THE IMPORTANCE OF METAGENOMICS RESEARCH IN HUMAN ECOLOGICAL NICHES AND THEIR ROLE IN THE DIAGNOSIS OF NONINFECTIOUS DISEASES

Marcin Weiner¹, Maria Kubajka²

¹Pope John Paul II State School of Higher Education in Biała Podlaska
²National Veterinary Research Institute in Puławy


Summary: Human body is a complex system that is affected by a significant number of microscopic organisms called the microbiomes. The dynamic development of science has led to innovative discoveries in the field of microbiology. This in turn has extracted new field, metagenomics, thanks to which it became possible to perform detailed analysis of individual groups of bacteria and to determine their effects on preserving a good health. One of the biggest scientific projects that would investigate the influence of microbiomes on humans is HMP (Human Microbiome Project). As part of it the research is being conducted leading to characterize human microbiome at the level of nucleotide sequence of the entire genomic DNA. The microflora of the skin, oral cavity, respiratory tract, digestive tract (intestines), genitourinary system has an essential role in the homeostasis. In the last year the carried research proved that it is a vital part of the human organism in preserving a good health. Any changes in its composition may lead to systemic diseases. Pathological changes affect the outcome of the interaction within the microflora that includes species of commensal and pathogenic bacteria, as well as immunology and genetics of the host. Metagenomics research will contribute not only to the recognition of new, so far unidentified by the bacteriological methods microorganisms, but most of all they will serve as a basis to understand the relationships between the human organism and in-dwelling microorganisms. Thanks to the development of the metagenomics or the NGS (Next Generation Sequencing) it will be possible to discover new metabolic pathways and bidirectional links of bacteria with human metabolism. This will help in finding new therapeutic methods in the treatment of many noninfectious diseases so far considered as civilization diseases or genetically conditioned.

Keywords: noninfectious diseases, HMP, metagenomics, NGS

Introduction

Human body is a complex system affected by a large number of microscopic organisms called the microbiomes (Frańczuk, Jagusztyn-Krynicka 2012). Microorganism cells count ten times more than human cells and the large group of them exist in the digestive tract, and in particular in the intestines (10¹¹ – 10¹² bacterium cells) (Olszewska, Jagusztyn-Krynicka 2012). Pasteur had already conducted the first study on the role of intestinal microorganisms (Pasteur 1885). A large number of bacteria influence the functioning of human organism, causing the changes in the immune system, affecting the proliferation of epithelial cells or improving metabolism that affects significantly the physiology and human health (Frańczuk, Jagusztyn-Krynicka 2012, Handelsmann et al. 1998). To date the studies on the various kinds of microorganisms have been limited to the species that could grow in the laboratory conditions (Stahl et al. 1985). Dynamic development has led to innovative discoveries in the field of microbiology and it identified a new field of science thanks to which it became possible to perform a detailed analysis of the particular groups of bacteria and to determine their effects on the health. Metagenomics, also called the ecogenomics, or the microorganisms population genomics, is a new field of science that consists in cloning of DNA coming directly from the natural environments and then sequencing a huge number of genomics libraries (Woese 1987). In order to explore the human microbiome large research programs have arisen in Europe and the USA. One of the biggest research project that investigates the influence of the microbiome on human is the HMP (Human Microbiome Project), under which various tests are carried out that characterize...
human microbiome on the nucleotide sequence level of the entire genomic DNA of the human microorganisms (including non-laboratory populations).

