东方蜜蜂微孢子虫的 SNP 与 InDel 位点鉴定及分析*

张文德[1**] 蔡宗兵[1**] 隆 琦[1] 吴 鹰[1] 孙明会[1] 康育欣[1]
胡 颖[1] 赵 萧[1] 陈大福[1, 2***] 郭 睿[1, 2***]

（1. 福建农林大学动物科学学院（蜂学学院），福州 350002；2. 福建农林大学蜂疗研究所，福州 350002）

摘 要 【目的】 东方蜜蜂微孢子虫 *Nosema ceranae* 是一种专性侵染成年蜜蜂中肠上皮细胞的单细胞真菌病原，广泛感染世界各地的蜂群。本研究拟利用已获得的东方蜜蜂微孢子虫纯净孢子的高质量转录组数据进行单核苷酸多态性（Single nucleotide polymorphism，SNP）和插入缺失（Insertion-Deletion，InDel）位点的鉴定和分析，旨在丰富东方蜜蜂微孢子虫的 SNP 和 InDel 信息，并为新型分子标记的开发提供基础。**【方法】** 使用 GATK 软件识别东方蜜蜂微孢子虫的 SNP 和 InDel 位点。采用 SnpEff 软件预测变异位点发生的基因组区域及变异产生的影响。通过相关生物信息学软件分别将 SNP 和 InDel 位点所在基因分别比对 GO 和 KEGG 数据库，获得相应的功能和通路注释。**【结果】** 共鉴定到 28 195 个 SNP 位点，其中发生转换和颠换的 SNP 位点分别有 21 403 和 6 792 个；上述 SNP 位点的突变类型有 12 种，其中最丰富的突变类型为 C/T；分布在 CDS 区的 SNP 位点最多，其次是基因间区、上游区、下游区和内含子区；最丰富的密码子突变类型是同义突变；SNP 位点所在基因可注释到代谢进程、细胞组分和催化活性等 43 个 GO 条目以及代谢途径、核糖体和等次生代谢产物的生物合成等 85 条 KEGG 通路。共鉴定到 2 831 个 InDel 位点，其中分布在基因间区 InDel 位点最多，分布在 CDS 区的 InDel 位点最少；最丰富的密码子突变类型是移码突变；InDel 位点所在基因可注释到细胞进程、细胞和结合等 38 个 GO 条目以及代谢途径、次生代谢产物的生物合成及核糖体等 73 条 KEGG 通路。**【结论】** 东方蜜蜂微孢子虫中存在大量的 SNP 和 InDel 位点，SNP 位点的突变类型主要为转换，与其他物种类似；SNP 与 InDel 位点的基因组功能元件分布规律和突变类型具有明显差异；SNP 和 InDel 位点所在基因与东方蜜蜂微孢子虫适应宿主细胞内环境及病原增殖过程具有潜在关系。

关键词 东方蜜蜂微孢子虫；单核苷酸多态性；插入缺失；转录组；分子标记

Identification and analysis of SNP and InDel loci in *Nosema ceranae*

ZHANG Wen-De[1**] CAI Zong-Bing[1**] LONG Qi[1] WU Ying[1] SUN Ming-Hui[1]
KANG Yu-Xin[1] HU Ying[1] ZHAO Xiao[1] CHEN Da-Fu[1, 2***] GUO Rui[1, 2***]

(1. College of Animal Sciences (College of Bee Science), Fujian Agriculture and Forestry University, Fuzhou 350002, China;
2. Apitherapy Research Institute, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract [Objectives] *Nosema ceranae* is a widespread, single-cell, fungal pathogen that exclusively infects honeybee midgut epithelial cells worldwide. The objective of this study is to identify and analyze single nucleotide polymorphism (SNP) and insertion and deletion (InDel) loci in *N. ceranae* using high-quality transcriptome data obtained from clean *N. ceranae* spores, with the aim of developing novel molecular markers. **[Methods]** SNP and InDel loci were detected using GATK

*资助项目 Supported projects：国家自然科学基金面上项目（32172792）；国家现代农业产业技术体系建设专项资金项目（CARS-44-KXJ7）；福建农林大学杰出青年科研人才项目（xjq201814）；福建农林大学硕士生导师团队项目（郭睿）；福建省病原真菌与真菌毒素重点实验室开放课题（郭睿）

**共同第一作者 Co-first authors，E-mail：wdebee@163.com；caizongbing@126.com

***共同通讯作者 Co-corresponding authors，E-mail：dfchen826@fafu.edu.cn；ruiguo@fafu.edu.cn

收稿日期 Received：2021-07-07；接受日期 Accepted：2021-11-08

software. SnpEff software was used to predict genomic regions with mutation sites and the effects caused by mutation. Genes containing SNP or InDel loci were respectively aligned to the GO and KEGG databases to annotate them to corresponding function and pathways. **[Results]** A total of 28 195 SNP loci were identified in *N. ceranae*, 21 403 of which were conversion loci and 6 792 of which were transversion loci. These SNP loci had 12 types of mutation, the most abundant of which was C/T. SNP loci were mainly distributed in the CDS region, followed by the intergenic region, upstream region, downstream region and the intron. In addition, the most common type of mutation codon in SNP loci was synonymous mutation. Genes containing SNP loci were annotated to 43 GO terms, including metabolic process, cellular process and catalytic activity, and 85 KEGG pathways, such as metabolic pathway, ribosome and biosynthesis of secondary metabolites. 2 831 InDel loci were identified, most of which were distributed in the intergenic region with fewest found in the CDS region. In addition, the most abundant type of codon mutation was the frameshift mutation. Genes containing Indel loci were annotated to 38 GO terms, including cellular process, cell and binding, and 73 KEGG pathways including metabolic pathways, the biosynthesis of secondary metabolites and ribosomes. **[Conclusion]** There are high numbers of SNP and InDel loci in *N. ceranae* and, similar to other species, the most common SNP mutation is conversion. The genomic distribution of the functional elements and mutation types of SNP loci are obviously different from those of InDel loci. Genes containing SNP and InDel loci are potentially involved in the adaptation of *N. ceranae* to the intracellular environment and the proliferation process of this pathogen.

