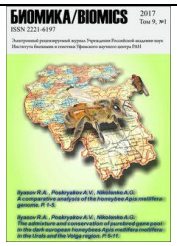
Биомика, 2017, Том 9, № 1, 1-5

БИОМИКА/BIOMICS

http://biomics.ru



A COMPARATIVE ANALYSIS OF THE HONEYBEE *APIS MELLIFERA* GENOME

R.A. Ilyasov, A.V. Poskryakov, A.G. Nikolenko

Institute of Biochemistry and Genetics, Ufa Scientific Centre, Russian Academy of Sciences, Ufa, 450054
E-mail: apismell@hotmail.com

RESUME

The analysis of the nuclear and the mitochondrial genomes of the honeybee *Apis mellifera* in comparison with the well-annotated, finished fruit fly *Drosophila melanogaster* genome presented in this article. The nuclear genome of the honeybee has about 245 millions bp, which distributed in 16 chromosomes and contains about 10 thousands genes. The mitochondrial genome of the *A. mellifera* has about 16 thousands bp, which located in mitochondrions and contains 35 genes. The nuclear genome of the fruit fly has about 144 millions bp, which distributed in 4 chromosomes and contains about 17 thousands genes. The mitochondrial genome of the *D. melanogaster* has about 19 thousands bp, which located in mitochondrions and contains 37 genes. Despite the full sequencing of the nuclear and the mitochondrial genomes of the *A. mellifera* the function of several genes and loci of *A. mellifera* do not disclosed fully. A comparative analysis of the genomes of *A. mellifera* and *D. melanogaster* using bioinformatics techniques allowed revealing the features of the structure and function of the honeybee *A. mellifera* genome. The genome of *A. mellifera* have more similarity with the vertebrate genome than *D. melanogaster*. The genome of *A. mellifera* contains less genes of the native immunity, of detoxification enzymes, of cuticle proteins and taste receptors compared with *D. melanogaster*. However, *A. mellifera* contains new genes associated with olfactory receptors, the processing of pollen and nectar, poison organs, wax glands, caste determination and labour division which absent at *D. melanogaster*. Probably, this is due to the ecology of bees and their social evolution.

Keywords: nuclear genome, mitochondrial genome, *Apis mellifera*, *Drosophila melanogaster*, fruit fly, fruitfly, honeybee, honey bee, genes, chromosomes.

There was only two sequenced genomes of the two Dipterans species *Drosophila melanogaster* [Adams et al., 2000] and *Anopheles gambiae* [Holt et al., 2002] before the recent sequencing of the honeybee *Apis mellifera* (Hymenoptera), *Tribolium castaneum* and *Bombyx mori* genomes (fig. 1). Hymenoptera diverged from Dipterans about 300 million years ago, and recent phylogenetic evidence implies that the *Apis* are the most distant group of holometabolous insects from *Drosophila* [Whiting, 2002; Krauss et al., 2005; Dearden et al., 2006].

Figure 1. **A.** A Dipteran species of the fruit fly *Drosophila melanogaster.*

B. A Hymenopteran species of the honeybee *Apis mellifera.*

The honeybee genome was compared with the well-annotated, finished *D. melanogaster* genome. The insect *D. melanogaster* genome is most studied of all the genomes. Despite honeybee *A. mellifera* is very economically important insect a few studies of full genome its have been published [Aronstein et al., 2016; Fried, Fried, 2016; Madras–Majewska et al., 2016; Pritchard, 2016; Skonieczna, 2016; Thunman, 2016]. Therefore, comparative analysis of two genomes the *A. mellifera* and *D. melanogaster* is very interesting [Ilyasov et al., 2015; Ilyasov, 2016].