**HMP**

HMP was created in 2007 and its objective was to carefully examine human microflora focusing not only on the intestine but also on other ecological niches such as: genitourinary tract, respiratory tract, skin and oral cavity (Binek 2012). Initially the study involved pathogenic microorganisms but the significance of the beneficial effects of the microflora on the physiology and organism immunity was omitted. The project was created in order to generate the research resources that would enable the comprehensive characteristic of the human microflora and its analysis in relations to the population, age, health condition, living environment, nutrients, impact on various diseases and taking medicines. Advances in the DNA sequencing technology have created a new field of study on microorganisms, allowing the standardization of methods including the collection of samples. In the future, the knowledge acquired in this way is due to improve the health of population through monitoring and manipulation of human microbiome. Thanks to these modern techniques scientists are able to isolate the entire unknown until now groups of microorganisms. The HMP project includes the sequencing of about 600 bacterial genomes where the total of the entire group of the reference genomes accounts to 1000. This will lead to the continuing metagenomics analysis that would characterize the complexity of microorganisms, in particular, ecological niches (i.e. digestive tract, genitourinary tract, respiratory tract and skin) (Gill et al. 2006).

The objectives of the HMP have been designated by the National Institute of Health in the USA and include (http://commonfund.nih.gov/hmp):

- Defining of the „core” microflora for humans, i.e. group of intestinal microorganisms forming a pool of genes present in the given environment (Fig. 1).
- Assessment of the correlation of human microflora changes with changes in the health condition.
- Genetic analysis of the available reference strains that contain information about the complexity of human microorganisms. Other advanced technologies include transcriptomics that describes the biological properties of whole microorganisms providing insight into the interaction on the human organism.
- Increase of the capacity to generate the whole sequences of bacteria genome, the development of new techniques for culture.
- Development of computer analysis that will be able to gather a large number of data (new generation sequencing platform that produces large amounts of sequences together with the method of its analysis).

Establishing of data analysis and coordination CENTER (DACC), on this page there will be a huge amount of information on the project, along with its findings and conclusions. DACC will perform several specific functions, such as: storage and distribution of raw data, coordination of data analysis, development of the data download tools for the scientific community, coordination of the development of standards, establishing operations in international projects, as well as solving ethical, legal and social issues related to the subject of research on the human flora.

**Figure 1.** Factors influencing changes in the balance of group of microorganisms (Turnbaugh et al. 2007)
HMP is carried out in three stages. The first step consists of obtaining samples that need to be diverse in terms of race, ethnicity and demographic and then the reference bacterial genomes of human microflora, collected from five areas of the body, are created (digestive tract, genitourinary tract, oral cavity and skin). During this stage it is important to perform a sequence of 16S rRNA gene on the samples from five areas of the body for women and four areas of the body for male volunteers, which will be used to create a widely available database. An innovative method for the culture of microorganisms, which were not able to grow in vitro and the improvement of existing methods are also important. The second stage consists of a detailed study of the microflora. Detailed sequence analysis and the number of examined people influence a definition of a “full” human microbiome. New species of bacteria, which are similar to those in human organism or those that can be found in various animal species, are also taken into account. The third step aims to compare samples taken from people that are different from each other in terms of origin, race, social affiliation etc. This stage is also affected by the general condition of the patient. In the course of research a relationship between disease state and the current (changing during the development of the disease) microflora of the organism is being looked for.

The obtained information will create diagnostic tests, new technological methods, changes in global nutritional network and new guidelines for health education (Peterson, Guyer 2009). The ultimate goal of HMP is an indication that there are opportunities to improve human health through monitoring or manipulation of a human microbiome (Peterson, Guyer 2009, Turnbaugh et al. 2007).

NGS – Next Generation Sequencing

Over the past 30 years Sanger method was the only effective method for DNA sequencing. The first commercial launch of the pyrosequencing platform in 2005 marked the beginning of a new era of genomic analysis of high throughput, defined as next-generation sequencing. NGS has the ability to sequence DNA with exceptional speed; it is also characterized by an unprecedented throughput, allowing scientists to study biological systems at a level not previously possible (Pallen et al. 2010). With each step, NG3 systems performance triples, reducing at the same time by double the price of one sequence. NGS is a technique for massive parallel sequencing on a plate in wells with a picoliters volume, which allows you to make a million sequences at once, in one cycle distribution depending on the platform. The first test was conducted on the 454 sequencer (Roche), which generated fifty times more DNA sequences for sixth of the sum of the costs compared to the previous method. NGS has had a huge impact on the study of microbial diversity in environmental and clinical samples. Genomic DNA is extracted from the sample and converted into NGS library and then it is sequenced. The output sequence is consistent with the known reference sequences for microorganisms which, as expected, are present in the sample. In addition, it becomes important to create a new directory in order to confirm the presence of known and potential new species. Qualitative genomic studies rely on the analysis of the read sequence, while the quantitative interpretation applies to individual species of the microorganisms.