Key words *Nosema ceranae*; single nucleotide polymorphism (SNP); insertion and deletion (InDel); transcriptome; molecular marker

东方蜜蜂微孢子虫 *Nosema ceranae* 是一种专性侵染成年蜜蜂中肠上皮细胞的单细胞真菌病原，对蜂王、工蜂、雄蜂和蜜蜂幼虫均具有感染性（Traver and Fell，2012；Eiri *et al.*，2015）。被感染的蜜蜂寿命缩短，哺育力下降，蜂群繁殖力和生产力均受到较大影响（Martín-Hernández *et al.*，2018）。目前，东方蜜蜂微孢子虫的参考基因组（Cornman *et al.*，2009）和全长转录组（陈华枝等，2020）均已公布，为其分子生物学和组学研究的深入开展奠定了坚实基础。

单核苷酸多态性（Single nucleotide polymorphism，SNP）是指在基因组水平上由单个核苷酸的变异所引起的 DNA 序列多态性，它是物种可遗传变异中最常见的一种，也是生物进化的重要途径（赵辉等，2006）。插入缺失（Insertion-Deletion，InDel）突变作为另一种新型分子标记，指基因组中小片段序列的插入或缺失（Wajnberg and Passetti，2016），相对于拷贝数变异（Copy number variations，CNV），InDel 可直接对基因多态性进行检测（Cui *et al.*，2018），同时具有高准确性、高稳定性和结果清晰易判等优点（杨洁等，2016）。另外，InDel 因在种内和种间都具有多态性而具有较强的通用性，可广泛用于图位克隆、基因定位、遗传图谱构建、群体遗传分析及分子辅助遗传育种等方面（Bhattramakki *et al.*，2002；Väli *et al.*，2008；Lu *et al.*，2009）。

近十多年来，二代测序技术获得了长足发展，已成为生命科学研究中不可或缺的常用工具。因具有高通量、高灵敏度、短周期和低成本等优势，二代测序技术已成功应用于分子标记的鉴定和开发（Blythe *et al.*，2012；Danielson *et al.*，2017）。目前，基于二代测序技术的 SNP 和 InDel 研究已在鸡 *Gallus domesticus*、玉米 *Zea mays* 及绒山羊 *Capra hircus* 等物种中见诸报道（赵森等，2020；刘小红，2021；张洁慧等，2021）。然而，东方蜜蜂微孢子虫的 SNP 和 InDel 研究至今仍然缺失。

此前，笔者所在团队利用 Illumina 测序技术对东方蜜蜂微孢子虫的纯净孢子进行了转录组测序，并对 lncRNA 和 circRNA 进行了系统鉴定和分析（Guo *et al.*，2018a，2018b）；还对参考基因组的注释基因进行了结构优化，并鉴定和分析了未注释的新基因（熊翠玲等，2019）。本研究拟利用已获得的高质量转录组数据对东方蜜蜂微孢子虫的 SNP 和 InDel 位点进行鉴定和分析，以期为其基因图位克隆、遗传多样性和遗传图谱等研究提供候选分子标记，并为其他真菌物

种和非模式物种提供思路和方法借鉴。

1 材料与方法

1.1 供试生物材料

本研究使用的东方蜜蜂微孢子虫孢子由福建农林大学蜂学学院蜜蜂保护实验室制备和保存（Guo *et al.*，2018a）。

1.2 东方蜜蜂微孢子虫的深度测序数据来源

笔者所在团队前期在实验室条件下进行了东方蜜蜂微孢子虫孢子的纯化，并利用 Illumina Hiseq Xten 平台对纯净的病原孢子进行双端测序，3 个孢子样品共测得 65 632 198 952 条原始读段（raw reads），经严格质量得到 218 462 218 条有效读段（clean reads），平均 Q30 达到 90.22%（Guo *et al.*，2018a）。测序原始数据已上传 NCBI SRA 数据库，Bioproject 号：PRJNA395264。高质量的转录组数据可用于本研究中 SNP 和 InDel 位点的鉴定与分析。

1.3 SNP 和 InDel 位点的鉴定与分析

参照吴鹰等（2022）的方法，使用 GATK 软件（McKenna *et al.*，2010）识别转录组数据与东方蜜蜂微孢子虫参考基因组（assembly ASM98816v1）之间的碱基错配，进而鉴定潜在的 SNP 和 InDel 位点。GATK 识别标准如下：（1）35 bp 范围内连续出现的单碱基错配不超过 3 个；（2）经过序列深度标准化的质量值大于 2.0。各样品分别按照以上条件筛选，最终获得可靠的 SNP/InDel 位点。SnpEff（Cingolani *et al.*，2012）是一款用于注释变异（SNP、InDel）和预测变异影响的软件。根据变异位点在参考基因组上的位置及参考基因组上的基因位置信息预测变异位点在基因组发生的区域及变异产生的影响（同义突变或非同义突变等）。依据碱基替换的方式可将 SNP 位点的突变类型分为转换（Transition）和颠换（Transversion）。