Differences between *A. mellifera* and *D. melanogaster* caused by not only the nucleotide polymorphism of the genes but also by their different epigenetical regulation. Superficially, *A. mellifera* development is similar to that of *D. melanogaster*, in that it is a holometabolous. However, *A. mellifera* are different in their development and biology from the *D. melanogaster* in a number of ways. There is a hypothesis that all the differences that are observed between *A. mellifera* and *D. melanogaster*, have occurred since their divergence. This hypothesis is confirmed by the differences observed between *A. mellifera* and *D. melanogaster* in the early stages of development [Crozier, Crozier, 1993; Whitfield et al., 2006; Weinstock et al., 2007]. Thus, *A. mellifera* use haplodiploidy to determine sex, a process different from that of sex determination in *D. melanogaster*. The adult honeybees *A. mellifera* has several novel evolutionary innovations not present in *D. melanogaster*, including poison organs and wax glands. Most important are the caste determination and labour division associated with the social nature of the honeybee [Dearden et al., 2006; Ilyasov, 2016].

The nuclear genome of the honeybees *A. mellifera* has 246 927 000 bp which subdivided into 16 chromosomes and containing 10 157 genes (GeneBank access AADG00000000) [Whitfield et al., 2006; Weinstock et al., 2007]. The mitochondrial genome of the honeybees has 16 343 bp which represented by a circular molecule of DNA and containing 35 genes (GeneBank access NC_001566) [Crozier, Crozier, 1993]. All chromosomes of the honeybees has different sizes: LG 1 (NC_007070) 30 000 bp contains 1669 genes (25 non coding genes); LG 2 (NC_007071) 15500 bp - 814 genes (27 non coding genes); LG 3 (NC_007072) 13200 bp - 735 genes (20 non coding genes); LG 4 (NC_007073) 12700 bp - 709 genes (46 non coding genes); LG 5 (NC_007074) 14400 bp - 874 genes (13 non coding genes); LG 6 (NC_007075) 18500 bp - 844 genes (15 non coding genes); LG 7 (NC_007076) 13200 bp - 596 genes (9 non coding genes); LG 8 (NC_007077) 13500 bp - 873 genes (33 non coding genes); LG 9 (NC_007078) 11100 bp - 584 genes (17 non coding genes); LG 10 (NC_007079) 13000 bp - 768 genes (11 non coding genes); LG 11 (NC_007080) 14700 bp - 968 genes (16 non coding genes); LG 12 (NC_007081) 11900 bp - 504 genes (14 non coding genes); LG 13 (NC_007082) 10300 bp - 418 genes (13 non coding genes); LG 14 (NC_007083) 10300 bp - 612 genes (8 non coding genes); LG 15 (NC_007084) 10200 bp - 730 genes (30 non coding genes); LG 16 (NC_007085) 7200 bp - 420 genes (26 non coding genes) (GenBank - http://www.ncbi.nlm.nih.gov, EnsemblMetazoa - http://metazoa.ensembl.org).

For comparison, the nuclear genome of the fruit flies *D. melanogaster* has 143 726 000 which subdivided into four chromosomes (2 large, 1 small autosomes and the X/Y sex chromosomes) and containing 17 651 genes (3 384 non coding genes) (GeneBank access GCA_000001215.4) [Adams et al., 2000]. The mitochondrial genome of the fruit flies has 19 524 bp which represented by a circular molecule of DNA and containing 37 genes (GeneBank access NC_024511.2) [Dearden et al., 2006]. All chromosomes of the fruit flies has different sizes: 2L (NT_033779.5) 23510 bp contains 3485 genes (770 non coding genes); 3L (NT_037436.4) 28110 bp contains 3453 genes (666 non coding genes); 4 (NC_004353.4) 1350 bp contains 112 genes (26 non coding genes); X (NC_004354.4) 23 540 bp contains 2661 genes (408 non coding genes); Y (NC_024512.1) 3 670 bp contains 113 genes (28 non coding genes) (GenBank - http://www.ncbi.nlm.nih.gov, EnsemblMetazoa - http://metazoa.ensembl.org).

The nuclear and mitochondrial genome of *A. mellifera* differ from *D. melanogaster* by high containing of AT-rich regions. Since the *A. mellifera's* nuclear genome contains 67% and the mitochondrial genome - 85% AT whereas *D. melanogaster's* nuclear genome contains 58% and the mitochondrial genome - 79% AT nucleotides [Jukes, Bhushan, 1986; Ilyasov et al., 2015].