Next-generation sequencing opens up new opportunities to improve the knowledge on the microbial diversity of the environment, as well as allows obtaining data on human microflora, enabling fast, efficient and cost-effective identification of microorganisms. The enormous scale of the tests, parallel to accuracy that will be skillfully used, will lead to an earlier prediction and prevention of the diseases with their simultaneous diagnostics. All of these discoveries will improve the process of prevention, provided that the obtained data will be skillfully linked and analyzed. The dynamics of research on the human genome in terms of understanding of the etiology of particular diseases has a positive prospect for the development of medicine (Chan 2005, Bisecker 2013).

The importance of the microflora of the skin, oral cavity, respiratory tract, digestive tract (intestines) and the genitourinary system in the homeostasis

The viruses, bacteria and Eukaryote colonizing the human body are called the microbiota, and the genes that are encrypted by them are called microbiome (Handelsman 2004). A significant part of the microbiota can be found on the skin, in the oral cavity and in the vagina, however, its major part is in the digestive tract. Shaping of human microbiome composition begins just a few minutes after the birth, when it comes to colonize the digestive tract and skin and then it changes throughout life (Mira et al. 2010).

The number of microorganisms present on the skin of an adult comes up to 2 x 10^4 CFU/cm2, although in some areas they may not exceed 10^3 CFU/cm2. Most bacteria are present in the superficial layers of the epidermis and upperparts of thehair follicle, showing commensal, mutualistic or pathogenic properties. Human microbiological interaction with the external environment begins at birth. Initially, aseptic newborn gets in contact with the
environment during delivery and at this time microorganisms are penetrating the skin. On the surface of the skin there are coagulase-negative Staphylococci (S. epidermidis representing approx. 64% of the isolates, but also S. hominis, S. warneri, S. haemolyticus, S. saprophyticus and S. capitis) with Propionibacterium acnes, Corynebacterium spp., and Pseudomonas spp. Initially, this composition has been considered as skin microflora, but after the molecular analysis of HMP (Human Microbiome Project) the presence of the microorganisms that are difficult to culture in vitro has also been indicated. In the skin microbiome bacteria of the genus Acinetobacter, Brevibacterium, Micrococcus, Gemella, Prevotella, Enterococcus, anaerobes and fungi of the genus Malassezia are also present. On the surface of healthy or changed skin, the biofilm is the phenotype of population of different microorganisms. Thanks to modern research techniques it is being proved that the biofilm is associated with the onset of various skin layers infections. On the surface of still exfoliating undamaged epidermis, the microorganisms of the natural skin flora have sufficient conditions to settle and therefore fulfil the protective and immunomodulatory functions. The physical receptors barrier protects the host tissue against colonization of pathogenic microorganisms. Also mediates in the process of adhesion of microorganisms to the eukaryotic cells and supports the biochemical modification of the environment by the products of primary colonizers. Skin microbiome competes with pathogenic microorganisms producing substances of biostatic (biocidal) nature - bacteriocins or BLS (Bacteriocin like substances). The bacteriocins are the protein metabolites produced by Gram-positive and Gram-negative bacteria, showing antimicrobial activity. The role of immunomodulatory natural microflora allows maturation of the immune system of the host and maintains it at a permanent level of activation. Between minor wounds arising on the skin surface and the microorganism's biofilm the homeostatic state is being established. The microorganisms reach the state of mutualism or commensalism, received by the host organism as a natural microbiome functioning in changed tissues (Różalska et al. 2013).