1.4 SNP 和 InDel 位点所在基因的数据库注释

采用基迪奥生物云平台（https://www.omicshare.com/tools/home/index/index）的相关生物信息学工具将东方蜜蜂微孢子虫的 SNP 位点所在基因和 InDel 位点所在基因比对 GO（Gene Ontology）数据库（https://www.omicshare.com/tools/Home/Soft/gogsea）和 KEGG（Kyoto Encyclopedia of Genes and Genomes）数据库（https://www.omicshare.com/tools/Home/Soft/pathwaygsea）比对，以获得相应的功能条目（Term）和通路（Pathway）注释。

2 结果与分析

2.1 东方蜜蜂微孢子虫 SNP 位点的鉴定与分析

共鉴定到东方蜜蜂微孢子虫的 28 195 个 SNP 位点。部分 SNP 位点的详细信息如表 1 所

表 1 东方蜜蜂微孢子虫的部分 SNP 的详细信息

Table 1 Detailed information of partial SNPs in *Nosema ceranae*

染色体 ID Chromosome ID	位置 Position	参考碱基 Reference base	突变碱基 Mutant base	位置类型 Structure type	基因 ID Gene ID
NW_003314216.1	12 908	C	T	CDS	Gene1973
NW_003314154.1	3 572	A	G	CDS	Gene1485
NW_003308797.1	128	G	C	Intergenic	–
NW_003308797.1	223	A	G	Intergenic	–
NW_003314154.1	3 685	A	G	Upstream	Gene1485
NW_003314145.1	2 933	C	T	Upstream	Gene1431
NW_003311411.1	528	T	C	Downstream	Gene105
NW_003314154.1	3 882	A	G	Downstream	Gene1486
NW_003314145.1	6 785	A	T	Intron	Gene1422

示。统计结果显示，分别有 21 403 和 6 792 个 SNP 发生转换和颠换。东方蜜蜂微孢子虫的 SNP 位点分为 12 个突变类型，包括 C/T、G/A、A/G、T/C、G/T、A/T、C/A、T/A、T/G、A/C、C/G 和 G/C；其中最丰富的突变类型为 C/T（5 422），最少的突变类型为 C/G（302）（图 1：A）。SNP 在参考基因组功能元件的分布统计结果显示，分布在 CDS 区（Coding sequence region）的 SNP 位点最多（10 865），然后是基因间区（10 508）、上游区（3 465）、下游区（3 356）和内含子区（1）（图 1：B）。进一步分析分布在 CDS 区的 SNP 位点的密码子突变类型，最丰富的密码子突变类型为同义突变（6 448），然后是错义突变（4 357）、终止子增加（38）、终止子减少（11）、启动子减少（9）及同义终止子转换（2）（图 1：C）。

