De novo identification and taxonomic analysis of human blood peptides, not being a product of known human genes

M.Osetrova[1], G.Arapidi[1,2], O.Ivanova[2], I.Butenko[3], V.Govorun[1,2,3]

[1] Moscow Institute of Physics and Technology (State University), Institutskii Per. 9, Moscow Region , Dolgoprudny 141700, Russian Federation
[2] Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Miklukho-Maklaya 16/10, Moscow 117997, Russian Federation
[3] Federal Research and Clinical Center of Physical-Chemical Medicine of the Federal Medical & Biological Agency, Malaya Pirogovskaya 1a, Moscow 119435, Russian Federation

According to recent estimates, less than 50% of total cell count in human organism are actually human cells, while the rest of the cells present symbiotic microorganisms that inhabit our intestine, mucosa and skin. Such proportion allows us to speak about the so-called "superorganism": the aggregate of many different species, only one of which is human [1, 2]. Since blood is the connective tissue that is in contact with the majority of organs and tissues, it contains proteins, reflecting health status of the whole body, and, potentially, proteins and peptides of symbiotic organisms [3].

Majority of up-to-date techniques to analyze human metaproteome have a fundamental disadvantage: the identification of microbial peptides strongly depends on the choice of protein sequence databases. To overcome these limitations, we propose a new algorithm for interpretation of de novo peptide identification.

Experimental group consisted of 17 blood plasma samples from healthy donors: 10 of them belonged to healthy men (4 individual and 6 pooled samples from 10 donors) and 6 healthy women (4 individual and 3 pooled samples from 10 donors). Peptide extraction method included the following steps: fractionation on cation-exchange particles, desorption of peptides from the surface of major plasma proteins by heat treatment, ultrafiltration through a cartridge with a transmission threshold of 10 kDa and solid phase extraction on the reversed phase sorbent. Eluates were analyzed by a mass spectrometer Sciex TripleTOF 5600+. De novo identification of the obtained mass spectra was carried out using PEAKS 7.5 software [4].

All peptide identifications, filtered by average local confidence, were searched for exact match in the NCBInr database of all annotated genomes. For organisms, whose protein fragments were detected in the samples we then extracted taxonomic information.

Phylogenetic analysis and clustering of de novo identifications revealed peptides related to symbiotic microorganisms in human blood samples. The most represented in the blood peptides belonged to organisms represented in the human microbiota: Proteobacteria, Firmicutes and Actinobacteria phyla. De novo identification results were also verified by human microbiota database search.

Thus, de novo approach to peptide identification from mass spectrometry data revealed exogenous peptides in the composition of human blood plasma. Taxonomic analysis showed a significant correlation between the distribution of exogenous peptides and the distribution of human symbiotic microorganisms on phylum level.

This work was supported by the Russian Science Foundation (project No. 14-50-00131).

1. *Manimozhiyan Arumugam et al.* Enterotypes of the human gut microbiome // Nature 2011. № 473 p. 174-180

2. *Tyakht, A.V. et al.* Human gut microbiota community structures in urban and rural populations in Russia // Nat. Commun. 2013. №4, article number: 2469

3. *Schulz-Knappe P. et al.* The peptidomics concept // Comb. Chem. & High Throughput Screen. 2005. №8, p. 697-704.

4. *Ma B. et al.* PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry // Rapid Commun. Mass Spectrom. 2003. №17, p. 2337-2342.