The oral cavity defines the beginning of the digestive tract, where over 400 species of bacteria can be found on the surface of the teeth, in the saliva, on the tongue, hard palate, and in gingival pockets. Differentiation of the surface supports the development of the microflora, due to different biological and physicochemical properties, i.e. neutral pH, unchanging temperature, or different microelements. In the oral cavity there are bacteria of the genus: Staphylococcus, Streptococcus, Enterococcus, Streptopeptococcus, Lactobacillus, Fusobacterium, Corynebacterium. With sinus and middle ear infections S. pneumoniae is present, and with periodontal infections - Porphyromonas gingivalis. Bacteria that form plaque have characteristic role. It is a natural form of the tooth surface and multispecies biofilm formed by various pathogens. In the case of imbalance in the flora of the oral cavity, bacteria cause the appearance of tooth decay and inflammation of the soft tissues (Fiedurek, 2014).

Microbiome of the respiratory tract was being analyzed within the framework of a new project of the National Heart, Lung and Blood Institute, USA (NHLBI) from 2011. Because of the previous inability to breed bacteria, it is being believed that the respiratory tract below the larynx is in the sterile physiological state. Currently, based on a molecular identification of the microorganisms of the 16S rRNA sequences, various bacteria inhabiting the lower respiratory tract are being detected. The research on the respiratory microbiome was focused on people with asthma, cystic fibrosis and on healthy smokers and non-smokers. In the course of the study, researchers encountered difficulties to collect material from the lowest parts of the respiratory tract and to protect it against contamination. Scientists studying lung microbes used a specimen collection technique using two fiber-optic bronchoscopes. One was being used to take samples from the oral cavity and nasopharynx whereas the second one, from the lower respiratory tract, by means of a protected brush. During the entire operation, each step must have been completed with particular care in order to prevent contamination from the various sections of the respiratory tract. It was being found that in healthy smokers and non-smokers various kinds of bacteria were being isolated. In 75% of cases the detected bacteria were: Firmicutes, Proteobacteria, Bacteroidetes and Fusobacteria and the species of Streptococcus, Pseudomonas, Fusobacterium Prevotella, whereas in more than 50% of cases Porphyromonas, Veillonella and Haemophilus were being identified. The microbiome of upper respiratory tract examined from sputum and bronchial aspirate is significantly different from the microbiome of the lower bronchial tree, which showed greater diversity. There is a group of microorganisms common to the whole respiratory tract, but there are bacteria that colonize the various sections of the system. During extensive research on the microbiome present in the lungs removed during transplantation, it is being found that in one lung there might be different bacterial flora. An interesting observation was performed on newborns vaccinated with a heptavalent pneumococcal conjugate vaccine. In microbiome of vaccinated infants, the participation of the strains contained in the vaccine has declined, while the participation of Haemophilus strain has increased as compared to unvaccinated infants. These studies have attracted attention of the medical world by showing that any interference can disrupt already unstable balance of the human microbiome. Mutual interaction of the human organism with the bacterial microflora of the respiratory tract has been associated with infectious, noninfectious and chronic diseases. Certainly there is a balance between human body and the microorganisms...
inhabiting it, but the origin of the respiratory microbiome is still unknown. Only hypothesis is being drawn that the source of the lungs colonization is the environment and the bacteria of the digestive tract. In a healthy organism, there is an enormous variety of bacterial flora which ensures homeostasis. In patient the diversity is being reduced causing the disturbance of the proportion between various bacterial species (Górecka et al. 2014).