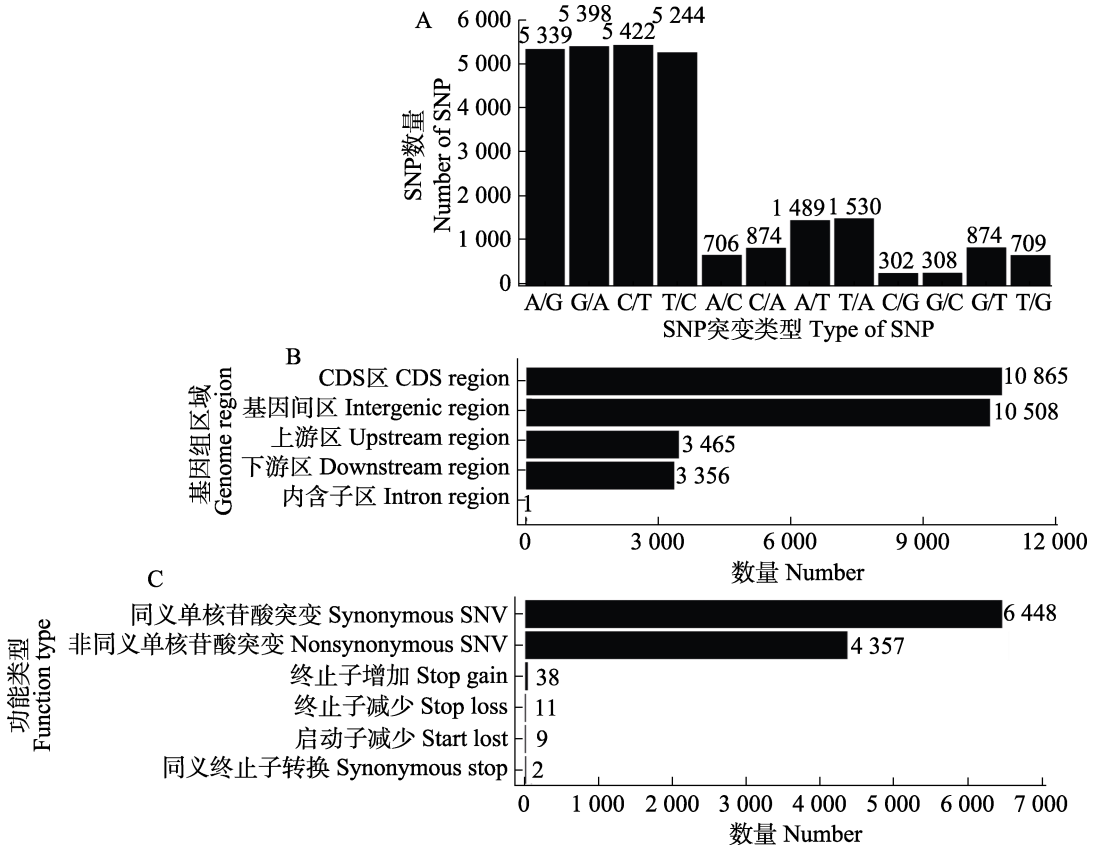

图 1　SNP 的突变类型（A）、基因组位置分布（B）及密码子突变类型（C）

Fig. 1　Mutation type of SNPs (A), location distribution in genome of SNPs (B) and mutation type of codons (C)

2.2　东方蜜蜂微孢子虫的 SNP 位点所在基因的数据库注释

GO 数据库注释结果显示，东方蜜蜂微孢子虫的 SNP 位点所在基因涉及 43 个功能条目，包括细胞进程（538）和代谢进程（514）等 22 个生物学进程相关条目，细胞（365）和细胞组分（365）等 11 个细胞组分相关条目，结合（512）和催化活性（467）等 10 个分子功能相关条目（图 2：A）。KEGG 数据库注释结果显示，SNP 位点所在基因涉及 85 条通路，注释基因最多的通路为代谢途径（92），其次是核糖体（56）、次生代谢产物的生物合成（45）及真核生物的核糖体生物发生（37）等（图 2：B）。

