

# КАРТИРОВАНИЕ ГЕНОМНЫХ СБОРОК АЗИАТСКОЙ ПЧЕЛЫ APIS CERANA

М.Д. Каскинова, Л.Р. Гайфуллина, Е.С. Салтыкова, У.Б. Юнусбаев, А.Г. Николенко

Институт биохимии и генетики – обособленное структурное подразделение Федерального государственного бюджетного научного учреждения Уфимского федерального исследовательского центра РАН, Уфа, Россия

В 2001 году Международный Консорциум по секвенированию опубликовал первую сборку генома человека. Расшифровка генома человека способствовала развитию технологий секвенирования и картирования, что позволило получить геномные сборки для других видов живых организмов, в том числе и для медоносной пчелы (Honeybee Genome Sequencing Consortium, 2006). Последняя сборка генома медоносной пчелы Amel\_HAv\_3.1 (Wallberg et al. 2019) была собрана на основе нескольких технологий секвенирования и картирования (Illumina, Pacbio, Bionano). Данная сборка представлена на хромосомном уровне и является наиболее полной версией генома медоносной пчелы. В Genebank представлены четыре сборки генома азиатской пчелы Apis cerana (ACSNU-2.0, ASM1110058v1, Apiscer\_1.0, ApisCC1.0), сестринского вида A. mellifera. Из них только одна собрана на хромосомном уровне (ASM1110058v1). Нами была поставлена задача перевести оставшиеся сборки на хромосомный уровень и выполнить аннотацию генов.

В качестве референсного генома нами был использован геном медоносной пчелы Amel\_HAv\_3.1 ([https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/003/254/395/GCF\\_003254395.2\\_Amel\\_HAv3.1/GCF\\_003254395.2\\_Amel\\_HAv3.1\\_genomic.fna.gz](https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/003/254/395/GCF_003254395.2_Amel_HAv3.1/GCF_003254395.2_Amel_HAv3.1_genomic.fna.gz) и [https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/003/254/395/GCF\\_003254395.2\\_Amel\\_HAv3.1/GCF\\_003254395.2\\_Amel\\_HAv3.1\\_genomic.gff.gz](https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/003/254/395/GCF_003254395.2_Amel_HAv3.1/GCF_003254395.2_Amel_HAv3.1_genomic.gff.gz)). В таблице 1 представлены основные характеристики исследуемых сборок. Две сборки A. cerana были собраны на основе длинных прочтений (long reads) PacBio и Oxford Nanopore и представляют наибольший интерес для исследователей.

Таблица 1 Характеристика исследуемых сборок

|                                  | Amel_HAv3.1                                                          | ACSNU-2.0         | Apiscer_1.0                                                               | ApisCC1.0           | ASM1110058v1       |
|----------------------------------|----------------------------------------------------------------------|-------------------|---------------------------------------------------------------------------|---------------------|--------------------|
| <b>Покрытие</b>                  | 192.0x                                                               | 152.0x            | 292x                                                                      | 122.0x              | 134.0x             |
| <b>Технология секвенирования</b> | PacBio; 10X Chromium; Phase Genomics HiC; Bionano                    | Illumina HiSeq    | Minion nanopore; illumina Hiseq2500                                       | Illumina GAIIx; 454 | PacBio Sequel      |
| <b>Метод сборки</b>              | FALCON v. 0.5.0; Arcs v. 1.0.1; Links v. 1.8.5; BioNano Solve v. 3.1 | AllPaths v. 42411 | SPAdе v. 3.10.1; Finishing Module in CLC genomics workbench Genome v. 1.7 | Velvet v. 1.2.03    | HGAP v. March-2019 |

Сборки A. cerana были переведены на хромосомный уровень с помощью RaGOO (Alonge et al. 2019). Для аннотации генов использовали Liftoff (Shumate and Salzberg 2020), используя gff файл Amel\_HAv\_3.1 как референс. В таблице 2 представлены данные по переносу исследуемых сборок генома на хромосомный уровень. В целом, размеры хромосом разных сборок совпадают.

В сборке Amel\_HAv\_3.1 аннотировано 12332 генов. Мы выполнили перенос этих генов на сборки A. cerana с помощью Liftoff (рис. 1). Самое большое количество генов картировалось на сборку ASM1110058v1, собранную из длинных прочтений PacBio. Park et al. (2015) аннотировали 10651 генов для своей сборки ACSNU-2.0. Тогда как с использованием Amel\_HAv\_3.1 нами было аннотировано 12089 генов.