The biggest habitat of the physiological microflora is the digestive tract, and the number of bacteria varies depending on the section, and so in the stomach and duodenum it reaches the number in the region of $10^1$ - $10^2$ CFU/ml, in the small intestine from $10^3$ - $10^5$ CFU/ml, and in the large intestine $10^{11}$ - $10^{12}$ CFU/ml. Metagenomes analysis of the human microbiome has revealed that the intestine takes 3.3 million genes, which is about 150 times more than in the human genome, and intestinal microbial diversity is estimated at over 1,000 species (Reiff and Kelly 2010). Intestinal microorganisms form a distinct phylogenetic group that has evolved with its host, entering into mutual cooperation. The attempts of quantitative and qualitative evaluation of intestinal microbiome were repeatedly undertaken but only on the basis of the 16S rRNA sequence, real results were being obtained.

It was being found that 80% of the microbial phenotypes was not being bred using techniques of classical bacteriology. Thanks to the analysis of the 16S rRNA the variation of the microflora was being demonstrated obtained. It was being found that 80% of the microbial phenotypes was not being bred using techniques of classical bacteriology. Thanks to the analysis of the 16S rRNA the variation of the microflora was being demonstrated.

Intestinal microorganisms are the components of the cell wall. They are present on the surface of intestinal epithelium cells, dendritic cells of the respiratory tract and macrophages (Binek 2012).

Vaginal microflora is made up from bacteria of the genus Lactobacillus, commonly known as lactic acid bacteria, since they lead to the formation of lactic acid that maintains the low pH and produces hydrogen peroxide. In such an environment multiplication of pathogens, that cause diseases in the reproductive system, is inhibited. In addition to that, they can adhere to the vaginal epithelium, protecting it in this way from other bacteria. Vaginal bacterial infection is characterized by imbalance in the number of different types of bacteria, which form the physiological vaginal environment. It leads to an excessive proliferation of bacteria and fungi inhibiting it, but the origin of the respiratory microbiome is still unknown. Only hypothesis is being drawn that the source of the lungs colonization is the environment and the bacteria of the digestive tract. In a healthy organism, there is an enormous variety of bacterial flora which ensures homeostasis. In patient the diversity is being reduced causing the disturbance of the proportion between various bacterial species (Górecka et al. 2014).

**Table 1.** Dependences of the intestinal microflora, providing biological balance (Fiedurek 2014)

<table>
<thead>
<tr>
<th>Intestinal microflora</th>
<th>Bacteria</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>absolutely pathogenic</td>
<td>Salmonella</td>
<td>Interferes with the functions of the digestive tract by producing toxins and carcinogens.</td>
</tr>
<tr>
<td></td>
<td>Shigella</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Staphylococcus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pseudomonas</td>
<td></td>
</tr>
<tr>
<td>relatively pathogenic</td>
<td>Klebsiella</td>
<td>Discloses a negative impact during disorders of the digestive tract.</td>
</tr>
<tr>
<td></td>
<td>Proteus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Escherichia</td>
<td></td>
</tr>
<tr>
<td>beneficial</td>
<td>Lactobacillus</td>
<td>Prevent the growth of pathogens, stimulating important functions of the digestive tract.</td>
</tr>
<tr>
<td></td>
<td>Bifidobacterium</td>
<td></td>
</tr>
</tbody>
</table>

Commensal bacteria together with the immune system have many common features with pathogenic microorganisms. Host cell receptors or membrane proteins, i.e., lipopolysaccharides, peptidoglycan or lipoteichoic acids are the components of the cell wall. They are present on the surface of intestinal epithelium cells, dendritic cells of the respiratory tract and macrophages (Binek 2012).
that were previously present in smaller quantities and in particularly the anaerobic bacteria such as *Gardnerella vaginalis*, the presence of which is considered as a marker of vaginal bacterial infections. A significant decrease in lactobacilli also has been noted.