The nuclear and mitochondrial genome of *A. mellifera* characterized by greater spatial heterogeneity of AT-rich areas, higher content of CpG islands and absence of the the most common families of transposones than at *D. melanogaster*. The genes of *A. mellifera* predominantly located in AT-rich areas and characterized by high content of GC nucleotides. The A and T nucleotides of AT-rich areas in protein coding genes of *A. mellifera* are located in second and third positions of codons predominantly [Whitfield et al., 2006; Weinstock et al., 2007].

The structure and localization of most common genes in *A. mellifera* differ from *D. melanogaster*. In the *A. mellifera* mitochondrial genome 11 genes of tRNA have shift position as compared with *D. melanogaster*. The genetic code of *A. mellifera* similar to *D. melanogaster* but two anticodons of tRNA differ

($tRNA_{LYS}$ – TTT, $tRNA_{SER}$ - TCT in *A. mellifera* and $tRNA_{LYS}$ – CTT, $tRNA_{SER}$ - GCT in *D. melanogaster*) [Crozier, Crozier, 1993].

Some nuclear genes of the *A. mellifera* are orthologs to the *D. melanogaster* genes, which has differences in sizes [Weinstock et al., 2007; Wang et al., 2014]. Thus, in *A. mellifera* the genes of Yellow/Major Royal Jelly Protein are larger, the genes of cuticular proteins are smaller, the genes of odorant receptors are larger, the genes of gustatory receptors are smaller, the genes of immunity are smaller, the detoxification genes are smaller than in *D. melanogaster*.

In the *A. mellifera* genes the transversions occured more frequently than transition whereas in *D. melanogaster* it is conversely. In the *A. mellifera* genes transversions occurred on third position of codons. Some genes of *A. mellifera* arisen as a result of evolutionary changes of the genes of common with *D. melanogaster* ancestors. Thus, the gene of *A. mellifera* encoding the major protein of royal jelly is derived from the ancient gene yellow, which presented in *D. melanogaster*. Many genes of *A. mellifera* and *D. melanogaster* is similar, but some genes of *D. melanogaster* is absent in *A. mellifera*. For example, in *A. mellifera*, the genes of WNT cell signalling pathways as *HEDGEHOG* (*HH*), *TRANSFORMING GROWTH FACTOR-B* (*TGF-B*), *RECEPTOR TYROSINE KINASE* (*RTK*), *NOTCH*, *JANUS KINASE* (*JAK*), *SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION* (*STAT*) are similar with *D. melanogaster*. However, the genes of cell signalling systems (*TERMINAL EMBRYO FATE*, *TRUNK*, *TORSO)*, of component of the dorso-ventral signalling system (*GURKEN*), of the G-protein-coupled receptor (mGluR-like) family (*BOSS*) are missing from the *A. mellifera* genome [Weinstock et al., 2007].

Some genes of the *D. melanogaster* has novel features in the *A. mellifera*. Thus, the gene of the Glucose-methanol-choline oxidoreductases family (*NINAG*) in *A. mellifera* presents as two putative *NINAG*-like genes, the gene of the receptor protein tyrosine kinase family (*INR*) in *A. mellifera* is duplicated; the gene of the phospholipase C family (*NORPA*) in *A. mellifera* is duplicated; the gene of the photoreceptor-cell-specific nuclear receptor family (*PNR*) in *A. mellifera* presents as three genes versus two genes in *D. melanogaster*; the gene of the TRPA subfamily of transient receptor potential channels family (*TRPA1*) are missing in *D. melanogaster*, but has two extra TRPA channels (GB14005 and GB16385) in *A. mellifera;* the gene of the ligand-gated ion channels family (*NACR*) in *A. mellifera* presents as 11 subunits instead of 10 in *D. melanogaster;* the gene of the ligand-gated ion channels family (*NMDAR*) in *A. mellifera* presents as 3 genes instead of 2 in *D. melanogaster;* the gene of the excitatory amino acid transporters family (*EAAT*) in *A. mellifera* presents as 5 genes instead of 2 in *D. melanogaster* [Weinstock et al., 2007].

