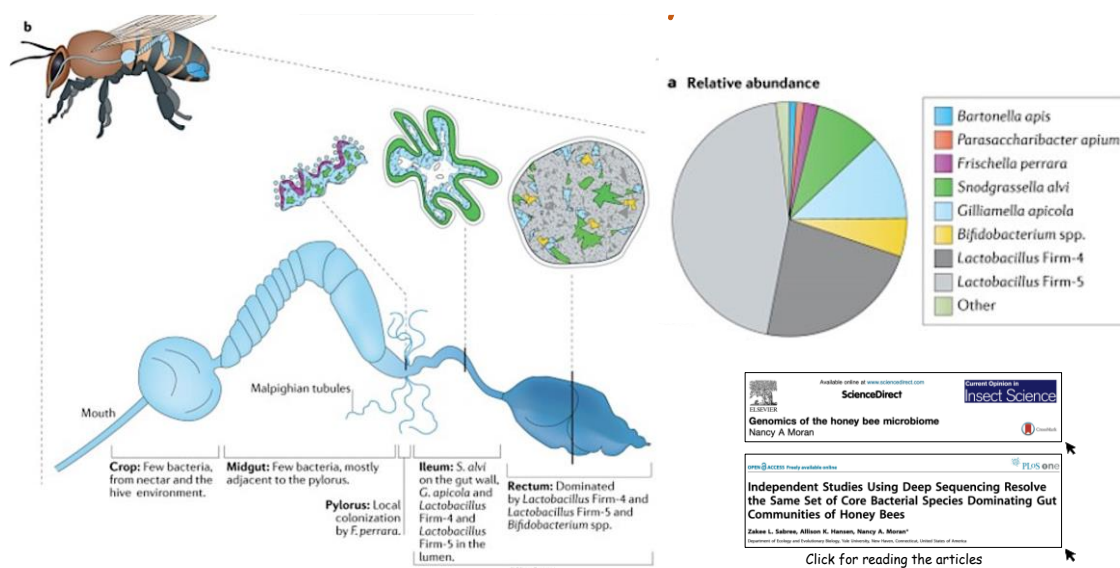


Fig. 17. The core bee microbiota

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> Mori2 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.