**Change the microflora and metabolic diseases**

In recent years, detailed studies of microbiota led to the conclusion that it is a necessary component of the human body to maintain a good health. Any changes in its composition may lead to systemic diseases. Pathological changes affect the outcome of the interaction within the microflora that includes species of commensal and pathogenic bacteria, as well as immunology and genetics of the host. The knowledge of the nucleotide sequence of the genetic material of microorganism's populations not only helps to identify the taxonomic levels of prokaryotes, but also it contributes to the creation of catalogs of genes and their corresponding functions. The literature provides new evidence that the increased prevalence of obesity, diabetes, and cancer is attributed to the changes in the human genome. It has to be taken into account that the microflora of the body is a new important environmental factor. Recent research may affect the treatment or prevention of most diseases of civilization, opening up a new area of prevention or treatment involving manipulation of the current composition of the microflora (Griece, Segre 2011).

**Obesity**

Obesity is the result of collection, storage and energy distribution processes occurring in the body. This positive energy balance makes the excess of energy supplied in the form of liquid foods, semi-liquid and solids must be stored in a physiologically programmed to this location i.e. the adipose tissue. Results of studies in recent years have shown that in the processes of food assimilation and energy regulation, microorganisms constituting the intestinal flora play an important role (Hattori, Taylor 2009). Pioneering research consisted of implantation of the human intestinal microflora to the intestines of mice, and the introduction of high-fat diet, which resulted in the change of the bacterial flora already in the first day. The activity of metabolic pathways have also changed, which could also affect the change in the expansion of the genes present in the intestinal microorganisms. The analysis of the final part of the digestive tract has helped to characterize the microorganisms dwelling there. The analysis has proved that bacteria inhabiting the intestines of the obese mice contain the enzymes that enable the degradation of polysaccharides which are the part of food. Stools of the obese subjects have contained a larger number of the end products such as acetic acid and butyric acid and were low in calories. The study led to the conclusion that the intestinal microflora in the obese can cause the acquisition of additional calories from ingested food. 16S rRNA gene analysis confirms the participation of the intestinal microflora in the pathogenesis of obesity. The number of bacteria belonging to the *Bacteroidetes* fell by 50% in obese with a proportional increase of the phylum *Firmicutes* bacteria. The pH level is a very important factor affecting the count and distribution of bacteria. Another difference observed in the microflora of the obese is an increased number of methanogenic bacteria that remove hydrogen from the environment and improve the fermentation process carried out by bacteria (Stachowicz 2013).

**Diabetes**

Type 2 Diabetes is a disease characterized by chronic hyperglycemia due to peripheral tissue insulin resistance or insufficient production of insulin by the β cells of the pancreas. The cooperation of genetic and environmental factors is influenced by the insulin-producing cell dysfunction and insulin resistance. In patients with diabetes a small number of *Faecalibacterium prausnitzii* has been detected. Study was carried out which showed the relationship between LPS in the blood plasma, the accumulation of triglycerides, weight gain and type 2 diabetes. LPS is a component of the cell wall of Gram-negative bacteria which constitutes the intestinal microflora and one that may also accompany the development of inflammation in the type 2 diabetes. This dependence may lead to the conclusion that LPS is involved in the development of diabetes. After the introduction of antibiotics (polymyxin B), which was directed mainly against Gram-negative bacteria, it mainly caused the reduction of LPS in blood plasma, reduction of the occurrence of the fatty liver and reduction frequency of type 2 diabetes (Furet et al. 2010).

Research on human microbiome contributes to the understanding of not only new microorganisms unidentified so far with the bacteriological methods, but most of all it will be the basis for the understanding of the relationship between the human organism and the microorganisms living in it. Thanks to the development...
of the metagenomics or the NGS it will be possible to discover new genes, metabolic pathways, and bidirectional links between human metabolism and bacteria, what will allow to find in the near future new therapeutic approaches in the treatment of many diseases previously considered only as civilization, or genetic diseases.

References:

17. Pasteur L. (1885), Observations relatives à la note précédente de M. Duclaux. CR Acad. Sci. 100, s.68.

Submitted: 18.01.2015
Accepted: 11.03.2015