In *A. mellifera* 96 homeobox domains were found in 74 genes, similar to *D. melanogaster*. More than 90% identity represented by homeobox genes (*SEX COMBS REDUCED* (*SCR*), *ANTENNAPEDIA* (*ANTP*), *ABDOMINAL-A* (*ABD-A*); *ENGRAILED* (*EN*), *MUSCLE SEGMENT HOMEOBOX* (*MSH*)). For the remaining *A. mellifera* genes, a *D. melanogaster* homologue is not known. This indicates that structurally homologous genes are involved in the control of *A. mellifera* and *D. melanogaster* development [Walldorf et al., 1989; Weinstock et al., 2007].

The nuclear genes of *A. mellifera* which responsible for circadian rhythms (*CRY-M, CLK, CYC, PDP1, VRI, PER*), RNA interference (RNAi) and DNA methylation (381 genes in eggs and sperm of *A. mellifera* with CpG methylation) have more similarity with genes of vertebrate than with genes of *D. melanogaster* [Toma et al., 2000; Rubin et al., 2006; Elango et al., 2009; Drewell et al., 2014]. The circadian rhythms genes *TIMELESS* (*TIM1*) and *CRYPTOCHROME* (*DCRY*) of *D. melanogaster* are absent in *A. mellifera* genome. The similarity with vertebrate may be explained by the parallel evolution of the some genes during adaptation to the environment conditions. The genome of *A. mellifera* contains less genes of the native immunity, of detoxification enzymes, of cuticle proteins and taste receptors compared with *D. melanogaster*. However, *A. mellifera* contains new genes associated with olfactory receptors, the processing of pollen and nectar which absent at *D. melanogaster*. Probably, this is due to the ecology of bees and their social organization [Dearden et al., 2006; Wallberg et al., 2014].

The rate of the evolutionary transformations of the nuclear and mitochondrial genome of *A. mellifera* less than in *D. melanogaster*. However, the genome of *A. mellifera* diverged more distantly from common ancestor than *D. melanogaster* [Crozier et al., 1989; Crozier, Crozier, 1992]. Probably, this is due to the small effective population size of *A. mellifera* and to low rate of the reverse mutation compared with *D. melanogaster* [Crozier, 1980].

Micro RNA (miRNA) of the nuclear genome of *A. mellifera* plays an important role in the regulation of social organization and caste differentiation via post-transcriptional regulation of gene expression. About 300 honeybee miRNAs deposited in miRBase (http://www.mirbase.org) [Ashby et al., 2016]. For example, differentially expressed miRNAs between 4-day-old queen and worker larvae of honeybees: up-regulated in queen larvae (ame-bantam, ame-let-7, ame-mir-10, ame-mir-100, ame-mir-6001-3p); equally expressed in queen larvae (ame-mir-11, ame-mir-1175, ame-mir-190, ame-mir-6065, ame-mir-989); down-

regulated in queen larvae (ame-mir-13b, ame-mir-252a, ame-mir-2765-5p, ame-mir-996, ame-mir-9a) [Shi et al., 2015]. In the nuclear genome of *A. mellifera* found miRNA, which characterized by caste specific expression: the miRNA C5599F most expressed in the queens, C689F - in the pupaes, C5560 - in the pupaes of workers [Whitfield et al., 2006].

Thus, the genome of *A. mellifera* have more similarity with the vertebrate genome than *D. melanogaster*. The genome of *A. mellifera* contains less genes of the native immunity, of detoxification enzymes, of cuticle proteins and taste receptors compared with *D. melanogaster*. However, *A. mellifera* contains new genes associated with olfactory receptors, the processing of pollen and nectar, poison organs, wax glands, caste determination and labour division which absent at *D. melanogaster*. Probably, this is due to the ecology of bees and their social evolution. A comparative analysis of the genomes of *A. mellifera* and *D. melanogaster* using bioinformatics techniques allowed revealing the features of the structure and function of the honeybee *A. mellifera* genome.

ACKNOWLEDGMENTS

This research was supported by the Russian Foundation for Basic Research (grant No. 14-04-97084 r_povolzhye_a) and was carried out on equipment available at the Biomica Center for Collective Use of the Division of Biochemical Methods of Analysis and Nanobiotechnology of the Agidel Resource Center for the Collective Use.