2.3　东方蜜蜂微孢子虫的 InDel 位点的鉴定与分析

共鉴定到东方蜜蜂微孢子虫的 2 831 个 InDel 位点。部分 InDel 位点详细信息如表 2 所示。

InDel 位点分布最丰富的功能元件为基因间

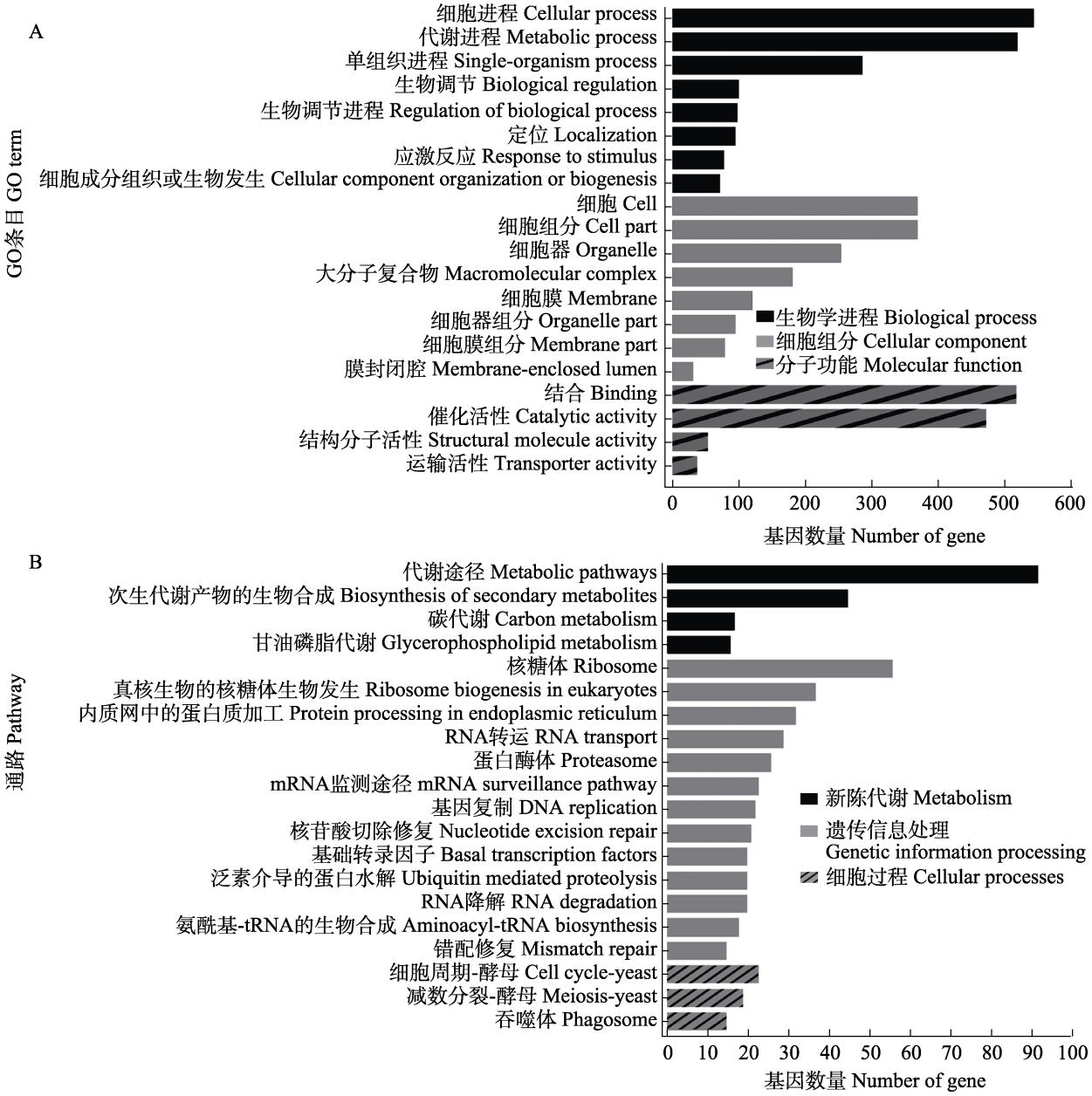

图 2　东方蜜蜂微孢子虫 SNP 所在基因的 GO 和 KEGG 数据库注释

Fig. 2　GO and KEGG database annotation of genes with *Nosema ceranae* SNPs

A. GO 数据库注释的前 20 位条目；B. KEGG 数据库注释的前 20 条通路。

A. Top 20 terms annotated in GO database; B. Top 20 pathways annotated in KEGG database annotation.

区，总数为 1 501 个，其次为上游区、下游区和 CDS 区，数量分别为 593 个、493 个和 244 个（图 3：A）。进一步对分布在 CDS 区的 InDel 位点的密码子突变类型进行分析，最丰富的密码子突变类型是移码突变（227），其次为密码子缺失（8）、密码子插入（4）、终止子增加（2）、启动子减少（2）和终止子减少（1）（图 3：B）。

2.4　东方蜜蜂微孢子虫的 InDel 位点所在基因的数据库注释

上述 InDel 位点所在基因可注释到 38 个功能条目，包括代谢进程（212）和细胞进程（212）等 20 条生物学进程相关条目，细胞（136）和细胞进程（136）等 10 条细胞组分相关条目，结合

表 2　东方蜜蜂微孢子虫 InDel 位点的详细信息

Table 2　Detailed information of partial InDels in *Nosema ceranae*

染色体 ID Chromosome ID	位置 Position	基因 ID Gene ID	参考碱基 Reference base	突变碱基 Mutant base	位置类型 Structure type
NW_003309261.1	200	–	A	ATT	Intergenic
NW_003308803.1	14	–	TA	T	Intergenic
NW_003312054.1	66	Gene154	T	TA	Upstream
NW_003312054.1	66	Gene154	T	TA	Upstream
NW_003311908.1	192	Gene143	TA	T	Downstream
NW_003313537.1	1 286	Gene434	AG	A	Downstream
NW_003313774.1	2 393	Gene603	ATAT	A	CDS
NW_003314116.1	1 679	Gene1272	A	AT	CDS

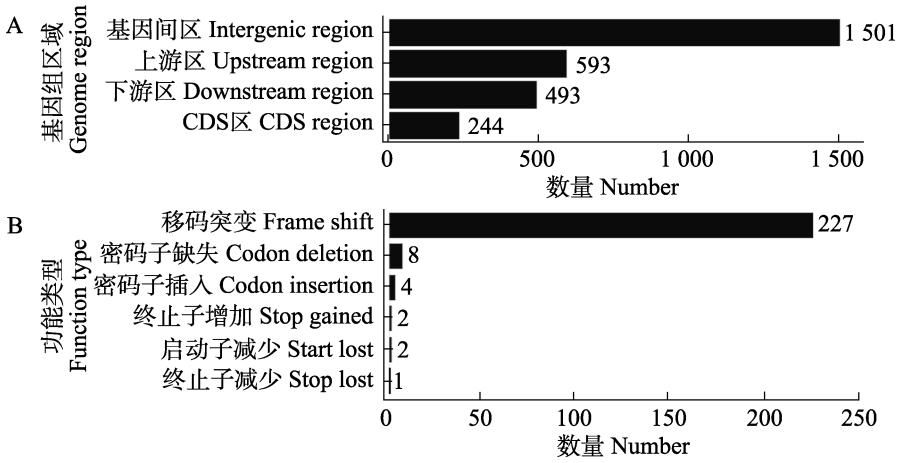

图 3　InDel 的基因组位置分布（A）和密码子突变类型（B）

Fig. 3　Location distribution in genome of InDels (A) and mutation type of codons (B)

（205）和催化活性（204）等 8 条分子功能相关条目（图 4：A）；此外还可以注释到代谢途径（45）、次生代谢产物的生物合成（23）及核糖体（19）等 73 条通路（图 4：B）。

3　讨论

基于已获得的高质量转录组数据，本研究首次对东方蜜蜂微孢子虫的 SNP 和 InDel 位点进行鉴定与分析，共鉴定到 28 195 个 SNP 位点，其中高达 75.91%的 SNP 位点突变类型为转换，这与前人在鸡（张洁慧等，2021）和向日葵锈菌 *Puccinia helianthi*（王妍等，2018）等物种上的研究结果类似。微孢子虫是基因组结构最简化的一类真核生物，其生存和繁殖高度依赖于宿主细胞内环境，因此相比于基因间区，与编码蛋白直接相关的 CDS 区会受到更大的环境选择压力（Williams *et al.*，2016）。本研究中，分布 SNP 位点最多的功能元件为 CDS 区，与上述结论一致。进一步对分布在 CDS 区的 SNP 位点进行突变类型分析，发现高达 59.35%的 SNP 位点为同义突变，即氨基酸未发生改变的中性突变，与玉米（刘小红，2021）和山核桃（董昂，2018）等物种中的研究结构一致。这表明东方蜜蜂微孢子虫在侵染过程中虽受到宿主细胞内的较强环境选择压力，但能通过某些方式维持编码蛋白的稳定性。