Таблица 2 Размеры хромосом в сборках генома Amel\_HAv3.1 и Apis cerana

|       | Amel_HAv3.1 | ACSNU-2.0 | Apiscer_1.0 | ApisCC1.0 | ASM1110058v1 |
|-------|-------------|-----------|-------------|-----------|--------------|
| Chr1  | 27754200    | 30493594  | 27245775    | 29586723  | 26802524     |
| Chr2  | 16089512    | 16993187  | 16631585    | 16221043  | 16892262     |
| Chr3  | 13619445    | 13698954  | 14122599    | 14146147  | 15590573     |
| Chr4  | 13404451    | 13125209  | 12022470    | 15077644  | 15483476     |
| Chr5  | 13896941    | 14424151  | 15114023    | 14929130  | 13581003     |
| Chr6  | 17789102    | 17345009  | 16738848    | 18123026  | 13422283     |
| Chr7  | 14198698    | 13958837  | 12181940    | 13796122  | 13101753     |
| Chr8  | 12717210    | 13288447  | 12123437    | 12418764  | 13171213     |
| Chr9  | 12354651    | 14328994  | 10553639    | 12938676  | 11764534     |
| Chr10 | 12360052    | 11153287  | 11707081    | 11820133  | 11894629     |
| Chr11 | 16352600    | 16487584  | 14253970    | 16878668  | 11548277     |
| Chr12 | 11514234    | 11749249  | 10816915    | 10161981  | 10523944     |
| Chr13 | 11279722    | 11427994  | 11150747    | 10986969  | 10566570     |
| Chr14 | 10670842    | 11361364  | 9899063     | 12622570  | 10402947     |
| Chr15 | 9534514     | 10259240  | 9120600     | 10226317  | 9174731      |
| Chr16 | 7238532     | 7736732   | 6518383     | 7009582   | 7099464      |

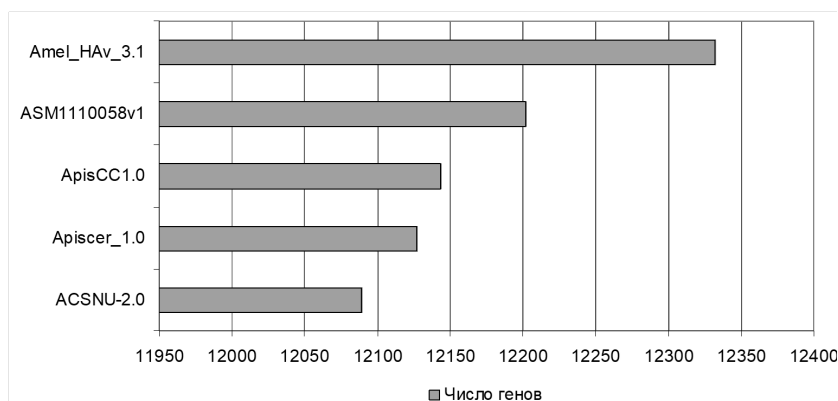


Рисунок 1 Аннотация генов в сборках Apis cerana с помощью Liftoff (Amel\_HAv3.1 в качестве референса)

Таким образом, не смотря на то, что эволюционное разделение азиатской *A. cerana* и европейской *A. mellifera* произошло 6–25 млн лет назад (Tihelka et al., 2020), у них выявлено большое количество общих генов. Полученные данные можно использовать для сравнительного анализа генов двух видов *Apis* и для оценки внутривидового полиморфизма *A. cerana*.

**Работа выполнена при финансовой поддержке гранта РФФИ № 19–54–70002.**

### Литература

- Alonge M., Soyk S., Ramakrishnan S., Wang X., Goodwin S. et al. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol.* 2019. 20: 224.
- Shumate A., Salzberg S.L. Liftoff: accurate mapping of gene annotations. *Bioinformatics.* 2020. 15: btaa1016.
- Wallberg A., Bunikis I., Pettersson O.V., Mosbech M. – B., Childers A.K. et al., A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. *BMC Genomics.* 2019. 20: 275.
- Honeybee Genome Sequencing Consortium. Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature.* 2006. 443: 931–949.
- Tihelka E., Cai C., Pisani D. et al. Mitochondrial genomes illuminate the evolutionary history of the Western honey bee (*Apis mellifera*). *Sci Rep.* 2020. 10: 14515.
- Park D., Jung J.W., Choi B.S. et al. Uncovering the novel characteristics of Asian honey bee, *Apis cerana*, by whole genome sequencing. *BMC Genomics.* 2015. 16: 1.