REFERENCES

1. Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D. The genome sequence of *Drosophila melanogaster* // Science. 2000. V. 287. P. 2185-2195.
2. Ashby R., Foret S., Searle I., Maleszka R. MicroRNAs in honeybee caste determination // Scientific Reports. 2016. V. 6. P. 187-194.
3. Crozier R. H., Crozier Y. C., Mackinlay A. G. The CO-I and CO-II region of the honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates // Molecular Biology and Evolution. 1989. V. 6. P. 399-695.
4. Aronstein K., Colby D., Boykin D. P. Honey bee stock genotypes do not affect the level of physiological responses to chalkbrood fungus, Ascosphaera apis // Biomics. 2016. V. 8. P. 1 - 19.
5. Fried B., Fried P. Bee races and protected areas in Switzerland // Biomics. 2016. V. 8. P. 20 - 26.
6. Madras–Majewska B., Skonieczna L., Sokól R., Michalczyk M., Lisowska Z., Ochnio L. Health condition of bees inhabiting wild beehives and logs located in the forest districts of North–Eastern Poland // Biomics. 2016. V. 8. P. 48 - 53.
7. Pritchard D. The BIBBA–SICAMM anniversary conference, Llangollen (Wales) and SICAMM perspective // Biomics. 2016. V. 8. P. 54 - 60.
8. Skonieczna L. Conservation of Apis mellifera mellifera in Poland // Biomics. 2016. V. 8. P. 61 - 64.
9. Thunman P. Rescue project for the native bee, (Apis mellifera mellifera L.) in Sweden // Biomics. 2016. V. 8. P. 65 - 68.
10. Crozier R. H., Crozier Y. C. The cytochrome b and ATPase genes of honeybee mitochondrial DNA // Molecular Biology and Evolution. 1992. V. 9. P. 474-482.
11. Crozier R. H., Crozier Y. C. The mitochondrial genome of the honeybee Apis mellifera: complete sequence and genome organization // Genetics. 1993. V. 133. P. 97-117.
12. Crozier R. H. Genetical structure of social insect populations // Evolution of Social Behavior: Hypotheses and Empirical Tests. Weinheim: Verlag Chemie. 1980. P. 129-146.
13. Dearden P. K., Wilson M. J., Sablan L., Osborne P. W., Havler M. Patterns of conservation and change in honeybee developmental genes // Genome Research. 2006. V. 16. P. 1376-1384.
14. Drewell R. A., Bush E. C., Remnant E. J., Wong G. T., Beeler S. M., Stringham J. L., Lim J., Oldroyd BP The dynamic DNA methylation cycle from egg to sperm in the honeybee *Apis mellifera* // Development. 2014. V. 141. P. 2702-2711.
15. Elango N., Hunt B., Goodisman M., Yi S. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera* // Proceedings of the National Academy of Sciences of the United States of America. 2009. V. 106. P. 11206-11211.
16. Holt R. A., Subramanian G. M., Halpern A., Sutton G. G., Charlab R. The genome sequence of the malaria mosquito *Anopheles gambiae* // Science. 2002. V. 298. P. 129-149.
17. Ilyasov R. A., Poskryakov A. V., Nikolenko A. G. Nucleotide polymorphism of the gene VG of honeybees // Biomics. 2015. V. 7. P. 54-61.
18. Ilyasov R. A. Features of the honeybee *Apis mellifera* genome versus fruit fly *Drosophila melanogaster* // Journal of Investigative Genomics. 2016. V. 3. P. 00044 (1-3).
19. Ilyasov R. A., Kosarev M. N., Neal A., Yumaguzhin F. G. Burzyan wild-hive honeybee *A. m. mellifera* in South Ural // Biomics. 2014. V. 6. P. 196-201.
20. Jukes T. H., Bhushan V. Silent nucleotide substitutions and G + C content of some mitochondrial and bacterial genes // Journal of Molecular Evolution. 1986. V. 24. P. 39-44.