本研究共鉴定到东方蜜蜂微孢子虫的 2 831 个 InDel 位点，与 SNP 位点的分布特点不同，多

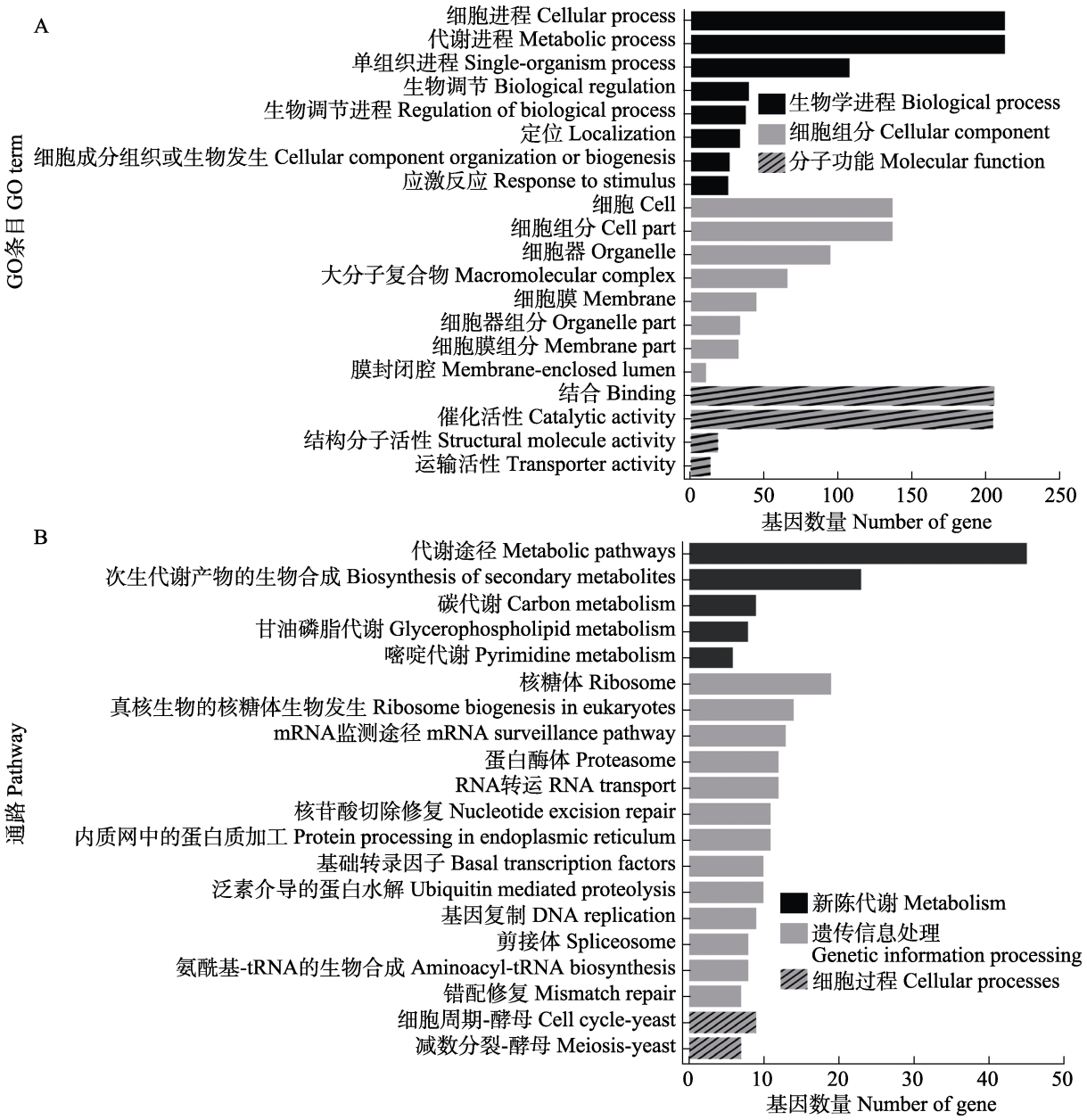

图 4　东方蜜蜂微孢子虫 InDel 所在基因的 GO 和 KEGG 数据库注释

Fig. 4　GO and KEGG database annotation of genes with InDel in *Nosema ceranae*

A. GO 数据库注释的前 20 位条目；B. KEGG 数据库注释的前 20 条通路。

A. Top 20 terms annotated in GO database; B. Top 20 pathways annotated in KEGG database annotation.

数（53.02%）InDel 位点分布在基因间区，而分布在 CDS 区的 InDel 位点最少（8.62%），与其他物种的研究报道相似（Bhangale *et al.*，2005；Brandstrom and Ellegren ，2007）。究其原因，相比于 SNP，InDel 如发生在 CDS 区易造成移码突变，进而会导致连续转录错误，转录提前终止或编码蛋白改变，一个物种为保证基因功能及编码蛋白的稳定性，InDel 突变较少发生在 CDS 区，这是长期自然选择的结果。前人在研究玉米时发现分布在 CDS 区的 InDel 位点碱基数多为 3 的倍数，即以密码子为单位进行突变（Liu *et al.*，2015）。有趣的是，本研究发现，对于东方蜜蜂微孢子虫的 InDel，发生在 CDS 区的移码突变占比高达 93.03%，而以密码子为单位进行的突变