21. Krauss V., Pecyna M., Kurz K., Sass H. Phylogenetic mapping of intron positions: a case study of translation initiation factor eIF2γ // Molecular Biology Evolution. 2005. V. 22. P. 74-84.
22. Rubin E. B., Shemesh Y., Cohen M., Elgavish S., Robertson H. M. Molecular and phylogenetic analyses reveal mammalian-like clockwork in the honeybee (*Apis mellifera*) and shed new light on the molecular evolution of the circadian clock // Genome Research. 2006. V. 16. P. 1352-1365.
23. Shi Y. Y., Zheng H. J., Pan Q. Z., Wang Z. L., Zeng Z. J. Differentially expressed microRNAs between queen and worker larvae of the honeybee (*Apis mellifera*) // Apidologie. 2015. V. 46. P. 35-45.
24. Toma D. P., Bloch G., Moore D., Robinson G. E. Changes in period mRNA levels in the brain and division of labour in honeybee colonies // Proceedings of the National Academy of Sciences of the United States of America. 2000. V. 97. P. 6914-6919.
25. Wallberg A., Han F., Wellhagen G. A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera* // Nature Genetics. 2014. V. 46. P. 1081-1088.
26. Walldorf U., Fleig R., Gehring W. J. Comparison of homeobox-containing genes of the honeybee and Drosophila // Proceedings of the National Academy of Sciences of the United States of America. 1989. V. 86. P. 9971-9975.
27. Wang Y., Odemer R., Rosenkranz P., Moussian B. Putative orthologues of genetically identified *Drosophila melanogaster* chitin producing and organising genes in *Apis mellifera* // Apidologie. 2014. V. 45. P. 733-747.
28. Weinstock G. M., Robinson G. E., Gibbs R. A. Insights into social insects from the genome of the honeybee *Apis mellifera* // Nature. 2007. V. 443. P. 931-949.
29. Whitfield C. W., Behura S. K., Berlocher S. H., Clark A. G., Johnston J. S., Sheppard W. S., Smith D. R., Suarez A. V., Weaver D., Tsutsui N. D. Thrice out of Africa: Ancient and recent expansions of the honeybee, *Apis mellifera* // Science. 2006. V. 314. P. 642-645.
30. Whiting M. F. Phylogeny of the holometabolous insect orders: Molecular evidence // Zoologica Scripta. 2002. V. 31. P. 3-15.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЕНОМА МЕДОНОСНОЙ ПЧЕЛЫ *A. M. MELLIFERA* L.

Р.А. Ильясов, А.В. Поскряков, А.Г. Николенко

Институт биохимии и генетики Уфимского научного центра Российской академии наук
450054, г. Уфа, Пр. Октября, 71, E-mail: apismell@hotmail.com

АННОТАЦИЯ

В статье представлен сравнительный анализ ядерного и митохондриального геномов медоносной пчелы *Apis mellifera* и плодовой мушки *Drosophila melanogaster*. Ядерный геном медоносной пчелы имеет размер около 245 миллионов пн, который распределен в 16 хромосом и содержит около 10 тысяч генов. Митохондриальный геном *A. mellifera* имеет размер около 16 тысяч п. н., который расположен в митохондриях и содержит 35 генов. Ядерный геном плодовой мушки имеет размер около 144 миллионов пн, который дифференцирован в 4-х хромосомах и содержит около 17 тысяч генов. Митохондриальный геном *D. melanogaster* имеет размер около 19 тысяч п. н., который находится в митохондриях и содержит 37 генов. Несмотря на полное секвенирование ядерных и митохондриальных геномов *A. mellifera*, функции многих генов и локусов медоносной пчелы до сих пор не раскрыты полностью. Проведенный сравнительный анализ геномов *A. mellifera* и *D. melanogaster* с помощью методов биоинформатики позволил выявить отличительные особенности структуры и функции геномов медоносной пчелы. Геном *A. mellifera* имеет большее сходство с геномом позвоночных, чем с геномом дрозофилы. Геном *A. mellifera* содержит меньше генов естественного иммунитета, ферментов детоксикации, белков кутикулы и вкусовых рецепторов по сравнению с дрозофилой. Однако, *A. mellifera* содержит новые гены, связанные с обонятельными рецепторами, переработкой пыльцы и нектара, ядовитыми железами, восковыми железами, кастовой детерминацией и разделением труда, которые отсутствуют у дрозофилы. Вероятно, это связано с экологией пчел и их социальной эволюцией.

Ключевые слова: ядерный геном, митохондриальный геном, *Apis mellifera*, *Drosophila melanogaster*, фруктовая мушка, медоносная пчела, гены, хромосомы