仅占 4.92%。东方蜜蜂微孢子虫的基因组仅含有 2 614 个基因,基因密度为 0.6/kb(Cornman *et al.*, 2009)。推测东方蜜蜂微孢子虫的 InDel 中移码突变占比很高是由于基因组高度简化且基因间区大幅缩短所致。

本研究发现，东方蜜蜂微孢子虫的 SNP 位点所在基因涉及代谢进程、细胞组分和催化活性等 43 个功能条目，以及代谢途径、核糖体和次生代谢产物的生物合成等 85 条通路; InDel 位点所在基因可注释到生长发育、新陈代谢和环境适应等相关的 38 个 GO 条目和 73 条 KEGG 通路。以上结果表明 SNP 和 InDel 突变与东方蜜蜂微孢子虫适应宿主细胞内环境及病原增殖过程具有潜在的重要关系。

参考文献 (References)

Bhangale TR, Rieder MJ, Livingston RJ, Nickerson DA, 2005. Comprehensive identification and characterization of dialle-lic insertion-deletion polymorphisms in 330 human candi-date genes. *Hum. Mol. Genet.*, 14(1): 59–69.

Bhattramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register JC 3rd, Tingey SV, Rafalski A, 2002. Insertion-deletion polymorphisms in 3′ regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Mol. Biol.*, 48(5/6): 539–547.

Brandström M, Ellegren H, 2007. The genomic landscape of short insertion and deletion polymorphisms in the chicken (*Gallus gallus*) Genome: A high frequency of deletions in tandem duplicates. *Genetics*, 176(3): 1691–1701.

Blythe MJ, Malla S, Everall R, Shih YH, Lemay V, Moreton J, Wilson R, Aboobaker AA, 2012. High through-put sequencing of the *Parhyale hawaiensis* mRNAs and microRNAs to aid comparative developmental studies. *PLoS ONE*, 7(3): e33784.

Chen HZ, DU Y, Fan XX, Zhu ZW, Jiang HB, Wang J, Fan YC, Xiong CL, Zheng YZ, Fu ZM, Xu GJ, Chen DF, Guo R, 2020. Construction and annotation of the full-length transcriptome of *Nosema ceranae* based on the third-generation nanopore sequencing technology. *Acta Entomologica Sinica*, 63(12): 1461–1472. [陈华枝, 杜宇, 范小雪, 祝智威, 蒋海宾, 王杰, 范元婵, 熊翠玲, 郑燕珍, 付中民, 徐国钧, 陈大福, 郭睿, 2020. 基于第三代纳米孔测序技术的东方蜜蜂微孢子虫全长转录组构建及注释. 昆虫学报, 63(12): 1461–1472.]

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM, 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 6(2): 80–92.

Cornman RS, Chen YP, Schatz MC, Street C, Zhao Y, Desany B, Egholm M, Hutchison S, Pettis JS, Lipkin WI, Evans JD, 2009. Genomic analyses of the microsporidian *Nosema ceranae*, an emergent pathogen of honey bees. *PLoS Pathog.*, 5(6): e1000466.

Cui Y, Yan H, Wang K, Xu H, Zhang X, Zhu H, Liu J, Qu L, Lan X, Pan C, 2018. Insertion/Deletion within the *KDM6A* gene is significantly associated with litter size in goat. *Front. Genet.*, 9: 91.

Danielson KM, Rubio R, Abderazzaq F, Das S, Wang YE, 2017. High throughput sequencing of extracellular RNA from human plasma. *PLoS ONE*, 12(1): e0164644.

Dong A, 2018. Construction of genetic linkage maps and analysis of DNA methytiaon-related gene expression in hickory. Master dissertation. Hangzhou: Zhejiang A&F University. [董昂, 2018. 山核桃遗传图谱构建及甲基化相关基因表达分析. 硕士学位论文. 杭州: 浙江农林大学.]

Eiri DM, Suwannapong G, Endler M, Nieh JC, 2015. *Nosema ceranae* can infect honey bee larvae and reduces subsequent adult longevity. *PLoS ONE*, 10(5): e0126330.

Guo R, Chen DF, Xiong CL, Hou CS, Zheng YZ, Fu ZM, Liang Q, Diao QY, Zhang L, Wang HQ, Hou ZX, Kumar D, 2018a. First identification of long non-coding RNAs in fungal parasite *Nosema ceranae*. *Apidologie*, 49(5): 660–670.

Guo R, Chen DF, Chen HZ, Xiong CL, Zheng YZ, Hou CS, Du Y, Geng SH, Wang HP, Zhou DD, Guo YL, 2018b. Genome-wide identification of circular RNAs in fungal parasite *Nosema ceranae*. *Curr. Microbiol.*, 75(12): 1655–1660.

Liu J, Qu J, Yang C, Tang D, Li J, Lan H, Rong T, 2015. Development of genome-wide insertion and deletion markers for maize, based on next-generation sequencing data. *BMC Genomics*, 16(1): 601.

Liu XH, 2021. Analysis on SNP and InDel markers based on RNA-seq technology in maize. *Mol. Plant Bre.*, 19(4): 1182–1189. [刘小红, 2021. 基于 RNA-seq 技术的玉米 SNP 和 InDel 标记分析. 分子植物育种, 19(4): 1182–1189.]

Lu BR, Cai X, Jin X, 2009. Efficient Indica and Japonica rice identification based on the InDel molecular method: Its implication in rice breeding and evolutionary research. *Progress in Natural Science*, 19(10): 1241–1252.

Martín-Hernández R, Bartolomé C, Chejanovsky N, Le Conte Y, Dalmon A, Dussaubat C, García-Palencia P, Meana A, Pinto MA, Soroker V, Higes M, 2018. *Nosema ceranae* in *Apis mellifera*: A

12 years postdetection perspective. *Environ. Microbiol.* , 20(4): 1302–1329.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA, 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9): 1297–1303.

Traver BE, Fell RD, 2012. Low natural levels of *Nosema ceranae* in *Apis mellifera* queens. *Journal of Lnvertebrate Pathology*, 110 (3): 408–410.

Väli U, Brandström M, Johansson M, Ellegren H, 2008. Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. *BMC Genet.* , 9: 8.

Wajnberg G, Passetti F, 2016. Using high-throughput sequencing transcriptome data for InDel detection: Challenges for cancer drug discovery. *Expert Opinion Drug Discovery*, 11(3): 257–268.

Wang Y, Jing L, Li XC, 2018. SNP mining and gene functional annotation in *Puccinia helianthi* transcriptome group. *Acta Agric. Boreali-Sin.* , 33(3): 92–98. [王妍, 景岚, 李鑫淳, 2018. 向日葵锈菌转录组 SNP 位点挖掘及所在基因功能注释. 华北农学报, 33(3): 92–98.]

Williams TA, Nakjang S, Campbell SE, Freeman MA, Eydal M, Moore K, Hirt RP, Embley TM, Williams BA, 2016. A recent whole-genome duplication divides populations of a globally distributed microsporidian. *Mol. Biol. Evol.* , 33(8): 2002–2015.

Wu Y, Cai ZB, Xu YJ, Guo YL, Bao JY, Kang YX, Ye YP, Qian JJ, Zhang KY, Chen DF, Guo R, 2022. Identification of single nucleotide polymorphism (SNP) and insertion-deletion (InDel) mutation loci in *Apis cerana cerana*. *Chinese Journal of Applied Entomology*, 59(3): 608–617. [吴鹰, 蔡宗兵, 许雅静, 郭意龙, 鲍佳益, 康育欣, 叶亚萍, 钱加珺, 张凯遥, 陈大福, 郭睿, 2022. 中华蜜蜂的单核苷酸多态性和插入缺失突变位点鉴定. 应用昆虫学报, 59(3): 608–617.]

Xiong CL, Tong XY, Chen HZ, Geng SH, Zhuang TY, Zheng YZ, Fu ZM, Chen DF, Zhao HX, Guo R, 2019. Optimization of gene structure and identification of novel genes in *Nosema ceranae*. *Journal of Environmental Entomology*, 41(2): 148–154. [熊翠玲, 童新宇, 陈华枝, 耿四海, 庄天艺, 郑燕珍, 付中民, 陈大福, 赵红霞, 郭睿, 2019. 东方蜜蜂微孢子虫的基因结构优化及新基因鉴定. 环境昆虫学报, 41(2): 148–154.]

Yang J, He J, Wang DB, Shi E, Yang WY, Geng QF, Wang ZS, 2016. Progressinresearch and application of InDel markers. *Biodiversity Science*, 24(2): 237–243. [杨洁, 赫佳, 王丹碧, 施恩, 杨文宇, 耿其芳, 王中生, 2016. InDel 标记的研究和应用进展. 生物多样性, 24(2): 237–243.]

Zhang JH, Dai TS, Zhou S, Wang CQ, Ma W, 2021. SNP/InDel analysis of pituitary and hypothalamus of Dagu chicken based on RNA-seq technology. *Biotic. Resources*, 43(2): 160–165. [张洁慧, 戴天姝, 邹爽, 王春强, 马巍, 2021. 基于 RNA-seq 技术对大骨鸡垂体及下丘脑组织的 SNP/InDel 分析. 生物资源, 43(2): 160–165.]

Zhao H, Li QZ, Li J, Ceng ZQ, Hu SN, Yu J, 2006. Research on the transition or transversion of adjacent base components and SNP produced in plant genome. *Sciencein China Series C:Life Sciences*, 36(1): 1–8. [赵辉, 李启寨, 李俊, 曾长青, 胡松年, 于军, 2006. 相邻碱基组分与产生 SNP 的转换或颠换在植物基因组中的研究. 中国科学 C 辑: 生命科学, 36(1): 1–8.]

Zhao S, Wang Z, Bai YY, Guo K, Li ZH, Zhangsun WG, Qu L, Lan XY, Song XY, 2020. Correlation between InDel of cashmere goat *MARCH*1 gene and body size traits. *Journal of Domestic Animal Ecology*, 41(3): 27–32. [赵森, 王真, 白洋洋, 郭凯, 李增辉, 长孙伟光, 屈雷, 蓝贤勇, 宋晓越, 2020. 绒山羊 *MARCH*1 基因 InDel 位点与体尺性状的关联研究. 家畜生态学报, 41(3): 27–32.]